ELSEVIER

Clinica Chimica Acta 357 (2005) 140–150

www.elsevier.com/locate/clinchim

Functional proteomics

Maria Monti, Stefania Orrù, Daniela Pagnozzi, Piero Pucci*

CEINGE Biotecnologie Avanzate scarl and Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, via Comunale Margherita 482, Naples 80145, Italy

Received 4 March 2005; accepted 9 March 2005
Available online 8 June 2005

Abstract

Background: With the increase in the number of genome sequencing projects, there is a concomitant exponential growth in the number of protein sequences whose function is still unknown. Functional proteomics constitutes an emerging research area in the proteomic field whose approaches are addressed towards two major targets: the elucidation of the biological function of unknown proteins and the definition of cellular mechanisms at the molecular level.
Methods: The identification of interacting proteins in stable complexes in vivo is essentially achieved by affinity-based procedures. The basic idea is to express the protein of interest with a suitable tag to be used as a bait to fish its specific partners out from a cellular extract. Individual components within the multi-protein complex can then be identified by mass spectrometric methodologies.
Results and conclusions: The association of an unknown protein with partners belonging to a specific protein complex involved in a particular mechanism is strongly suggestive of the biological function of the protein. Moreover, the identification of protein partners interacting with a given protein will lead to the description of cellular mechanisms at the molecular level. The next goal will be to generate animal models bearing a tagged form of the bait protein.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Functional proteomics; Mass spectrometry; Protein–protein interactions; Affinity-based strategies

1. Introduction

As the human genome sequencing project came to a successful end, it appeared immediately clear that the knowledge of the entire DNA sequence of an organism certainly represented a wealth of information but constituted more a starting point than the "end of the story" in understanding the function of living cells at the molecular level. In the last years, the challenge has then shifted to the protein side,

Abbreviations: AldA-NRE, Aldolase A negative regulatory element; FCP1, TFIIF-associating component of CTD phosphatase; KRAB-ZFPs, Krüppel-like zinc-finger proteins; MEP50, methylosome protein 50; RNAPII-CTD, RNA polymerase II-carboxy terminal domain; TAP, tandem affinity purification; ZnF224, zinc-finger protein 224.

* Corresponding author. Tel.: +39 0813722896; fax: +39 0813722808.
E-mail address: pucci@unina.it (P. Pucci).

0009-8981/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cccn.2005.03.019

giving rise to the so-called "proteomic era," with the aims of identifying and localising proteins within a given organelle, cell, or even organism, as well as unraveling protein pathways in vivo [1–8]. These new goals, however, cannot be easily achieved as intrinsic difficulties increase by several orders of magnitude when moving from genome to proteome research. The static nature of the genome, in fact, cannot be compared to the dynamic properties of the proteome; protein expression profiles change several times during the cell cycle and are heavily affected by a number of intracellular and extracellular stimuli (temperature, stress, apoptotic signals, etc.) [1]. Moreover, the occurrence of alternative splicing and post-translational modifications led to a complete rethinking of the old paradigm "one gene–one protein," which does not reflect anymore the real nature of the cellular proteome.

Current proteome investigations are essentially focused on two major areas: expression proteomics, which aims to measure the up-regulation and down-regulation of protein levels, and functional proteomics, which aims to characterise protein activities, multiprotein complexes, and signaling pathways [9–12]. Typically, expression proteomics studies are addressed to the investigation of the expression protein patterns in abnormal cells (i.e., malignant, stimulated by drug treatment, etc.), in comparison with normal cells. In biomedical application, this comparative approach is usually employed to identify proteins that are upregulated or downregulated in a disease-specific manner for use as diagnostic markers or therapeutic targets [13–16]. In these studies, a reliable analysis of quantitative changes in protein expression is crucial. Such changes are often obtained from the staining intensities of protein spots on gels, a labor-intensive method that is prone to error. Recently, better and more reliable results were achieved using stable isotope methodologies or dual fluorescent techniques [17–20].

Functional proteomics constitutes an emerging research area in the proteomic field that is "focused to monitor and analyse the spatial and temporal properties of the molecular networks and fluxes involved in living cells" [1]. With the increase in the number of genome sequencing projects that are being carried out, there is a concomitant exponential growth in the number of protein sequences whose function is still unknown. Biological sciences are then experiencing a sort of paradoxical situation in which the protein sequence, the corresponding coding gene, its chromosomic localization, or even the regulation mechanisms may have been elucidated, but the biological role of the protein in the cell is still completely obscure.

Functional proteomics approaches are addressed towards two major targets: the elucidation of biological functions of unknown proteins and the definition of cellular mechanisms at the molecular level. In the cells, many proteins display their biological functions through the rapid and transient association within large protein complexes [21]. Understanding protein functions as well as unraveling molecular mechanisms within the cell then depend on the identification of the interacting protein partners. The association of an unknown protein with partners belonging to a specific protein complex involved in a particular mechanism would, in fact, be strongly suggestive of its biological function [22,23]. Furthermore, a detailed description of the cellular signaling pathways might greatly benefit from the elucidation of protein–protein interactions in vivo [24].

