Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

Mapping multiprotein complexes by affinity purification and mass spectrometry

Mark O Collins and Jyoti S Choudhary

The combination of affinity purification and tandem mass spectrometry (MS) has emerged as a powerful approach to delineate biological processes. In particular, the use of epitope tags has allowed this approach to become scaleable and has bypassed difficulties associated with generation of antibodies. Single epitope tags and tandem affinity purification (TAP) tags have been used to systematically map protein complexes generating protein interaction data at a near proteome-wide scale. Recent developments in the design of tags, optimisation of purification conditions, experimental design and data analysis have greatly improved the sensitivity and specificity of this approach. Concomitant developments in MS, including high accuracy and high-throughput instrumentation together with quantitative MS methods, have facilitated large-scale and comprehensive analysis of multiprotein complexes.

Addresses
Proteomic Mass Spectrometry, The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

Corresponding author: Choudhary, Jyoti S (jc4@sanger.ac.uk)

Current Opinion in Biotechnology 2008, **19**:324–330

This review comes from a themed issue on
Protein technologies
Edited by John Timms

Available online 16th July 2008

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2008.06.002

Introduction

The development of protein-tagging methods and protein identification by tandem mass spectrometry (MS) has revolutionised the way we look at cell biology. The combined use of protein tagging and MS has been used to systematically map protein interactions in many species including yeast [1,2,3••,4••], *Escherichia coli* [5,6] and human cell lines [7••]. In addition, this approach has enabled mapping of specific biological pathways such as the TNFα/NF-κβ [8] and WNT/β-catenin [9] pathways as well as specific interactomes such as the human transcription machinery [10] and the embryonic stem cell pluripotency network [11]. The major requirements for this type of protein interaction mapping are specific, efficient and scaleable protein complex purification methods, sensitive and accurate MS acquisition platforms for reproducible and high-throughput protein identification and data analysis pipelines that exploit qualitative and quantitative features of MS data that effectively identify false protein identifications with minimum loss of sensitivity.

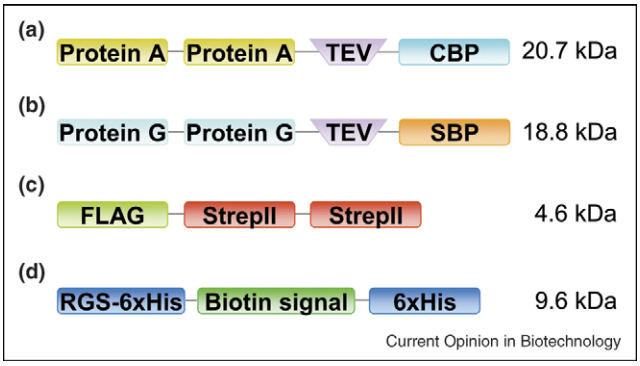
In this review we focus on technological developments over the past two years, in particular methods for generation and analysis of tagged-protein complexes. This encompasses developments in the design of tandem affinity purification (TAP) tags, large-scale protein-interaction mapping, novel strategies for the identification of transient protein interactions and structural analysis of protein complexes.

Novel TAP-tagging strategies

The original TAP tag developed for yeast consisted of a Protein A tag and a calmodulin-binding peptide (CBP) tag separated by a tobacco etch virus (TEV) protease cleavage site (Figure 1A) [12]. Although it has been used to purify complexes from mammalian cells, yields from these purifications are low and consequently large numbers of cells are needed as starting material [13•]. Burckstummer *et al.* [13•] compared the efficiency of this yeast TAP tag with that of three other TAP tags containing combinations of Protein A, Protein G, CBP and streptavidin-binding peptide (SBP) with TEV cleavage sites. The GS-TAP (Protein G and SBP combination, Figure 1B) strategy produced a 10-fold increase in the efficiency compared to the conventional TAP tag and allowed purification of complexes from one-tenth of the starting material needed for conventional TAP purifications. This refinement of the TAP tag strategy will be useful for studies where starting material is limited or for the purification of complexes of low abundance.

