EXPERT REVIEWS

Puzzle of protein complexes *in vivo*: a present and future challenge for functional proteomics

Expert Rev. Proteomics 6(2), 159–169 (2009)

Maria Monti, Marianna Cozzolino, Flora Cozzolino, Giuseppina Vitiello, Roberta Tedesco, Angela Flagiello and Piero Pucci†

†Author for correspondence
Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
Tel.: +39 081 373 7896; +39 081 674 318
Fax: +39 081 373 7808; +39 081 674 313
pucci@unina.it

Complete description of the complex network of cellular mechanisms and use of the network to predict the full range of cellular behaviors are major goals of systems biology. A key role in contemporary biology can be played by functional proteomics, which focuses on the elucidation of protein functions and the definition of cellular mechanisms at the molecular level. The attainment of these targets is strictly dependent on the identification of individual proteins within functional complexes *in vivo*. Isolation of interacting proteins relies on either affinity-based or immunoprecipitation procedures in which the protein bait and its specific partners can be fished out by their specific binding to ligand molecules immobilized on insoluble supports. These approaches led to the final identification of several proteins belonging to distinct complexes endowed with different biological functions. Assignment of each protein to a specific complex constitutes a tremendous problem that can only be partially solved using protein–protein interaction databases and literature information. The development of prefractionation methodologies to separate individual protein complexes while preserving their native interactions might then represent an essential tool for the future of functional proteomics. Prepurification of single complexes can only be pursued under native conditions on the basis of their physicochemical features, such as size, dimension (gel filtration chromatography) and density (gradient ultracentrifugation). Following prefractionation, the complex associated to a specific biological function can be isolated using affinity purification techniques. Functional proteomics approaches able to describe individual proteins belonging to complexes involved in specific cellular functions will have a terrific impact on future systems biology studies.

Keywords: affinity purification • functional proteomics • gel filtration • immunoprecipitation • prefractionation strategies • systems biology • ultracentrifugation

The ability of any organism to survive depends on several complex cellular mechanisms that fulfill all the functions needed for cell life. Moreover, cells must be able to modify their patterns of gene expression in response to extra- and intracellular signals to ensure that the correct amount of the appropriate subset of genes is expressed at the proper time [1,2]. In the past, cellular mechanisms and signal transduction were thought to occur as a set of isolated processes or linear pathways where biological functions took place one after the other in a time-dependent fashion. Nowadays, however, it is well established that fundamental biological mechanisms are pursued by a multitude of proteins gathering together to transiently form large functional complexes [3] or cell pathways [4]. Having performed their functions, these complexes dissociate, making the individual components free to associate with other complexes and carry out distinct activities. To further complicate the picture, several mechanisms and many signaling pathways can interact with each other, thus forming an intricate network that integrates both extra- and intra-cellular signals. These dynamic processes constitute the molecular networks through which signaling information flows within the cell [5,6].

Complete description of this complex signal transcription and transduction network, and use of the network to predict the full range of cellular behaviors, are major goals of systems biology [7,8]. A key role in contemporary and future

 10.1586/EPR.09.7 ISSN 1478-9450

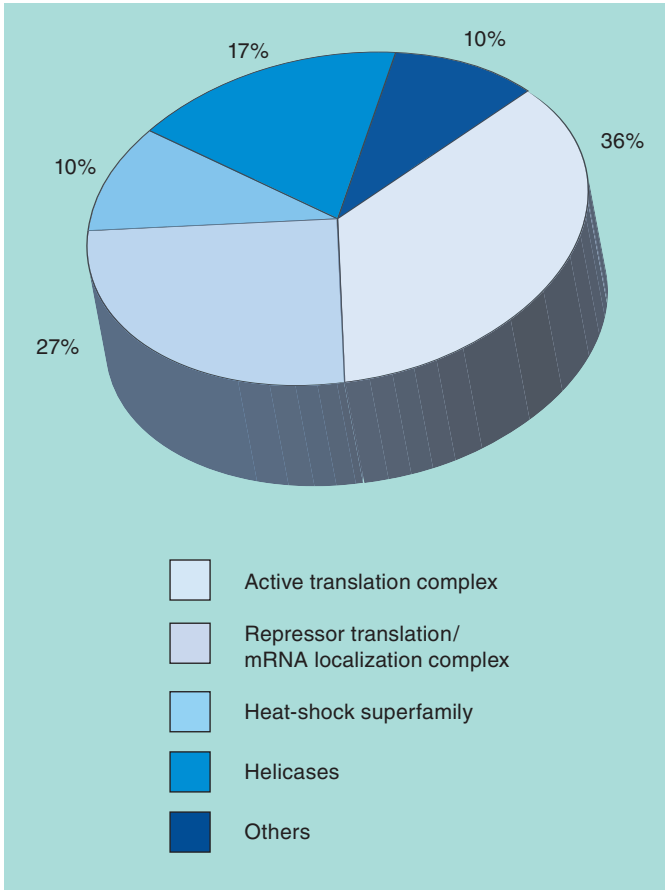

Figure 1. Assignment of the affinity-purified eukaryotic translation initiation factor 4E interactors to five different functional categories. These were the active translational complex, repressor translational/mRNA localization complex, heat-shock protein superfamily, helicases and other proteins. These findings confirmed that a single protein molecule may simultaneously belong to different functional complexes.

biological sciences can be played by functional proteomics, an emerging area in proteomics research addressing the elucidation of protein functions and the definition of cellular mechanisms at the molecular level [9,10]. A comprehensive description of cellular processes at the molecular level is in fact strictly dependent on the clear definition of the individual protein components involved in these functional entities. The association of an unknown protein with partners belonging to a specific protein complex involved in a particular mechanism would be strongly suggestive of its biological function [11]. Furthermore, a detailed description of the cellular signaling pathways might greatly benefit from the elucidation of protein–protein interactions *in vivo*.

Isolation of entire multiprotein complexes can be accomplished by affinity-based approaches (including 'pull-down' and 'tap-tag' experiments) in which a specific ligand molecule is used to bind the protein of interest and its protein partners [12–14]. Alternatively, immunoprecipitation strategies aimed at either the endogenous protein [15] or a tagged version of the bait can be used, taking advantage of the availability of several antitag systems endowed with high binding efficiency [16]. It should be mentioned that, in both cases, the design of good negative-control experiments is of the utmost importance due to the large number of possible false-positives. Finally, protein components specifically bound to the bait are eluted, fractionated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and identified by liquid chromatography (LC)-tandem mass spectrometry (MS/MS) techniques [17,18].