2. Protein identification by mass spectrometry methodologies

The identification of proteins fractionated by either 1D or 2D gel electrophoresis is essentially obtained through peptide mass fingerprinting using MALDI-MS (Fig. 1). Proteins are excised from the gel, reduced and alkylated with iodoacetamide to irreversibly block the cysteine residues, and digested in situ with trypsin. The resulting peptide mixture is then directly analysed by MALDI-MS using a reflectron instrument. Identification of the various proteins is carried out through the peptide mass fingerprinting procedure: the mass values, together with other information, such as the protease used for the hydrolysis and the protein molecular mass roughly estimated from the SDS PAGE gel, are introduced into different mass search programs (ProFound, Mascot, MS-Fit, etc.) available on the net. The mass values are compared to those originated from the theoretical digestion of all the proteins occurring in the database, leading to the identification of the protein(s).

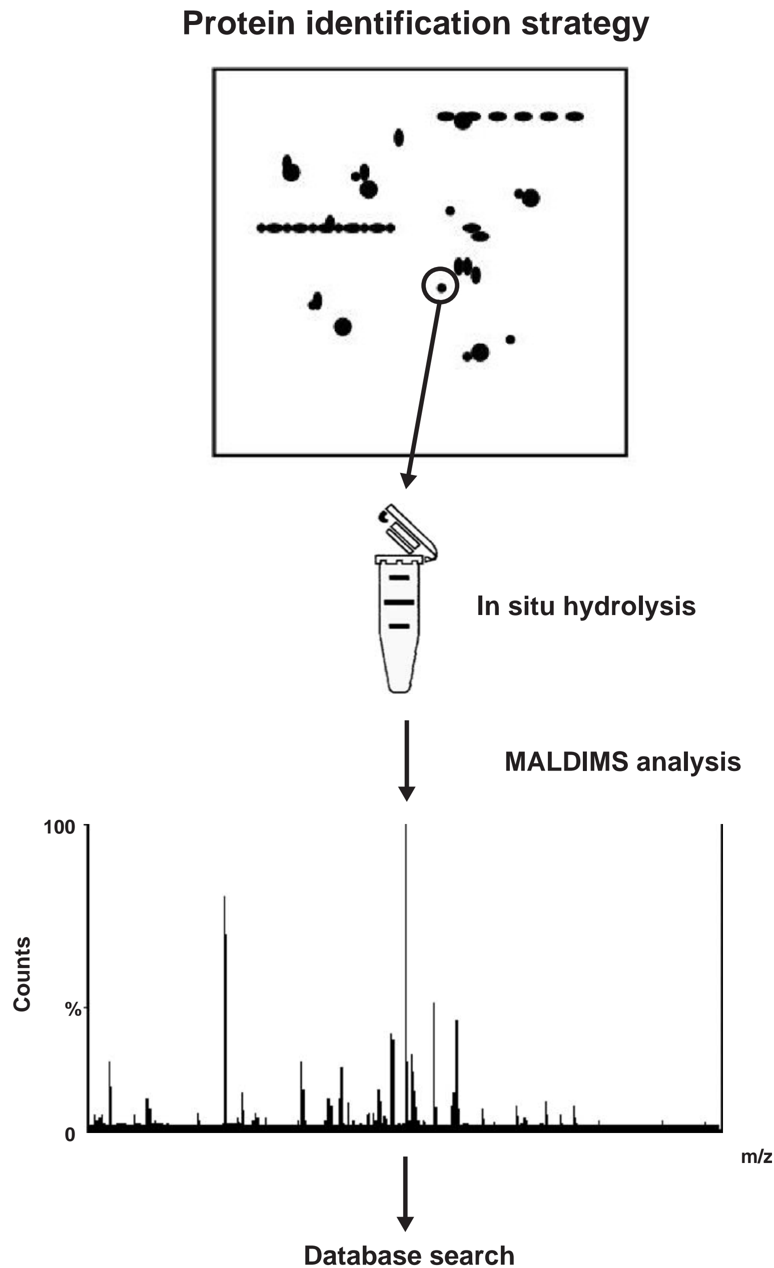

Fig. 1. Outline of protein identification strategy by peptide mass fingerprinting.

Alternatively, when the mass fingerprinting procedure is not sufficient to identify the proteins, liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) methods can be employed (Fig. 2). Peptide mixtures produced by in situ digestions are fractionated by capillary HPLC analysis, the fractions eluted from the column are directly inserted into the ES mass spectrometry source, and their mass values are determined. Peptide ions will simultaneously be isolated and fragmented within the mass spectrometer, producing daughter ion spectra from which sequence information on individual peptides can be obtained.