The addition of a TAP tag (in the region of 21 kDa) to the C-terminus or N-terminus of a protein may affect its folding and/or activity and hence the success of the purification and validity of identified protein interactions. In an effort to minimise such potential effects, Gloeckner *et al.* [14] designed a novel TAP tag (SF-TAP) composed of a tandem Strep-tag II and a Flag tag, reducing the size of the tag to 4.6 kDa (Figure 1C). This TAP tag allows elution with desthiobiotin in the first purification and with Flag peptide in the second purification. Of the three proteins that were tagged with the conventional TAP tag and the SF-TAP tag, the expression level of the TAP-tagged B-Raf was significantly lower than the SF-tagged version, indicating possible instability of this protein when fused to the larger tag.

Figure 1



Novel tandem affinity purification (TAP) tags. New combinations of epitope tags are shown with approximate molecular weight. **(a)** The original yeast TAP tag is shown for comparison. **(b)** Use of this optimised GS-TAP tag was found to result in 10-fold increased yields when compared to the original yeast TAP tag [13•]. **(c)** This small TAP tag was developed to minimise potential functional effects of protein tagging and allows fast purification times. **(d)** This combination of tags allows purification of protein complexes under denaturing conditions which is particularly useful for cross-linked complexes [23]. Variants of this tag contain one hexahistidine tag and a TEV cleavage site.

A limitation of cDNA tagging for the isolation of protein complexes is that tagged transgenes are not under the control of endogenous promoters. Poser *et al.* [15] have recently overcome this limitation by exploiting BAC transgenes (which are large enough to ensure that most regulatory elements are present) to express tagged proteins in mammalian cells. This ensures that expression of transgenes is as close to that of endogenous genes as possible without performing gene-targeting experiments. BAC transgenes were rapidly tagged using recombineering in a 96-well format to introduce TAP tag that consisted of an extended green fluorescent protein (EGFP), an S-peptide tag and precision and TEV protease cleavage sites (EGFP-S-peptide). This TAP tag allows cellular localisation (using EGFP) as well as protein complex purification and analysis of protein–DNA interactions by tag-based ChIP. In addition, this BAC-tagging approach can be applied to ES cells to create transgenic mice and is thus a versatile high-throughput approach for mammalian functional proteomics.

The composition of protein complexes in cell lines may not directly reflect the composition of the complex *in vivo*; this is especially true for the brain because of its complexity in terms of brain regions and multiple cell types. Angrand *et al.* [16] tackled this issue by generating a TAP-tagged transgenic mouse, in which TAP-tagged 14-3-3ζ was expressed under the control of a human ubiquitin C promoter. Although, the bait protein was not under the control of its endogenous promoter, the expression level was found to be in the same range as the endogenous gene. LC–MS/MS analysis of purified complexes identified 147 14-3-3ζ associated proteins including almost 40 novel interactors that were not identified in TAP-tagged 14-3-3ζ purifications from HEK293 cells [17], indicating that these may be tissue-specific interactors. Although this approach is clearly useful for ubiquitously expressed proteins, it may not be suitable for genes with more tightly regulated expression and therefore a strategy that targets the endogenous gene to generate TAP-knockin mice would be preferable. We have characterised a TAP-tag-knockin mouse line in which a synaptic-scaffolding protein was tagged by homologous recombination. Introduction of the tag did not produce any detectable abnormalities in the mice and LC–MS/MS analysis of endogenous protein complexes identified a core synaptic interactome (Fernandez *et al.*, unpublished data).

Systematic analysis of protein complexes

In the past two years a number of large-scale studies of protein complexes have been reported; more comprehensive analyses of yeast and *E. coli* interactomes as well as the first analysis of a human interactome. In 2006, two comprehensive studies of protein complexes in yeast were published [3••,4••]. Both of these studies employed systematic TAP tagging by homologous recombination and increased proteome coverage significantly (up to 72% [4••]) compared to initial large-scale TAP and Flag-tagging experiments in yeast [1,2]. Gavin *et al.* [3••] successfully tagged and characterised protein complexes associated with 1754 bait proteins revealing 2760 protein identifications. Computational analysis based on the basis of the raw protein identification data led to a core set of 1483 proteins organised into 491 distinct complexes. The reproducibility of protein complex characterisation was estimated by comparing the components of 139 protein complexes that were purified in duplicate. From this, 69% of proteins were common to replicate purifications, indicating that in any interaction study more replicates may be necessary to obtain a core set of high confidence PPIs.