These approaches usually result in the final identification of a very long list of interactors from a single functional proteomics experiment [19,20]. It is now clear, in fact, that a single protein might be involved in different complexes, each endowed with its own specific biological function. Therefore, the attempt to rationalize the list of interactors within a biologically significant picture constitutes, at present, a major problem in functional proteomics. Even the simple verification of the entire set of interacting proteins by different experimental procedures (e.g., coimmunoprecipitation) is often impossible due to the extremely large number of candidates. Two approaches are normally used to manage these lists: the simplest strategy involves the choice of one or a few interactors among the various protein candidates on the basis of previous data or literature information. The interaction of these proteins with the bait can then be confirmed and functionally investigated [21–23]. A second more challenging and time-consuming approach aims to assign each protein to a specific functional complex, a task analogous to solving various '100-piece' puzzles at the same time and with the same pieces [24]. As stated earlier, in fact, each protein molecule may simultaneously belong to different functional complexes. This tremendous problem can only be partially solved using protein–protein interaction databases and literature information but, when successful, the results may be suggestive of new, undisclosed biological functions of the protein bait [25]. The final goal of this second approach would be the achievement of a global functional overview of the protein bait interactome to be inserted into appropriate systems biology studies [26,27].

Despite the continuous growth of protein-interaction databases and software tools [28,29], the association of individual proteins to a specific protein complex still constitutes the 'bottleneck' of any functional proteomics approach. The development of different upstream strategies, based on prefractionation of protein extracts into separate individual protein complexes while preserving their native interactions, might then represent an essential tool for the future of functional proteomics. A brief overview of the main approaches presently used in functional proteomics studies, with particular emphasis on prefractionation methodologies using some examples from our laboratory, is described in the following paragraphs.

Affinity-based approaches

The identification of interacting proteins in stable or even transient complexes by affinity-based procedures takes advantage of the intrinsic affinity of a specific ligand for the protein of interest. The ligand molecule is covalently bound to an insoluble support (usually agarose/sepharose beads) and used as a bait to selectively

bind the protein of interest and its partners. Theoretically, there are no limitations to the choice of ligand molecules that can be employed in the pull-down experiments. Several examples have been reported using modified substrates and/or inhibitors of selected proteins with specific enzymatic activity [30–32]. In other cases, native and/or modified drugs addressed towards specific protein targets have also been used [33]. In this respect, a particularly intriguing application deriving from this strategy is the possibility of identifying new, previously unknown targets for a specific drug.

This strategy was applied to the identification of the protein partners of the eukaryotic translation initiation factor 4E (eIF4E) in an attempt to investigate the translation initiation complex in *Drosophila melanogaster* during oogenesis. Since eIF4E binds the 5′-cap structure of eukaryotic mRNAs [34–36], appropriate affinity experiments were designed using 7-methylguanosine (m7GTP)-conjugated agarose beads where the m7GTP mimics the mRNA cap moiety [22]. The whole protein extract from adult ovaries was then incubated onto the m7GTP sepharose beads and eIF4E and its interacting partners were selectively captured, leading to the identification of approximately 30 protein candidates. Among these, several previously known translation factors could be detected, together with a number of new interactors not previously described [22]. According to their reported biological activities, the identified interactors could be grouped into different functional categories (Figure 1), suggesting that eIF4E might be involved in different functional complexes.

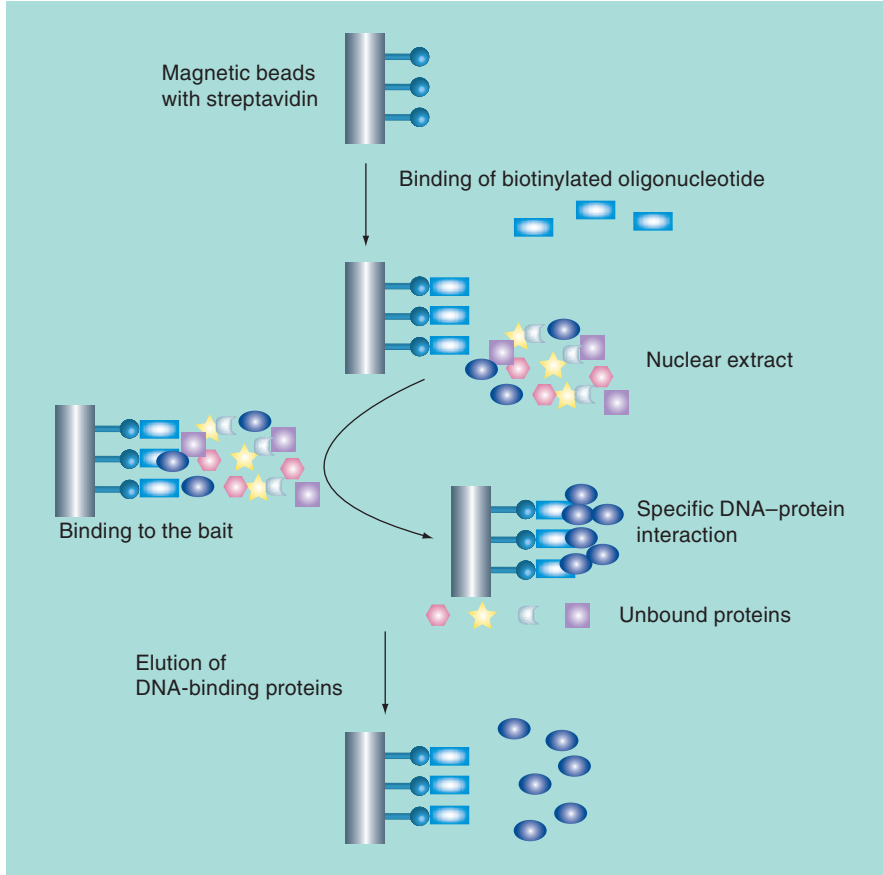

Figure 2. Affinity-based procedure for the isolation of DNA-binding protein complexes using a biotinylated oligonucleotide as bait. The biotinylated oligonucleotide is immobilized on the agarose beads by strepatavidin binding. The nuclear protein extract is then incubated with the bait that establishes noncovalent interactions with its specific partners. The unbound proteins are washed out, whereas the protein components specifically recognized by the bait are retained on the agarose beads. The specific DNA interactors are then eluted, fractionated by sodium dodecyl sulfate polyacrylamide gel electrophoresis, digested and identified by liquid chromatography-tandem mass spectrometry techniques.

In a more widespread variant of the fishing strategy, the protein of interest (bait) can be expressed with a suitable tag, such as glutathione *S*-transferase or a poly-His tail, for example, and immobilized on the insoluble support, taking advantage of the availability of several antitag systems endowed with high binding efficiency. The immobilized bait can then be used to fish its specific partners out from an entire cellular extract [16]. These affinity-tag systems provide a general applicability with a large number of proteins and a minimal effect on the tertiary structure and the biological activity of the bait, preventing complex instability.