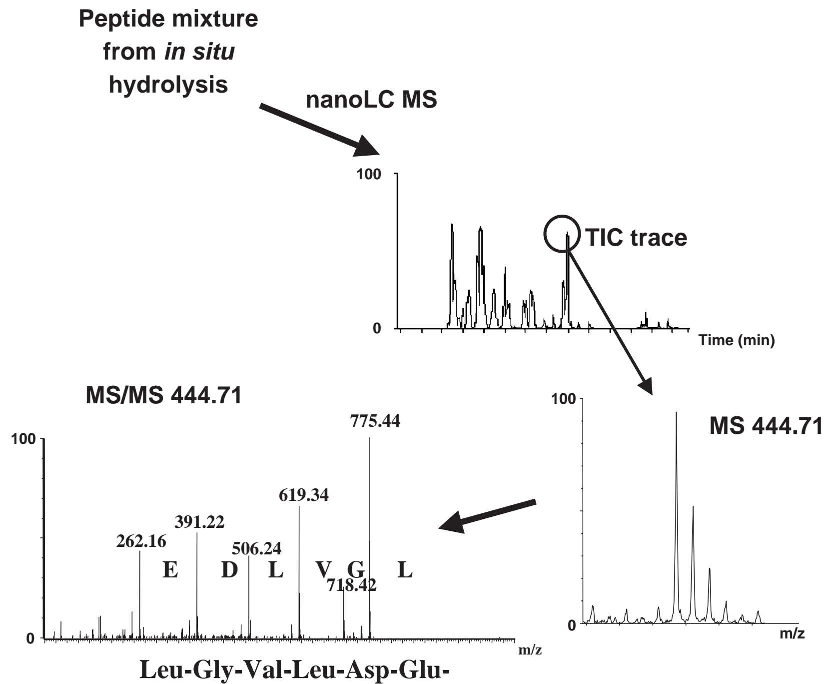

Fig. 2. Schematic description of protein identification performed by the LC-MS/MS approach.

This information is then used to search for protein databases, leading to the identification of the protein components.

3. Identification of protein partners by functional proteomics approaches

The identification of interacting proteins in stable complexes in vivo essentially relies on affinity-based procedures. The basic idea is to express the protein of interest with a suitable tag to be used as a bait to fish its specific partners out from a cellular extract. Isolation of the entire multi-protein complex can then be accomplished by taking advantage of the availability of several anti-tag systems, immobilised on agarose–sepharose supports, and by showing high binding efficiency, as illustrated in Table 1 [25]. Different strategies relying on this simple concept have been developed and a brief overview of the main approaches presently used in functional proteomics studies is described below.

3.1. Fishing for partners' strategies

Using commercially available protein expression systems, the protein bait can be produced as a hybrid protein fused to the glutathione *S*-transferase (GST-fused protein) or to small peptide epitopes (i.e., FLAG, HA, or c-myc), or containing a poly-His tail, or covalently modified with biotin. In all cases, the tagged bait can be immobilised onto agarose beads derivatised with the appropriate anti-tag ligand (glutathione, anti-epitope antibodies, nickel ions, streptavidin, etc.). All these affinity tag systems provided a

Table 1
Affinity tags and ligands commonly used in the isolation of multi-protein complexes

Tag	Ligand
Poly-His	Ni++
Biotin	Streptavidin
Calmodulin-binding peptide	Calmodulin (Ca++)
GST	Glutathione
Specific epitope (FLAG, c-myc, HA, etc.)	Monoclonal Ab