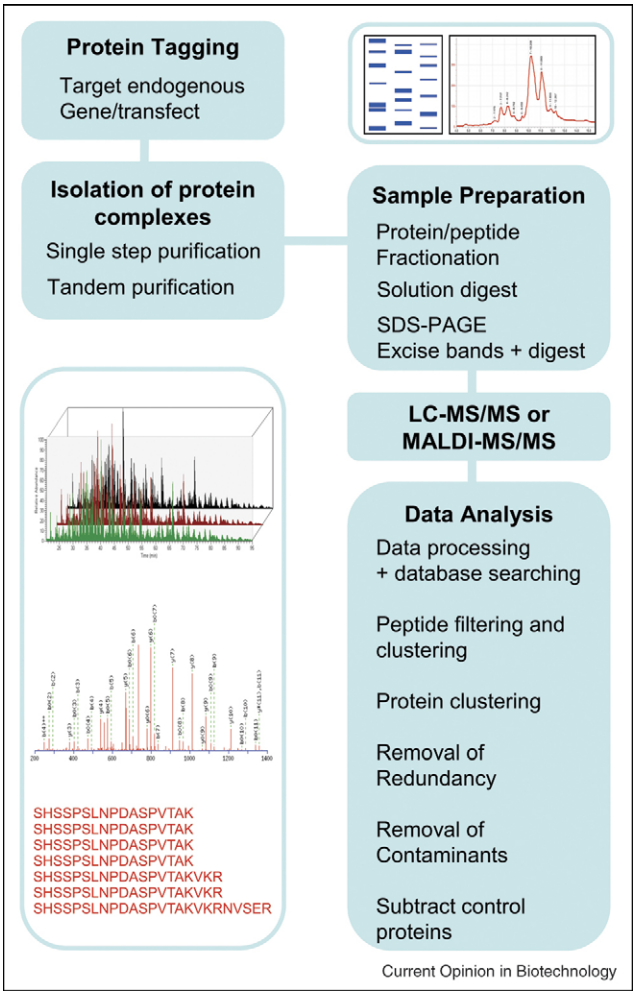
Krogan *et al.* [4••] successfully tagged and characterised protein complexes associated with 2357 yeast bait proteins revealing 4087 protein identifications. Application of a machine-learning algorithm based on published protein interaction data to this dataset, allowed a core set of 2708 proteins organised into 547 distinct complexes to be defined. Differences in the purification protocols, filtering of protein identifications (including removal of contaminants [4••]) and computational approaches to define protein complexes, makes these two yeast datasets difficult to compare [18]. Although, the majority of yeast proteins were tagged in these studies, only one-third led to successful purifications and protein identifications. Clearly, some protein classes such as transmembrane proteins require optimised conditions and tagging of some proteins may affect protein stability; however, up to 72% of the yeast proteome was amenable to identification by MS because of the presence of proteins in multiple complexes. The overlap between

these two datasets was just 1152 proteins (18%) but it would probably increase if the same filtering and analysis rules were applied to both datasets [18]. Although, this is an improvement in reproducibility compared to previous large-scale studies in yeast [1,2] and in *E. coli* [5,6], it is still a major issue in large-scale protein interaction experiments. Further work is necessary to address points of variability in the analysis pipeline as well as standardisation of data analysis and data representation [19].