The success of an affinity-based approach depends on the absence of excessive aspecific interactions, which, in turn, is related to the specificity of the bait-partners' recognition. It is not surprising then that the fishing strategy has found considerable application in the isolation of nucleic acid-binding proteins, where the binding specificity is extremely high, and in studies on transcription and/or translation complexes or the splicing machinery. Either synthetic oligonucleotides encompassing the specific base pair sequences or PCR-amplified DNA or RNA fragments can be covalently immobilized on commercially available inert beads and employed for the isolation of proteins involved in the specific binding of the nucleotide sequence [37–39]. A biotin tag can also be added to synthetic oligonucleotides; immobilization of the bait on the support will then take advantage of its high affinity for streptavidin molecules (Figure 2) [40]. Very recently, a modified version of the chromatin immunoprecipitation technique (mChIP) has been proposed for the discovery of chromatin-associated protein networks. mChIP consists of a single affinity-purification step, whereby chromatin-bound protein networks are isolated from mildly sonicated and gently clarified cellular extracts using magnetic beads coated with antibodies [41].

Protein components specifically recognized by the bait can then be eluted and fractionated by SDS-PAGE. Protein bands are in-gel digested and the resulting peptide mixtures analyzed by LC-MS/MS techniques, leading to the identification of the

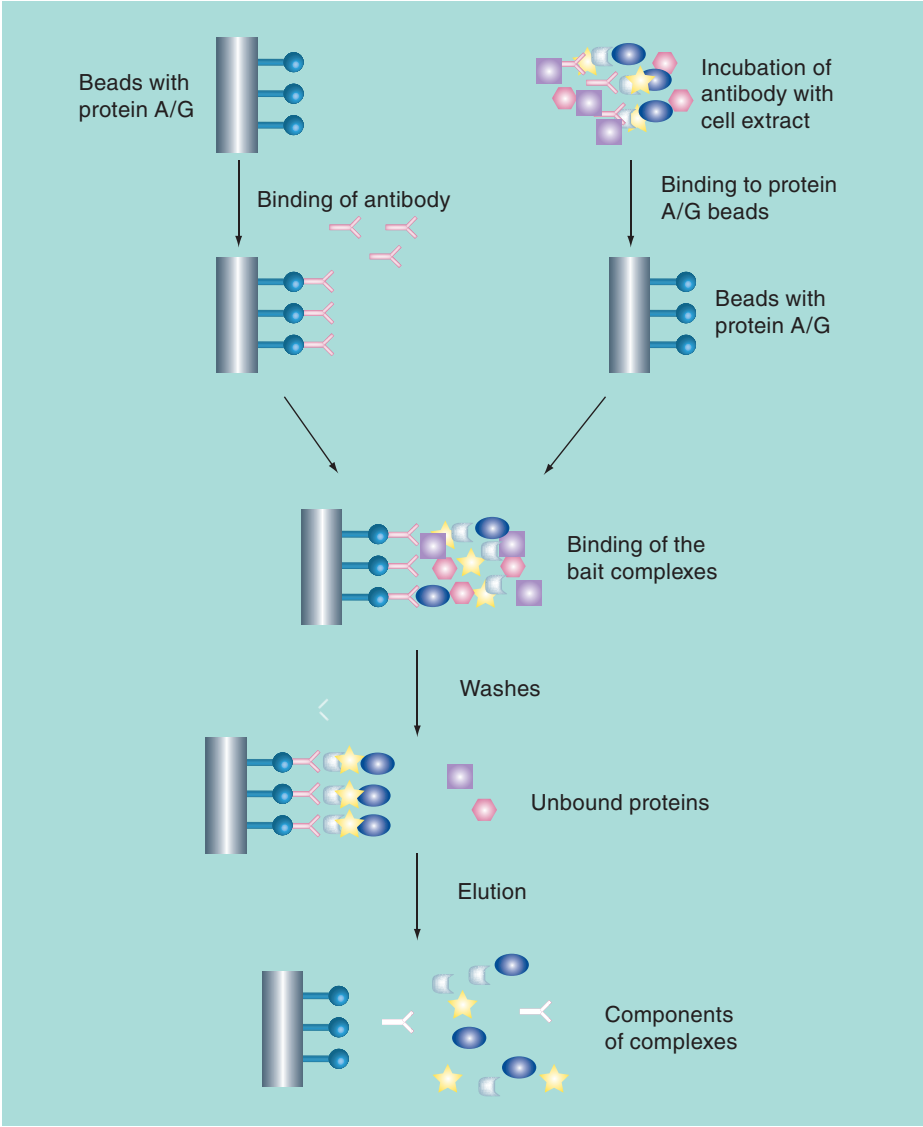

Figure 3. Endogenous immunoprecipitation strategy. The two possible approaches are outlined. In the first method, the antibait antibody is coupled to protein A/G-derivatized sepharose beads (left side). Alternatively, the protein bait is immunoprecipitated with its specific antibody and the entire complexes are then rescued by protein A/G-derivatized sepharose beads (right side). In both cases, nonspecific proteins are washed out and the bait interactors are eluted and identified.

proteins by database search [42]. This experimental sequence, SDS-PAGE/in-gel digestion/LC-MS/MS, is usually preferred to other existing methods because SDS-PAGE, although endowed with poor resolution, increases the recovery of low-solubility proteins and it is sensitive enough to allow analysis of the small amount of protein complexes isolated by affinity and/or immunoprecipitation (see later) procedures. Hence, since protein bands from SDS-PAGE often contain several proteins, LC-MS/MS methodologies, being able to handle very complex peptide mixtures, are routinely employed for protein identification.

Gavin *et al.* [26] and Krogan *et al.* [27] used a slightly different version of the affinity purification strategy known as tandem affinity purification (TAP)-tag for the isolation of a huge amount of interacting proteins in yeast. This method combines two different tags on the same protein, usually spaced by an enzyme-cleavable linker sequence. The original TAP-tag system consisted of two IgG-binding domains of *Staphylococcus aureus* protein A (ProtA) and a calmodulin-binding peptide, separated by a tobacco etch virus protease cleavage site. The recombinant vector replaces the endogenous wild-type gene, thus avoiding overexpression of the protein bait and allowing the recovery of protein complexes expressed at their own natural level.

Although widely used, the complex affinity capture suffers from a huge number of drawbacks. Besides a number of technical problems, during preparation of the cellular lysate, the architecture of the subcellular compartments is disrupted and proteins that are normally segregated in different organelles can come into contact, generating specific but nonphysiological interactions. More importantly, in these experiments, all the interactions take place on the derivatized beads in a noncellular environment where several conditions, such as the molar ratio among the bait and its protein partners, are largely altered.

Immunoprecipitation strategies

Several strategies relying on immunoprecipitation techniques have been introduced for protein complex isolation in order to overcome most of the drawbacks affecting the affinity-based approaches [43,44]. These procedures essentially fall into two categories: direct immunoprecipitation of the endogenous protein bait from cellular extracts or *in situ* production of a tagged form of the protein bait followed by immunoprecipitation with antitag antibodies.