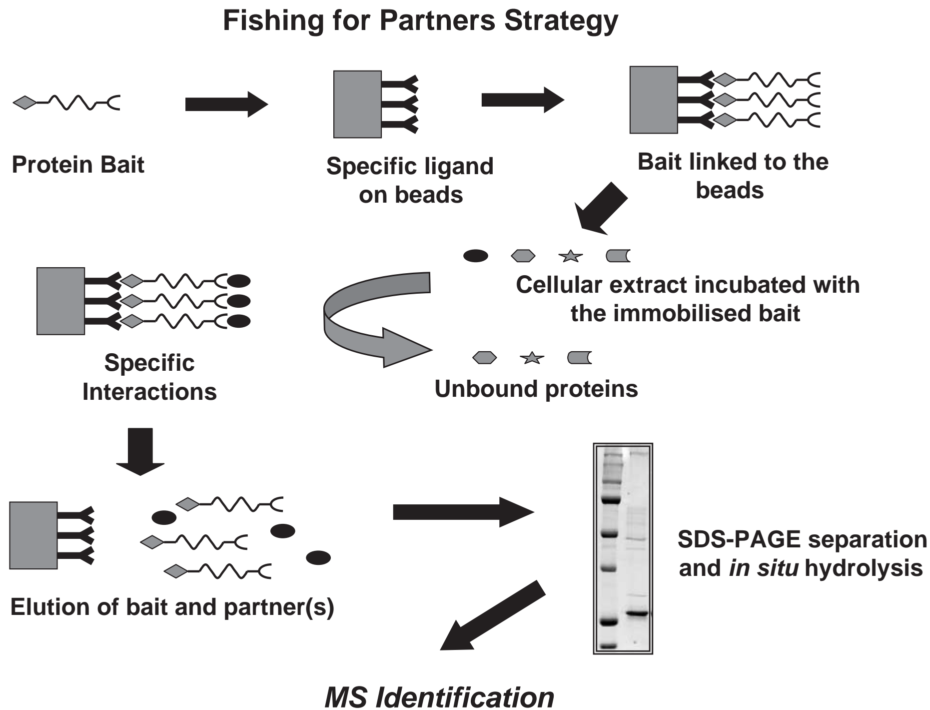

Fig. 3. Outline of the "fishing for partners" strategy.

general applicability with a large number of proteins and a minimal effect on the tertiary structure and biological activity of the bait, preventing complexes' instability. The entire cellular extract and/or, when appropriate, the extract from specific organelles can then be incubated with the immobilised bait. The immobilised protein forms stable non-covalent interactions with specific partners occurring in the cellular extract, whereas the unbound proteins will be eluted during the washing. The protein components specifically recognised by the bait and retained on the agarose beads can then be eluted and fractionated by SDS-PAGE. The protein bands detected on the gel are in situ enzymatically digested and the resulting peptide mixtures are analysed by MALDI/MS fingerprinting and/or LC-MS/MS techniques, leading to the identification of the proteins by database search. An outline of this approach is shown in Fig. 3.

This strategy was applied to the identification of the protein partners of ZnF224, a zinc-finger protein of about 82 kDa belonging to the "Krüppel-like" zinc-finger proteins family (KRAB-ZFPs), one of the largest classes of transcription factors (Fig. 4) [26]. ZnF224 specifically binds to the negative regulatory

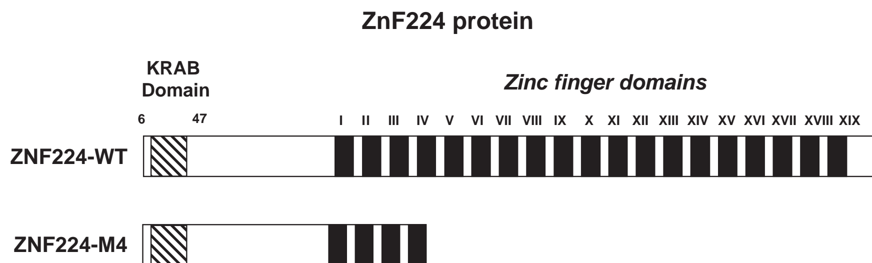

Fig. 4. Schematic structure of the zinc finger ZnF224 protein and its deletion mutant *ZnF224-M4*, lacking the last 15 zinc-finger domains.

element (AldA-NRE) located in the promoter region of the human aldolase A gene through its array of zinc fingers and inhibits the transcription by the 45-amino-acid KRAB-A domain. The proteins interacting with ZnF224 and involved in the transcriptional inhibition, as well as the molecular mechanisms of these negative regulation processes, are still unknown. In order to elucidate these aspects, the full-length cDNAs of ZnF224 and its deletion mutant *ZnF224-M4* containing only four zinc fingers were expressed as fused protein to the C-terminus of GST and purified on glutathione–sepharose resin. The purified chimeric bait was then linked to GSH-activated beads.

Nuclear 293 cell extracts were pre-purified by incubation with beads treated with GST alone in order to get rid of both matrix and GST-interacting proteins and to obtain a higher cleaning of the protein mixture. The pre-cleaned extract was then incubated with the GST-ZnF224 bait. After several washings to remove unbound proteins, the complex components were eluted from the beads, separated by SDS-PAGE, and stained with colloidal Coomassie. The resulting gel pattern is shown in Fig. 5; the poor quality of these results is strongly indicative of the limitations of this procedure, as discussed below.

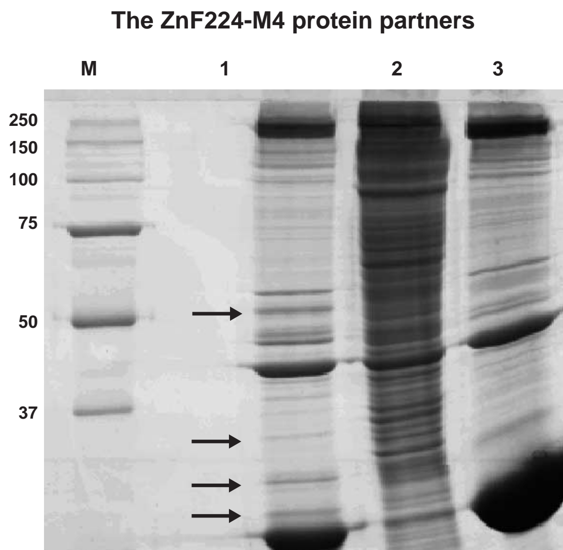

Fig. 5. SDS-PAGE analysis of the GST pulldown experiment carried out using GST-ZnF224-M4 as bait with pre-cleaning steps. Lane M: markers; lane 1: proteins bound to GST-M4; lane 2: unretained proteins; lane 3: proteins bound to GST alone.