Arifuzzaman *et al.* [5] extended the *E. coli* protein interaction network by hexahistidine (His)-tagging 4339 bait proteins and performing large-scale pull-down experiments which were analysed by MALDI-TOF MS to identify protein interactors. 2667 (61%) bait proteins were successfully overexpressed and purified and 2667 of these were associated with a binding protein. The remaining 39% of tagged proteins failed because of problems associated with overexpression and purification, particularly for membrane proteins which represented up to 50% of the failed bait proteins. After the subtraction of protein identifications in control experiments, 11 511 protein–protein interactions (PPIs) were identified for 2667 bait proteins. Previous TAP and SPA (sequential peptide affinity) tagging of *E. coli* proteins with expression under the control of native promoters, allowed the identification of over 5000 PPIs for 648 bait proteins [6]. The overlap of tagged bait proteins in these two studies was 521, with 1168 and 995 prey proteins identified by the TAP/SPA-tagging and His-tagging approaches, respectively. This corresponds to an overlap of PPI's of only 4 and 7% for the TAP/SPA and His-tagging studies, respectively. Although it was suggested that this reproducibility was in the range of that observed for yeast 2 hybrid experiments, it highlights concerns about the reproducibility of such large-scale studies. This important issue of reproducibility of identification of PPIs has been addressed recently and a novel interaction confidence score has been developed [7••].

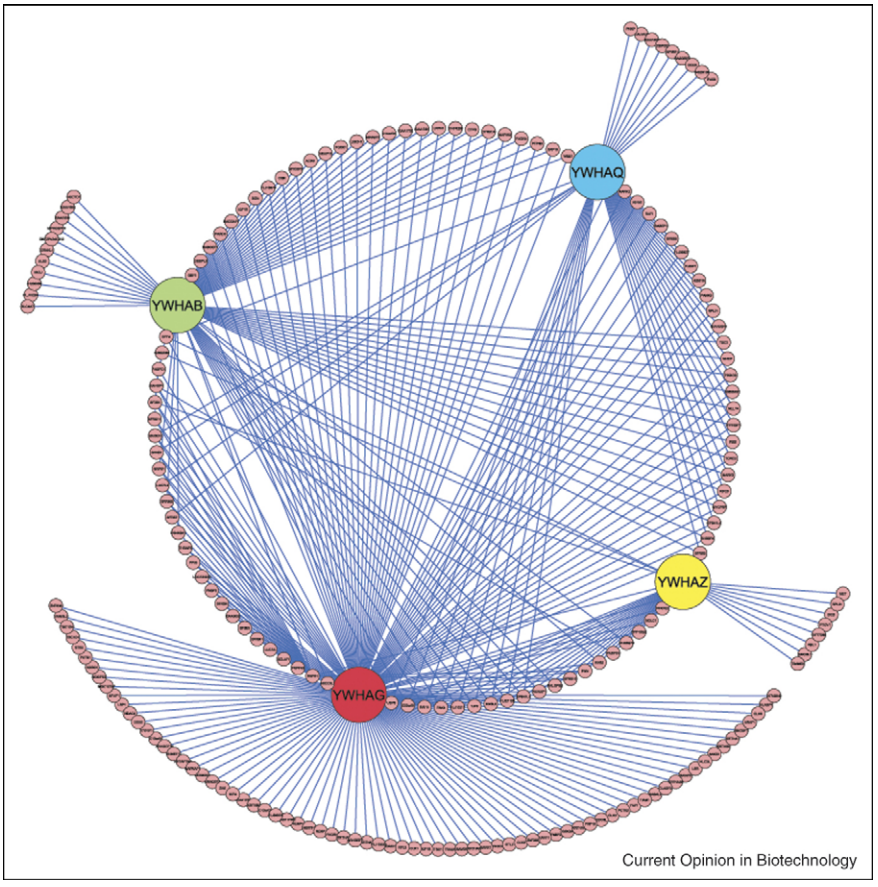
Ewing *et al.* [7••] performed the first large-scale analysis of protein complexes in human cells. 338 human bait proteins selected on the basis of known or suspected disease associations were Flag tagged and transfected into HEK293 cells. Large-scale single-step Flag tag purification of these proteins and analysis by tandem MS led to the identification of 24 540 PPIs amongst 2826 identified proteins. A number of data-filtering steps were used to improve the quality of the dataset (Figure 2), including the removal of bait–bait interactions and proteins identified in control purifications and proteins interacting with more than 5% of bait proteins, resulting in a final dataset of 6463 PPIs between 2235 proteins. This approach enabled high-quality protein interaction data to be generated from single-step protein complex purification thereby reducing the need for TAP. Additional analysis of the data and the development of a novel interaction confidence score