The classical approach to endogenous immunoprecipitation consists of the coupling of the antibody raised against the protein bait onto protein A- (or G)-derivatized sepharose beads that can be used to isolate the target protein and its native complexes. Alternatively, the protein bait is immunoprecipitated from the whole protein extract using a soluble form of the antibody and protein A- (or G)-derivatized sepharose beads are then added to the sample. In both cases, after several washes, the bait interactors can be eluted from the beads in denaturing conditions and the proteins identified, as reported earlier (Figure 3).

When no antibodies suitable for this procedure are available, the target protein is expressed with a small peptide epitome against which good antibodies exist (e.g., FLAG, V5, c-myc and hemagglutinin) located at either the N- or C-terminus, where a minimal effect on the tertiary structure and biological activity of the protein is expected. The tagged protein is then transfected into the appropriate cell line and immunoprecipitated using commercially available antitag antibodies covalently conjugated to agarose beads [16]. Following the removal of unspecific proteins, the interactors can be eluted by competition with a relevant concentration of the free form of the tag peptide.

Endogenous immunoprecipitation should always be preferred whenever the protein bait is expressed enough to allow a relatively large amount of complexes to be isolated. However, this second approach is certainly endowed with a number of advantages, including higher immunoprecipitation yields and low contamination of the sample, decreasing the possible occurrence of false-positives. Because the antibody is covalently bound to the beads, elution of the specific interactors can be performed in mild conditions; moreover, the final sample is not contaminated by the antibody chains, thus increasing protein identification. Overexpression of the tagged protein in the host cells should definitely be avoided since a high concentration of the bait alters the stoichiometric ratio with its natural partners, often leading to the formation of nonspecific and/or nonphysiological interactions.

The strategy of endogenous immunoprecipitation was applied to the isolation of functional complexes involving the eIF4E protein in *Drosophila* to evaluate possible differences with the affinity-based approach. Endogenous eIF4E-containing complexes were immunoprecipitated using a specific antibody and protein G-conjugated sepharose beads according to the procedure described earlier. Proteins specifically retained on the beads were eluted with SDS and fractionated by SDS-PAGE. The resulting coomassie-stained gel is shown in Figure 4 where the gel from the pull-down experiment is also reported for comparison. Although the initial amount of protein extract was the same in both experiments, the affinity-based purification clearly led to the isolation of higher numbers of protein bands than the immunoprecipitation approach. These data were also confirmed by protein identification; only 14 interactors could in fact be identified in the immunoprecipitation experiment.

This result could be explained by considering the higher yield of affinity purification compared with the immunoprecipitation procedure, which is strictly dependent on the performance of the antibody. Moreover, binding of the antibody to specific epitopes in the eIF4E native structure might compete with interacting proteins binding to the same regions, thus leading to dissociation of some protein–protein interactions [44].

In the affinity procedure, the selective capture of the eIF4E complexes was accomplished through the binding to m7GTP, a ligand involved in a specific process: the translation of mRNAs. In the immunoprecipitation strategy, the antibody recognized all the eIF4E-containing complexes, even if they are involved in different biological functions. Consistently, despite the lower number of interactors, the 14 identified proteins were assigned to distinct functional

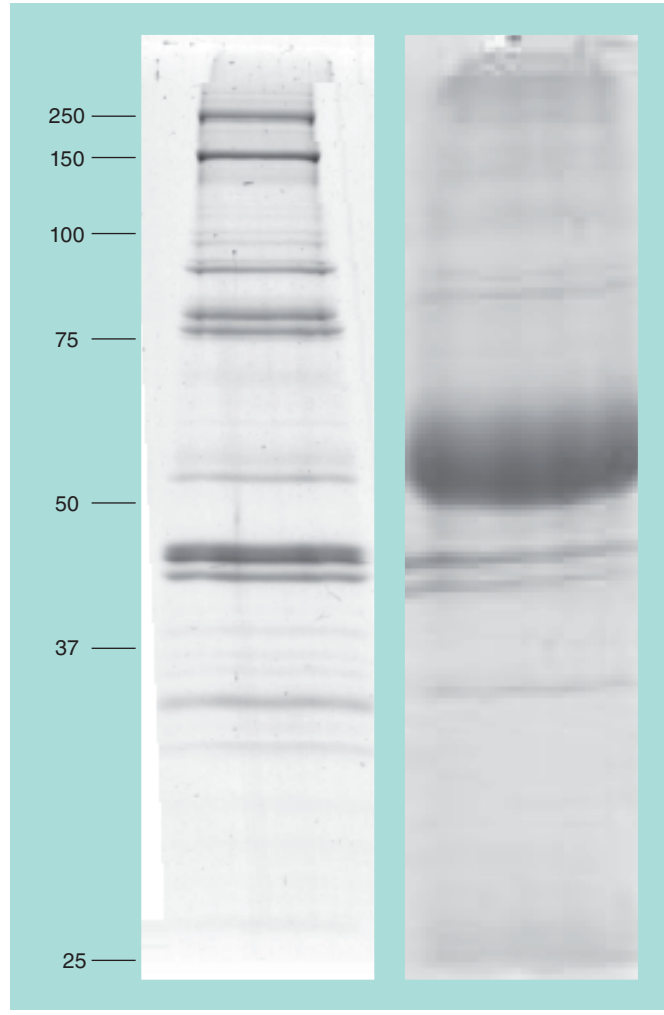

Figure 4. Coomassie-stained gels from the pull-down (left) and the immunoprecipitation (right) experiments for the isolation of eukaryotic translation initiation factor 4E interactors. Although an equal amount of protein extract was loaded onto the gels, several more protein bands could be detected in the affinity-based purification procedure than the immunoprecipitation approach. Only 14 interactors could be identified by this latter experiment compared with more than 30 proteins identifed by the affinity procedure. The large band occurring slightly higher than 50 kDa in the immunoprecipitation experiment corresponds to the antibody heavy chain contaminating the sample.

complexes on the basis of their reported biological activity (Figure 5). This might open up the way to understanding the novel roles of the eIF4E protein different from that exerted in protein synthesis.