The occurrence of an excessive background, the presence of several identical bands in both the sample and the control, and the difficulties in identifying the proteins specifically interacting with the bait immediately underline the need to optimize sample cleaning procedures. Protein bands indicated by arrows appeared to be present only in the sample and were then selected for identification by in situ digestion and LC-MS/MS analysis.

Although this approach has found large applications in the studies of protein–protein interactions, a conspicuous number of drawbacks exist. As mentioned above, extensive pre-cleaning of the extract is needed and a number of control samples have to be prepared. Moreover, when protein extracts are prepared by cellular lysis, the architecture of the subcellular compartments is disrupted and proteins that are normally segregated in different organelles can come in contact, generating nonphysiological interactions. However, the major criticism to this approach is that all the interactions among the bait and its protein partners are essentially in vitro interactions taking place outside the cell, on the derivatised beads. These drawbacks might originate a number of false positives that should be considered when using these procedures.

The success of an affinity-based approach then depends on the absence of excessive aspecific interactions, which in turn is related to the specificity of the bait partners recognition. When this binding specificity is extremely high, as in the case of DNA-binding proteins, very good results can be achieved. In this particular variant of the fishing strategy, the bait consists of a specific oligonucleotide linked to an insoluble support. Nuclear proteins can then be incubated with the bait in search for specific partners, following the strategy outlined above [27]. Control experiments can easily be designed using randomised oligonucleotides. This strategy was applied to the identification of the transcriptional repressor that specifically binds to the human aldolase A (AldA) negative regulatory element (NRE). This factor was eventually identified as the ZnF224 protein [26].

It is important to underline that when a putative candidate is provided by proteomic approaches, this identification has to be confirmed by independent methodologies. In the case of ZnF224, the ability to recognise both its specific DNA binding site and

the transcriptional repressor activity of the aldolase A gene expression had to be tested. Therefore, a band shift experiment with the AldA-NRE oligonucleotide region was carried out using the recombinant form of the wild type protein and two deletion mutants. Moreover, to test the ZNF224-mediated transcriptional repression, a classical CAT reporter gene assay was developed using the recombinant plasmid encoding ZNF224. The expressed ZNF224 protein negatively modulated the reporter CAT gene transcription in a dose-dependent manner. These results demonstrated that indeed ZnF224 was able to specifically recognise the DNA target and to repress AldA-NRE-mediated transcription of a heterologous promoter [26].

3.2. Immunoprecipitation strategies

To overcome most of the drawbacks affecting the affinity-based approaches, alternative strategies essentially relying on immunoprecipitation techniques have been introduced [28]. An outline of these strategies is reported in Fig. 6. The gene coding for the bait tagged with an epitope against which good antibodies exist (FLAG, HA, c-myc, etc.) is transfected into the appropriate cell line and expressed in the cognate host. Protein complexes are allowed to form in vivo within the cell and the cell extracts are immunoprecipitated with anti-tag monoclonal antibodies using suitable experimental conditions to avoid dissociation of the complexes. The immunoprecipitated material containing the protein bait and its interacting partners can then be fractionated by SDS-PAGE and the individual protein components identified by different mass spectrometric methodologies.

The immunoprecipitation strategy was employed to identify the protein partners of FCP1, a conserved phosphatase involved in the regulation of eukaryotic RNA polymerase II [29]. RNA polymerase II is subjected to reversible phosphorylation of the C-