Figure 2



Pipeline for the characterisation of tagged-protein complexes. This generic workflow begins with the tagging of a protein, which can be achieved either by transfecting a tagged gene into cells or by targeting the endogenous gene using homologous recombination. Tagged-protein complexes can then be purified by a single-step or by two isolation steps depending on the design of the tag. Eluted protein complexes can then be fractionated by SDS-PAGE or protein chromatography. Alternatively, a solution digest of the eluted complexes can be performed and peptides can be directly analysed by LC–MS/MS or by 2D-LC–MS/MS depending on the sample complexity. SDS-PAGE lanes are usually cut into slices and proteins digested with trypsin (or other protease) and peptides extracted. Each peptide fraction is analysed by LC–MS/MS (or MALDI-MS) in which typically the top 3 (up to 10) most intense precursor ions are selected for MS/MS fragmentation for every MS scan. Next, raw spectra are processed and are submitted for sequence database searching (e.g. MASCOT, SEQUEST, etc.) and spectra matching with sequences in the database are scored (by mass accuracy, fragment ion matches, etc.) and filtered to obtain high-confidence peptide identifications. Peptides are then clustered to protein sequences and scores assigned to protein identifications. Usually, protein identifications are clustered to remove redundancy and common contaminant protein identifications (Keratin, trypsin, TEV, etc.) as well as proteins identified in the control purifications are subtracted from the list of proteins identified in the tagged-protein purification. Quantitative MS data can be used to assist differentiation of contaminant proteins from transient protein interactions.

Figure 3



Overlapping but distinct interactomes of 14-3-3 isoforms. 14-3-3 primary interactions identified in a large-scale Flag-tagging study in human cells [7••]. 14-3-3 family proteins have a highly conserved protein interaction domain that binds to short linear phosphorylated peptide sequences. The conservation of this binding domain would suggest that protein interactions would be conserved between 14-3-3 family members. Analysis of protein interactions identified for four of these proteins shows a large network of common protein interactions as well as dimerisation of 14-3-3 proteins. However, a large number of protein interactions are specific to 14-3-3 family members, particularly 14-3-3 gamma (YWHAG), which specifically interacts with over 70 proteins, suggests either a more complex mode of binding specificity or differences in the abundance of 14-3-3 family members.

(which based on uses a number of parameters including MASCOT score, number of identified peptides and rank of the prey protein) allowed the authors to further validate protein interactions. A number of distinct well-connected protein complexes were characterised, including the components of the proteasome, spliceosome, chromatin remodelling complexes and 14-3-3 complexes. Comparison of 14-3-3 isoforms not only highlighted the overlapping specificity of these proteins but also led to the identification of many isoform-specific protein interactions indicating the sensitivity and resolution of their biochemical and computational pipeline for protein interaction mapping (Figure 3). Other approaches have been developed to discriminate between real and non-specific interactors in epitope-tagged-protein complexes. Rinner *et al.* [20] used label-free MS quantification (based on the basis of MS1 peak area calculations) to analyse serial dilutions of epitope-tagged-protein complexes spiked into control purifications and showed that while *bone fide* interactors increased in abundance, non-specific interactors remained relatively constant across the dilution series. Similarly, Sardiu *et al.* [21] used normalised spectral abundance factors (NSAFs) to identify proteins that were enriched in epitope-tagged purifications and hierarchical clustering of NSAF values in reciprocal pull-downs permitted core complexes and modules to be defined.