When high-throughput affinity and immunoprecipitation experiments are compared, a scarce coincidence in the results may be observed. A poor overlap between yeast protein complexes was similarly found in two studies that used the tap-tag approach and the FLAG immunoprecipitation procedure followed by MS identification of the isolated components [45,46]. Although surprising, these results are not without precedent. Both using the tap-tag approach, the studies by Gavin *et al.* [26] and Krogan *et al.* [27] were able to successfully purify and identify more than 3000 different proteins

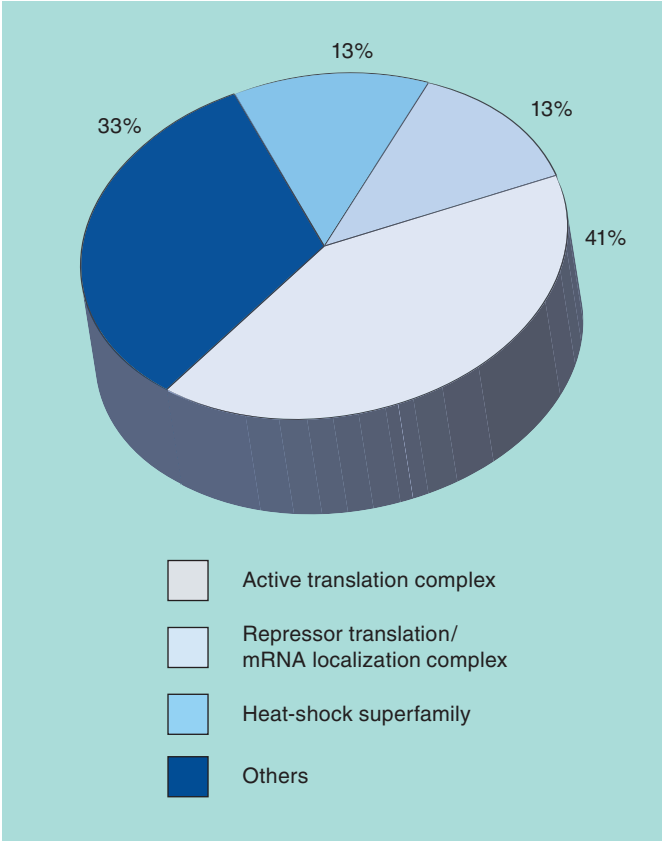

Figure 5. Assignment of eukaryotic translation initiation factor 4E interactors isolated by immunoprecipitation to different functional categories. Besides the active and repressive translational complexes, these findings suggested that the protein bait may be involved in other still unknown processes, supporting the idea that a single protein might be involved in different biological functions.

in total (i.e., 47% of the *Saccharomyces cerevisiae* proteome) [47]. However, when the data were compared, only 1000 proteins were found to be common to both datasets. In addition, only six complexes were identical between the two studies. Almost 200 of the complexes reported by Gavin and coworkers [26] did not share any protein with the complexes found by Krogan *et al.* [27]. Besides the effects of different bioinformatics tools and possible dissimilarities in the purification of the tap-tagged proteins, these disconcerting results might reflect the dynamics of cellular processes. Proteins transiently associate in functional complexes to execute a specific process and then rapidly dissociate, and each of them is free to form other complexes involved in different mechanisms.

Prefractionation of protein complexes

Whether carried out by affinity-based or immunoprecipitation procedures, a single functional proteomics experiment might lead to a long list of putative interactors that clearly indicate the occurrence of the protein bait in different complexes. As in other proteomic applications, the most challenging and time-consuming task in these experiments is the interpretation of the data and, particularly, the attempt to assign all the protein components to the specific complex they belong to. Despite the wealth of information provided by protein databanks and protein–protein interactions databases, and the availability of various software tools, the inference of a biological significance from the list of interactors still remains an elusive problem.

A possible solution to overcome these difficulties might be brought in by classical protein purification methodologies. Prefractionation of the total protein extract by methods able to preserve native protein interactions might be instrumental to isolating individual functional complexes, thus greatly decreasing the complexity of the functional proteomics data. Prepurification of individual complexes under native conditions can be achieved on the basis of their physicochemical features, including size and dimension (gel filtration chromatography) [48–50], density (gradient ultracentrifugation) [51–53] and electrophoretic mobility (Blue Native electrophoresis) [54–56]. These methods cannot include reverse-phase chromatography or other denaturing procedures to avoid complex dissociation. Following prepurification of the sample, the various complexes containing the protein bait are separated in different fractions and the occurrence of the protein bait within these fractions can simply be detected by western blot analysis. The positive immunostained fractions are selected while the others are discarded, and the complexes are then individually isolated using either the immunoprecipitation or the affinity techniques described earlier [48]. It should be underlined that loose or weaker interactors may become lost during the extra prefractionation step; on the other hand, these interactors would also dissociate during the stringent washing conditions needed to decrease unspecific binding in the absence of prefractionation. In some cases, cross-linking approaches can also be used to avoid the dissociation of weak interactors.

Decreasing the complexity of a cellular extract by prefractionation procedures results in the identification of a much lower number of putative interactors for each fraction. Moreover, all the components found in the same fraction are most likely to belong to the same functional complex, making the interpretation of the data much simpler. Finally, the removal of a huge amount of proteins not related with the bait positively affects the dynamic range of the method, thus leading to a great decrease in the background and to a higher overall confidence in identifications. The false-positives issue is so critical in proteomics experiments that a few years ago Matthias Mann and coworkers developed a smart strategy to discriminate specific and nonspecific interactors isolated during protein immunoprecipitation experiments. This strategy, termed quantitative immunoprecipitation combined with knockdown (QUICK) [57], combines stable-isotope labeling amino acids (SILACs) with immunoprecipitation and RNAi.

Gradient ultracentrifugation

Gradient ultracentrifugation, for example using sucrose, glycerol or CsCl, has found its principal application in subcellular organelle separation [58,59]. Vogelmann *et al.* separated membrane compartments using a linear 10–20–30% iodixanol density gradient [52]. Only a few examples of the application

of these methods to protein complexe separation have been reported to date [51,52]. In the standard procedure, the protein extract is loaded onto a suitable density gradient and subjected to ultracentrifugation for a fixed time. The recovered fractions are then tested for the presence of the protein bait, usually by western blot using the specific antibody for the protein of interest. However, protein identification by enzymatic activity or by using specific biological assays has also been reported [52].

Gradient ultracentrifugation was applied to the prefractionation of eIF4E complexes and the data were compared with previous experiments. The total protein extract was loaded onto a 10–50% glycerol gradient and, after ultracentrifugation, ten fractions were collected. Western blot analysis of the fractions with anti-eIF4E antibody showed that eIF4E had been spread in almost all the fractions, suggesting its involvement in complexes that differ in density and size. The gradient fractions were also tested for the presence of Cup, an already known eIF4E interactor with inhibitory activity, using a specific anti-Cup antibody. The western blot analysis clearly showed the coexistence of both proteins in a few fractions with medium–high density. The fractions containing eIF4E and Cup, as well as those at lower glycerol density containing only eIF4E, were collected and separately subjected to affinity purification on m7GTP sepharose beads, as described earlier.

The proteins eluted from each affinity step were fractionated by SDS-PAGE and identified by mass spectrometry. As expected, few interactors were identified in both fractions, making the interpretation of the data relatively easy. The eIF4E-interacting proteins occurring in the medium–high density fraction were associated with the translational complex. The contemporary presence of Cup and eIF4E clearly indicated that this complex is involved in the inhibition of translation [60–62]. Moreover, a number of proteins never associated to this process or reported to bind to eIF4E or Cup were also identified.

Proteins identified in the low-density fraction and particularly the presence of eIF4G, a well-known translational activator, indicated the occurrence of the active translation complex. Even in this case, a number of novel proteins not previously associated with this specific biological function could also be detected.