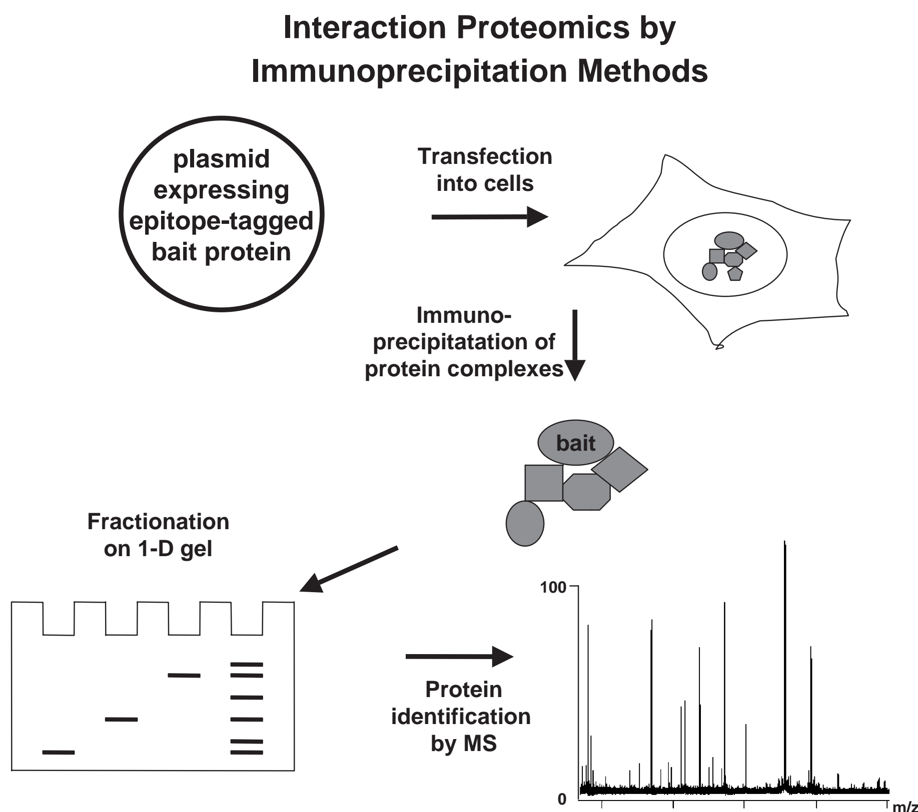

Fig. 6. Outline of the immunoprecipitation strategy.

terminal domain (CTD) of the largest subunit during the transcription cycle. While several kinases affect the CTD phosphorylation status during transcription, FCP1 protein is the only CTD phosphatase identified so far [30,31]. In order to elucidate the role of FCP1, a biochemical characterization of FCP1 associating factors using a 3xFLAG-tagged FCP1 stably expressing cell line was performed [32]. Following complex formation, nuclear extracts were collected and immunoprecipitated with anti-FLAG agarose-conjugated antibody. The protein bands separated on SDS-PAGE and stained by colloidal Coomassie were identified essentially by mass fingerprinting (Fig. 7). Among several other specific FCP1 partners, the methylosome protein 50, MEP50, was identified. This protein is involved in the dimethylation of arginine and belongs to the methylosome complex located in the cytoplasm, while the FCP1 protein is exclusively found in the nucleus. Independent verification experiments were designed, including coimmunoprecipitation of FLAG-FCP1 and a c-myc-tagged MEP50 recombinant form and sub-cellular localization by both sedimentation profiling and immunofluorescence techniques. Taken together, the results obtained demonstrated that MEP50 and FCP1 associate into the nucleus in a complex of the same size distinct from the 20*S* methylosome complex, and that FCP1 is able to interact with components of the pre-mRNA spliceosomal machinery. These findings add further support to the concept that there is functional intercommunication between the transcription and splicing machineries, and the RNAPII-CTD appears to play a pivotal role in coordinating transcription and pre-mRNA processing (Refs. [32,33] and references therein).

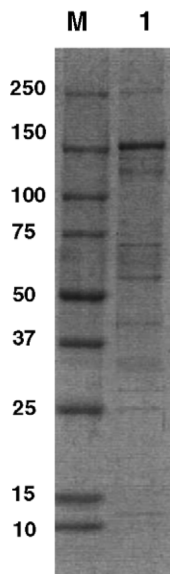

Fig. 7. SDS-PAGE analysis of the immunoprecipitation experiment carried out using 3xFLAG-FCP1 as bait. Lane M: markers; lane 1: retained proteins specifically bound to 3xFLAG-FCP1.

As described above, identification of protein partners by immunoprecipitation techniques has a number of advantages over the fishing for partners approach. However, some negative issues need to be discussed. Antibodies used in common immunostaining methods are not always suitable for immunoprecipitation protocols that require a more efficient (and quantitative) recognition of the proper antigen compared to Western blot or ELISA applications. Cross-recognition of aspecific antigens or aspecific binding of proteins to the antibodies, to the peptide tags, or to the insoluble agarose support can lead to false positives. Pre-immunoprecipitation of the cellular extract with antibodies of the animal host not yet immunised against the specific antigen is then strongly suggested.

According to recent literature, the use of antibodies specifically directed against the protein bait is not encouraged because the antibody might compete with the interacting proteins for binding to the bait epitope, thus leading to destabilization of protein–protein interactions and dissociation of the protein complexes, at least partially. The use of tagged proteins has greatly helped to avoid these problems but has introduced new debated questions. The presence of the tag can affect protein conformation, thus altering or impairing complexes formation. Preliminary experiments with protein baits tagged at either the N- or the C-terminus should be carried out or, alternatively, the three-dimensional structure of the bait, when available, should carefully be considered to decide where the tag should be posed. Finally, overexpression of the tagged protein in the host cells may definitively alter the stoichiometric ratio between the bait and the natural partners, often leading to the formation of nonspecific and/or nonnatural protein interactions with host proteins [34].

3.3. The TAP tag system

The tandem affinity purification (TAP) tag system was developed for the purification of protein complexes in high yield under native conditions [35]. This method represents an implementation of the tag complexes' affinity purification procedures described above, by combining two different tags on the same protein usually spaced by an enzyme-cleavable linker sequence. Since the introduction of the original methodology, a number of variant systems have been proposed. However, the first TAP tag system consisted of two IgG binding domains of *Staphylococcus aureus* protein A (ProtA) and a calmodulin binding peptide (CBP) separated by a TEV protease cleavage site [36]. Both affinity tags have been selected for highly efficient recovery of proteins present at low concentration. However, the system is very flexible and variations to the original strategy, including application of the tag cassette to either the N- or the C-terminal end of the protein, the introduction of alternative tags, and the tailoring of the system for various host organisms, can easily be obtained.

The N- or C-terminal TAP tags are introduced in-frame with the coding region of the protein of interest in an appropriate expression vector using standard DNA cloning procedures. The tagged construct can then be transiently or stably introduced into recipient cells or organisms. Optimally, the recombinant vector should replace the endogenous wild-type gene, although this condition might not be always possible. In all cases, overexpression of the protein bait is avoided since the TAP tag system was specifically designed for recovery of protein complexes expressed at their own natural level.

ProtA binds tightly to an IgG matrix, providing the first affinity purification step of the native complex. The use of the TEV protease is then required to elute the bound material under native conditions. The eluant of this first affinity step still containing the intact complex is then incubated with calmodulin-coated beads in the presence of calcium. Nonspecific protein contaminants and the excess of TEV protease are removed by repetitive washing and the bound material is released under mild conditions by elution with EGTA. The released material containing the protein bait and its interacting partners can then be fractionated by SDS-PAGE and the individual protein components identified by different mass spectrometric methodologies.

The introduction of a double affinity purification procedure significantly reduces the possible occurrence of nonspecific protein contaminants, thus decreasing both the aspecific background noise on the SDS gel and the possible presence of false positives. This method was originally developed in yeast [36] and has been found in large employment into the description of multiprotein complexes in *Saccharomyces cerevisiae* [22]. However, optimized conditions have been developed for the generic use of the TAP strategy. Recently, the system has been successfully optimised even in mammalian cells [37] by using alternative tags to further reduce aspecific interactions between tags and extract proteins [38].

4. Conclusions and perspectives

Understanding protein function and unraveling cellular mechanisms at the molecular level constitute a major need in modern biology. With the availability of full genome sequences, these goals can be achieved by determining which macromolecules interact with a given protein in a specific manner. We are witnessing a scientific period that is reminiscent of the early 1930s when scientists were able to describe "step by step" the metabolic pathways. For the first time, experimental procedures are available to define the transient macromolecular assembly of proteins needed to fulfill specific and fundamental functions within the cell, or to design the routes along which signaling mechanisms can take place. The functional proteomic approaches described in this paper have proven to be useful tools for the detection of interacting partners of a target protein, although each of them highlighted the occurrence of possible drawbacks. Particular attention should be paid to attain low levels of false positives, which might result in misleading interpretations; improvements and refining of the affinity-based procedures should take great care of this point. A further factor that is often neglected is the transient nature of protein complexes: protein assembly at the right place and the right time to fulfill a specific function. The complex then dissociates and individual components can participate in the formation of other complexes, following the occurrence of specific signals. Any

attempt to the description of protein interaction networks involved in cell functions should carefully consider the dynamics of the formation and dissociation of protein complexes, as the same protein can interact with different partners at different times. Finally, the goal to be pursued in the future will be to transfer these approaches to real in vivo systems by generating animal models bearing a tagged form of the given protein. If vital animals are obtained, homozygous embryos will then provide tissues and/or progenitor cells for immunoprecipitation assays. Proteomic analyses of the protein complexes occurring in vivo will disclose the identity of the individual components and whether they differ from a territory to another.

Acknowledgement

This work was supported, in part, by MIUR (PRIN 2003).

References

[1] Godovac-Zimmermann J, Brown LR. Perspectives for mass spectrometry and functional proteomics. Mass Spectrom Rev 2001;20:1–57.

[2] Panisko EA, Conrads TP, Goshe MB, Veenstra TD. The post-genomic age: characterization of proteomes. Exp Hematol 2002;30:97–107.

[3] Michaud GA, Snyder M. Proteomic approaches for the global analysis of proteins. Biotechniques 2002;33:1308–16.

[4] Souchelnytskyi S. Proteomics in studies of signal transduction in epithelial cells. J Mammary Gland Biol 2002;7:359–71.

[5] Taylor SW, Fahy E, Ghosh SS. Global organellar proteomics. Trends Biotechnol 2003;21:82–8.