A number of general considerations can be drawn by these experiments. First, the prefractionation procedure successfully resulted in the separation of at least two protein complexes, both containing the protein bait and each endowed with its own biological function. The presence of discriminant proteins, such as the mutually exclusive Cup and eIF4G, led to an easy identification of the repressor and the activation complex [63]. A second observation is that a consistent number of components were found in both pools of interacting proteins. This might be an obvious consequence of contamination of the complexes isolated in adjacent fractions due to the low resolution of ultracentrifugation. However, these data might also suggest that the two complexes consist of a common protein core that is addressed toward different processes by the association of a small number of different proteins.

Size-exclusion chromatography

Recently, Holaska and Wilson reported the isolation of several emerin complexes from HeLa nuclei in three subsequent chromatographic steps [64]. The emerin-containing complexes were first enriched by immunoprecipitation using anti-emerin antibodies and then fractionated by ion-exchange and size-exclusion chromatography. Besides a number of emerin-containing complexes, this latter analysis suggested the further presence of at least two complexes: one at high molecular weight that included emerin, nuclear myosin I, actin and RII-spectrin, and a smaller complex containing emerin, nuclear myosin I and actin [64]. In this approach, the immunoprecipitation step was performed upstream of the chromatographic fractionation of the complexes. This was possible because the emerin-containing complexes were eluted from the antibody by competition with the recombinant protein. This mild elution procedure allowed the protein complexes to be recovered in quasi-native conditions, preserving the most stable protein–protein interactions. Identification of the protein components within the emerin complexes provided support to the existence of a central protein core common to several different complexes to which discriminant proteins were linked.

In a slightly different approach, size-exclusion chromatography can be used as a preliminary purification procedure to fractionate the total protein extract prior to the immunoprecipitation step [48,49]. The fractions containing the protein of interest can then be further subjected to immunoprecipitation for the isolation of the specific bait-containing complexes [48]. This procedure was employed for the fractionation of the *Drosophila* ovary extracts before purification of the eIF4E-containing complexes. Fractions from the column were analyzed by western blot and showed a widespread distribution of eIF4E in almost all the fractions, in agreement with ultracentrifugation results. However, at least two peaks with different retention times were clearly detected in the profile, corresponding to fractions with high and low molecular mass, respectively, suggesting the involvement of eIF4E in complexes of different sizes. When the same fractions were tested with the anti-Cup antibody, the translational inhibitor clearly localized within the highest-molecular-weight complex. Consistently, affinity purification of this complex on 7mGTP sepharose beads revealed the presence of the repressive translational complex.

Fractions corresponding to the low-molecular-mass complex were also incubated with the 7mGTP sepharose beads and the retained proteins were identified. Surprisingly, neither Cup nor eIF4G were found in this pool of interactors. Moreover, the vast majority of these proteins were novel eIF4E interactors never before associated with the translation process. These findings strongly suggested the involvement of eIF4E within a further functional complex whose biological activity is still unknown.

Isolation of protein complexes by the chromatographic procedure showed many advantages in comparison with the ultracentrifugation experiment. Sample loading, fractionation of the complexes and fraction recovery is quite easy and less time-consuming than in the ultracentrifugation procedure, ensuring a higher preservation of native protein–protein interactions.

Moreover, gel filtration showed a higher resolution in the separation of native protein complexes that allowed the detection of low-molecular-mass complexes not observed in the ultracentrifugation experiment, very likely because they were excluded by the glycerol gradient.

Expert commentary

Protein assembly to form functional complexes is a dynamic process that takes place at the appropriate time. The transient complex dissociates soon after the biological mechanism has been completed and individual components can participate in the formation of other complexes driven by specific signals. Moreover, cellular processes occur at very defined sites within the cell and functional complexes assemble at the proper time and in the proper place to correctly fulfill a specific function. Functional proteomics experiments aimed at describing the network of cellular processes should then carefully consider the time distribution and intracellular localization of the specific mechanism to be investigated.

Modern biological sciences are experiencing a paradoxical situation in which, very often, the protein sequence, the corresponding coding gene and even the regulatory mechanisms related to gene expression are known but the biological role of the protein is completely unclear. This apparent paradox suggests that there is a need to change our approach to the comprehension of cell functioning. We must understand that a protein's activity and its biological role may be two very different aspects. The activity of a protein describes whether the protein itself is, for example, an enzyme, a transporter, a receptor, a chaperone or a scaffold, but this activity may be exerted in completely different biological processes according to the cell's needs. The biological function of a protein is then related to the specific process that the protein is actually participating in. Consequently, since the same protein can assemble with different partners to form different functional complexes, a protein may have a single activity but several biological functions.

Five-year view

It is relatively easy to predict that functional proteomics will play a key role in future systems biology studies. An accurate description of protein components within functional complexes is a fundamental prerequisite to understanding cell functioning and to unravelling the network of signaling pathways at the molecular level. Interpretation of proteomics data constitutes, at present, a major problem, impairing the association of each identified protein to a specific complex, thus greatly diminishing the impact of these studies. The development of sensitive prefractionation methods endowed with adequate resolution to separate individual protein complexes will decrease sample complexity and protein background, making the interpretation of the data relatively straightforward. Specific experiments can then be designed to give functional meanings to proteomic results. There is much room for creativity in connecting cell and molecular biological strategies with the powerful proteomics approach to solve questions that could not previously be addressed.

A further goal to be pursued in the future will be the attempt to perform functional proteomic experiments in real *in vivo* systems by generating animal models bearing a tagged version of the protein bait. Homozygous animals, when vital, will provide tissues and/or progenitor cells for the isolation of functional complexes occurring *in vivo*, thus disclosing the identity of their individual components and addressing the question of whether their composition may vary in different districts.

Financial & competing interests disclosure

This work was supported in part by grants from the Italian Ministry of Education (MIUR), Prin 2007 and Progetto FIRB RBRN07BMCT, 'Italian Human ProteomeNet'. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Fundamental biological mechanisms are pursued by a multitude of proteins gathering together to transiently form large functional complexes. A single protein may be involved in different complexes, each endowed with its own specific biological function.
- The major goal of systems biology is the complete description of the molecular networks of dynamic processes through which signaling information flows within the cell.
- Functional proteomics is an emerging area in proteomics research that addresses the elucidation of protein functions and the definition of cellular mechanisms at the molecular level.
- Isolation of entire multiprotein complexes in functional proteomics approaches is performed either by affinity-based strategies or by immunoprecipitation procedures.
- The association of each identified protein to a specific protein complex constitutes, at present, a major problem in functional proteomics.
- The development of sensitive prefractionation methods to separate individual protein complexes while preserving their native interactions may represent an essential tool for the future of functional proteomics.
- Cellular processes occur at very defined sites within the cell and functional complexes assemble at the appropriate time and in the appropriate place to correctly fulfill a specific function.
- The activity of a protein describes the intrinsic capacity of the molecule, whereas its biological function is related to the specific process where this activity is exerted according to the cell's needs.
- Since the same protein can assemble within different functional complexes, a protein may have a single activity but several biological functions.