[6] Patterson SD, Aebersold RH. Proteomics: the first decade and beyond. Nat Genet 2003;33:311–23.

[7] Dreger M. Proteome analysis at the level of subcellular structures. Eur J Biochem 2003;270:589–99.

[8] Godovac-Zimmermann J, Brown LR. Proteomics approaches to elucidation of signal transduction pathways. Curr Opin Mol Ther 2003;5:241–9.

[9] Neubauer G, Gottschalk A, Fabrizio P, Seraphin B, Lurhmann R, Mann M. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A 1997;94:385–90.

[10] Shevchenko A, Keller P, Scheiffele P, Mann M, Simons K. Identification of components of trans-Golgi network-derived transport vesicles and detergent-insoluble complexes by nanoelectrospray tandem mass spectrometry. Electrophoresis 1997;18:2591–600.

[11] Pandey A, Fernandez MM, Stehen H, Blagoev B, Nielsen MM, Roche S, et al. Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. J Biol Chem 2000;275:38633–9.

[12] Hinsby AM, Olsen JV, Bennett KL, Mann M. Signaling initiated by overexpression of the fibroblast growth factor receptor-1 investigated by mass spectrometry. Mol Cell Proteomics 2003;2:29–36.

[13] Soldes OS, Kuick RD, Thompson II IA, Hughes SJ, Orringer MB, Iannettoni MD, et al. Differential expression of Hsp27 in normal oesophagus, Barrett's metaplasia and oesophageal adenocarcinomas. Br J Cancer 1999;79:595–603.

[14] Banks RE, Dunn MJ, Hochstrasser DF, Sanchez JC, Blackstock W, Pappin DJ, et al. Proteomics: new perspectives, new biomedical opportunities. Lancet 2000;356:1749–56.

[15] Dunn MJ. Studying heart disease using the proteomic approach. Drug Discov Today 2000;5:76–84.

[16] Jungblut PR, Zimny-Arndt U, Zeindl-Eberhart E, Stulik J, Koupilova K, Pleissner KP, et al. Proteomics in human disease: cancer, heart and infectious diseases. Electrophoresis 1999;20:2100–10.

[17] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999;17:994–9.

[18] Zhou H, Ranish JA, Watts JD, Aebersold R. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. Nat Biotechnol 2002;20:512–5.

[19] Yan JX, Devenish AT, Wait R, Stone T, Lewis S, Fowler S. Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of *Escherichia coli*. Proteomics 2002;2:1682–98.

[20] Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 2003;3:36–44.

[21] Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 1998;92:291–4.

[22] Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 2002;415: 141–7.

[23] Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, et al. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 2002;415:180–3.

[24] Lewis TS, Hunt JB, Aveline LD, Jonscher KR, Louie DF, Yeh JM, et al. Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. Mol Cell 2000;6:1343–54.

[25] Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 2003;60:523–33.

[26] Medugno L, Costanzo P, Lupo A, Monti M, Florio F, Pucci P, et al. A novel zinc finger transcriptional repressor, ZNF224, interacts with the negative regulatory element (AldA-NRE) and inhibits gene expression. FEBS Lett 2003;534:93–100.

[27] Kadonaga JT, Tijan R. Affinity purification of sequence-specific DNA binding proteins. Proc Natl Acad Sci U S A 1986;83:5889–93.

[28] Cho S, Park SG, Lee DH, Park BC. Protein–protein interaction networks: from interactions to networks. J Biochem Mol Biol 2004;37:45–52.

[29] Kobor MS, Archambault J, Lester W, Holstege FC, Gileadi O, Jansma DB, et al. An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. Mol Cell 1999;4:55–62.

[30] Archambault J, Chambers RS, Kobor MS, Ho Y, Cartier M, Bolotin D, et al. An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 1997;94:14300–5.

[31] Archambault J, Pan G, Dahmus GK, Cartier M, Marshall N, Zhang S, et al. FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase IIO. J Biol Chem 1998;273:27593–601.

[32] Licciardo P, Amente S, Ruggiero L, Monti M, Pucci P, Lania L, et al. The FCP1 phosphatase interacts with RNA polymerase II and with MEP50 a component of the methylosome complex involved in the assembly of snRNP. Nucleic Acids Res 2003;31:999–1005.

[33] Maniatis T, Reed R. An extensive network of coupling among gene expression machines. Nature 2002;416:499–506.

[34] Swaffield JC, Melcher K, Johnston SA. A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein. Nature 1995;374:88–91.

[35] Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, et al. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods 2001;24:218–29.

[36] Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol 1999;17:1030–2.

[37] Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. Nat Cell Biol 2004;6:97–105.

[38] Knuesel M, Wan Y, Xiao Z, Holinger E, Lowe N, Wang W, et al. Identification of novel protein–protein interactions using a versatile Mammalian tandem affinity purification expression system. Mol Cell Proteomics 2003;2:1225–33.