References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest

1 Kobayashi Y, Weigel D. Move on up, it's time for change – mobile signals controlling photoperiod-dependent flowering. *Genes Dev.* 21, 2371–2384 (2007).

2 Qiu Z, Ghosh A. A calcium-dependent switch in a CREST–BRG1 complex regulates activity-dependent gene expression. *Neuron* 60, 775–787 (2008).

3 Michaud GA, Snyder M. Proteomic approaches for the global analysis of proteins. *Biotechniques* 33, 1308–1316 (2002).

4 Souchelnytskyi S. Proteomics in studies of signal transduction in epithelial cells. *J. Mammary Gland Biol. Neoplasia* 7, 359–371 (2002).

5 de la Fuente van Bentem S, Mentzen WI, de la Fuente A *et al.* Towards functional phosphoproteomics by mapping differential phosphorylation events in signalling networks. *Proteomics* 8, 4453–4465 (2008).

6 Lecaudey V, Cakan-Akdogan G, Norton WH *et al.* Dynamic FGF signaling couples morphogenesis and migration in the zebrafish lateral line primordium. *Development* 135, 2695–2705 (2008).

7 Pfeifer AC, Timmer J, Klingmüller U. Systems biology of JAK/STAT signalling. *Essays Biochem.* 45, 109–120 (2008).

8 Zak DE, Aderem A. Systems biology of innate immunity. *Immunol. Rev.* 227, 264–282 (2009).

9 Schlieker CD, Van der Veen AG, Damon JR *et al.* A functional proteomics approach links the ubiquitin-related modifier Urm1 to a tRNA modification pathway. *Proc. Natl Acad. Sci. USA* 105, 18255–18260 (2008).

10 Souza TA, Chen X, Guo Y *et al.* Proteomic identification and functional validation of activins and bone morphogenetic protein 11 as candidate novel muscle mass regulators. *Endocrinology* 22, 2689–2702 (2008).

11 Brown KA, Ham AJ, Clark CN *et al.* Identification of novel Smad2 and Smad3 associated proteins in response to TGF-β1. *J. Cell. Biochem.* 105, 596–611 (2008).

• **Discovery of novel Smad2- and/or Smad3-associated proteins reveals how these proteins are regulated and/or uncovers new functions of Smad2 and Smad3 in TGF-β1 signaling.**

12 Wright CW, Duckett CS. The aryl hydrocarbon nuclear translocator alters CD30-mediated NF-βB-dependent transcription. *Science* 323, 251–255 (2009).

•• **CD30, a TNF receptor family member, is upregulated in numerous lymphoid-derived neoplasias. An affinity purification strategy led to the identification of ARNT as a CD30-interacting protein, and functional assays revealed that ARNT operated in concert with RelB in a CD30-induced negative-feedback mechanism.**

13 Markillie LM, Lin CT, Adkins JN *et al.* Simple protein complex purification and identification method for high-throughput mapping of protein interaction networks. *J. Proteome Res.* 4, 268–274 (2005).

14 Wu F, Wang P, Young LC *et al.* Proteome-wide identification of novel binding partners to the oncogenic fusion gene protein, NPM-ALK, using tandem affinity purification and mass spectrometry. *Am. J. Pathol.* 174, 361–370 (2009).

15 Zada AA, Pulikkan JA, Bararia D *et al.* Proteomic discovery of Max as a novel interacting partner of C/EBPα: a Myc/Max/Mad link. *Leukemia* 20, 2137–2146 (2006).

16 Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 60, 523–533 (2003).

17 Meier N, Krpic S, Rodriguez P *et al.* Novel binding partners of Ldb1 are required for haematopoietic development. *Development* 133, 4913–4923 (2006).

18 Henkin JA, Jennings ME, Matthews DE *et al.* Mass processing – an improved technique for protein identification with mass spectrometry data. *J. Biomol. Tech.* 15, 230–237 (2004).

19 Orrù S, Aspesi A, Armiraglio M *et al.* Analysis of the ribosomal protein S19 interactome. *Mol. Cell Proteomics* 6, 382–393 (2007).

20 Scholten A, van Veen TA, Vos MA *et al.* Diversity of cAMP-dependent protein kinase isoforms and their anchoring proteins in mouse ventricular tissue. *J. Proteome Res.* 6, 1705–1717 (2007).

21 Fraldi A, Zito E, Annunziata F *et al.* Multistep, sequential control of the trafficking and function of the multiple sulfatase deficiency gene product, SUMF1 by PDI, ERGIC-53 and ERp44. *Hum. Mol. Genet.* 17, 2610–2621 (2008).

• **Functional proteomics approach revealed that SUMF1 interacts with PDI, ERp44 and ERGIC-53. Functional assays revealed that these interactions are crucial for controlling SUMF1 traffic and function.**

22 Pisa V, Cozzolino M, Gargiulo S *et al.* The molecular chaperone Hsp90 is a component of the cap-binding complex and interacts with the translational repressor Cup during *Drosophila* oogenesis. *Gene* 432, 67–74 (2009).

23 Bowen C, Stuart A, Ju JH *et al.* NKX3.1 homeodomain protein binds to topoisomerase I and enhances its activity. *Cancer Res.* 67, 455–464 (2007).

24 Köcher T, Superti-Furga G. Mass spectrometry–based functional proteomics: from molecular machines to protein networks. *Nat. Methods* 4, 807–815 (2007).

25 Lee PY, Bae KH, Kho CW *et al.* Interactome analysis of yeast glutathione peroxidase 3. *J. Microbiol. Biotechnol.* 18, 1364–1367 (2008).

26 Gavin AC, Aloy P, Grandi P *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631–636 (2006).

•• **Genome-wide screen for complexes in budding yeast, using affinity purification and mass spectrometry. This study provides the largest collection of physically determined eukaryotic cellular machines so far and a platform for biological data integration and modeling.**

27 Krogan NJ, Cagney G, Yu H *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637–643 (2006).

•• **Tandem affinity purification was used to process 4562 different tagged proteins for *Saccharomyces cerevisiae* interactome analysis. A Markov clustering algorithm organized the found interactions into 547 protein complexes. These data will help future studies on individual proteins as well as functional genomics and systems biology.**

28 Banks E, Nabieva E, Chazelle B *et al.* Organization of physical interactomes as uncovered by network schemas. *PLoS Comput. Biol.* 4, e1000203 (2008).

29 Zhang B, Park BH, Karpinets T *et al.* From pull-down data to protein interaction networks and complexes with biological relevance. *Bioinformatics* 24, 979–986 (2008).

30 Klein T, P Geurink P, S Overkleeft H, *et al.* Functional proteomics on zinc-dependent metalloproteinases using inhibitor probes. *ChemMedChem.* 4, 164-170 (2008).

31 Rix U, Hantschel O, Dürnberger G *et al.* Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. *Blood* 110, 4055–4063 (2007).

32 Remsing Rix LL, Rix U, Colinge J *et al.* Global target profile of the kinase inhibitor bosutinib in primary chronic myeloid leukemia cells. *Leukemia* DOI: 10.1038 (2008) (Epub ahead of print).

33 Aye TT, Mohammed S, Toorn HW *et al.* Selectivity in enrichment of PKA RI and RII isoforms and their interactors using modified cAMP affinity resins. *Mol. Cell Proteomics*. (2008) (Epub ahead of print).

•• **Combination of affinity chromatography and mass spectrometry-based quantitative proteomics has been employed to investigate specificity in PKA–AKAP interactions.**

34 Jackson RJ. Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem. Soc. Trans.* 33, 1231–1241 (2005).

35 Lasko P. The *Drosophila melanogaster* genome: translation factors and RNA binding proteins. *J. Cell Biol.* 150, 51–56 (2000).

36 Kozak M. Initiation of translation in prokaryotes and eukaryotes. *Gene* 234, 187–208 (1999).

37 Russo A, Cirulli C, Amoresano A *et al.* *Cis*-acting sequences and *trans*-acting factors in the localization of mRNA for mitochondrial ribosomal proteins. *Biochim. Biophys. Acta.* 1779, 820–829 (2008).

38 Lin JY, Li ML, Huang PN *et al.* Heterogeneous nuclear ribonuclear protein K interacts with the enterovirus 71 5′ untranslated region and participates in virus replication. *J. Gen. Virol.* 89, 2540–2549 (2008).

39 Medugno L, Costanzo P, Lupo A *et al.* A novel zinc finger transcriptional repressor, ZNF224, interacts with the negative regulatory element (AldA-NRE) and inhibits gene expression. *FEBS Lett.* 534, 93–100 (2003).

40 Hall-Pogar T, Liang S, Hague LK *et al.* Specific *trans*-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3′-UTR. *RNA* 13, 1103–1115 (2007).

41 Lambert JP, Mitchell L, Rudner A *et al.* A novel proteomic approach for the discovery of chromatin associated protein networks. *Mol. Cell Proteomics* (2008) (Epub ahead of print).

42 Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 422, 198–207 (2003).

43 Monti M, Orrù S, Pagnozzi D *et al.* Functional proteomics. *Clin. Chim. Acta* 357, 140–150 (2005).

44 Monti M, Orrù S, Pagnozzi D *et al.* Interaction proteomics. *Biosci Rep.* 25, 45–56 (2005).

45 Gavin AC, Bösche M, Krause R *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147 (2002).

46 Ho Y, Gruhler A, Heilbut A *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415, 180–183 (2002).

47 Goll J, Uetz P. The elusive yeast interactome. *Genome Biol.* 7, 223 (2006).

48 Parcellier A, Schmitt E, Gurbuxani S *et al.* hsp27 is a ubiquitin-binding protein involved in I-κBα proteasomal degradation. *Mol. Cell Biol.* 23, 5790–5802 (2003).

49 Loyer P, Trembley JH, Grenet JA *et al.* Characterization of cyclin L1 and L2 interactions with CDK11 and splicing factors. *J. Biol. Chem.* 283, 7721–7732 (2008).

50 Liu X, Yang WC, Gao Q *et al.* Toward chromatographic analysis of interacting protein networks. *J. Chromatogr. A* 1178, 24–32 (2008).

51 Tanese N. Small-scale density gradient sedimentation to separate and analyze multiprotein complexes. *Methods* 12, 224–234 (1997).

52 Vogelmann R, Nelson WJ. Fractionation of the epithelial apical junctional complex: reassessment of protein distributions in different substructures. *Mol. Biol. Cell* 16, 701–716 (2005).

53 Dörr J, Kartarius S, Götz C *et al.* Contribution of the individual subunits of protein kinase CK2 and of hPrp3p to the splicing process. *Mol. Cell Biochem.* 316, 187–193 (2008).

54 Schägger H, von Jagow G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* 199, 223–231 (1991).

55 Camacho-Carvajal MM, Wollscheid B, Aebersold R *et al.* Two-dimensional blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. *Mol. Cell Proteomics* 3, 176–182 (2004).

56 Krause F. Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (membrane) protein complexes and supercomplexes. *Electrophoresis* 27, 2759–2781 (2006).

57 Selbach M, Mann M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat. Methods* 3, 981–983 (2006).

58 Castle JD. Purification of organelles from mammalian cells. *Curr. Protoc. Immunol.* 8, 8.1B (2003).

59 Chang J, Ruiz V, Vancura A. Purification of yeast membranes and organelles by sucrose density gradient centrifugation. *Methods Mol. Biol.* 457, 141–149 (2008).

60 Wilhelm JE, Hilton M, Amos Q *et al.* Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. *J. Cell Biol.* 163, 1197–1204 (2003).

61 Nakamura A, Sato K, Hanyu-Nakamura K. *Drosophila* cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. *Dev. Cell* 6, 69–78 (2004).

62 Clouse KN, Ferguson SB, Schupbach T. Squid, Cup, and PABP55B function together to regulate gurken translation in *Drosophila*. *Dev. Biol.* 313, 713–724 (2008).

63 Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477–480 (2005).

64 Holaska JM, Wilson KL. An emerin "proteome": purification of distinct emerin-containing complexes from HeLa cells suggests molecular basis for diverse roles including gene regulation, mRNA splicing, signaling, mechanosensing, and nuclear architecture. *Biochemistry* 46, 8897–8908 (2007).

•• **Emerin-associated protein complexes from HeLa cells were purified by immunoprecipitation followed by ion-exchange chromatography and finally by size-exclusion chromatography.**

Affiliations

- Piero Pucci
  Full Professor of Biochemistry, Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7896; +39 081 674 318
  Fax: +39 081 373 7808; +39 081 674 313
  pucci@unina.it
- Maria Monti, PhD
  Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895; +39 081 674 474
  Fax: +39 081 373 7808; +39 081 674 313
  montimar@unina.it
- Marianna Cozzolino
  PhD student, Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895
  Fax: +39 081 373 7808
  cozzolinom@ceinge.unina.it
- Flora Cozzolino
  PhD student, Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895
  Fax: +39 081 373 7808
  cozzolinof@ceinge.unina.it
- Giuseppina Vitiello, BSc
  Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895
  Fax: +39 081 373 7808
  vitiello@ceinge.unina.it
- Roberta Tedesco
  Biologist, Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895
  Fax: +39 081 373 7808
  tedesco@ceinge.unina.it
- Angela Flagiello, BSc
  Dipartimento di Chimica Organica e Biochimica and CEINGE Biotecnologie Avanzate s.c.a r.l., Università di Napoli "Federico II", Via Comunale Margherita, 482 80145, Napoli, Italy
  Tel.: +39 081 373 7895
  Fax: +39 081 373 7808
  flagiello@ceinge.unina.it