Mapping transient protein interactions

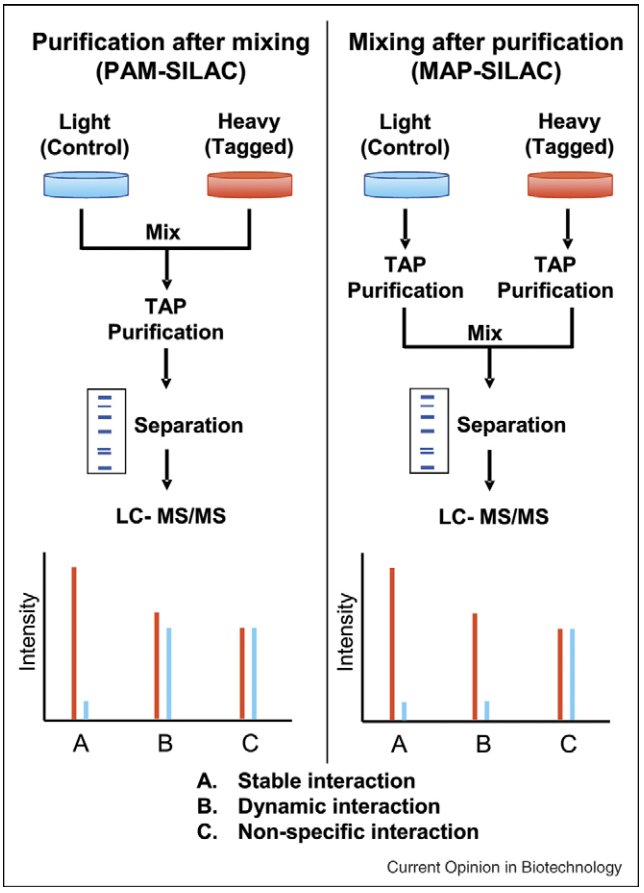
Maintenance and identification of transiently interacting proteins during affinity purification remains a challenge. These dynamic protein interactions are likely to be important in the context of signalling pathways where modulation of signalling can occur rapidly. It is thought that such interactions can be lost during the longer purification times necessary for tandem purifications, while single-step purifications may retain these inter-

actions to some degree. However, decreased purity of one-step methods may hinder identification of low abundance transient interactions. Guerrero *et al.* [22] developed a strategy called QTAX (quantitative analysis of tandem affinity purified *in vivo* cross-linked (X) protein complexes) to discover stable and transient interactions in the yeast 26S proteasome complex. The combination of a novel tag (HB tag) that contains a bacterially derived peptide that induces biotinylation *in vivo*, flanked by two hexahistidine tags (Figure 1D) and *in vivo* formaldehyde cross-linking allowed fully denaturing TAP of the complex [23]. Quantitative LC–MS/MS analysis of SILAC-labelled, tagged and control cells allowed non-specific interactions to be identified and excluded from the 26S proteasome interaction network. This approach identified 64 specific proteasome-interacting proteins (42 of which were novel) and represents an efficient and stringent method for capturing transient protein interactions.

Recently, complementary approaches that enable distinction between stable and dynamic interactions have been developed [24•,25•]. One approach consists of two types experiments: one, termed time course-purification after mixing-SILAC (Tc-PAM-SILAC) is a time course experiment with a range of incubation times with protein complex purification from mixed heavy and light SILAC-labelled cell extracts and another type of purification is termed mixing after purification-SILAC (MAP-SILAC) (Figure 4). This MAP-SILAC strategy is different to standard protocols because the tagged and control (heavy versus light) purifications are performed separately and eluted material is then pooled for digestion. When applied to the 26S proteasome, the time course experiments revealed that SILAC ratios for some proteins increased with shorter incubation times because longer incubations could lead to mixing of heavy and light forms from the tagged and control purifications, respectively. In order to avoid this mixing problem, purifications were performed separately and only pooled at the digestion stage. This allowed the identification of 35 stable subunits and 16 proteins with MAP-SILAC ratios >2-fold higher than Tc-PAM-SILAC, suggesting that these are dynamic interactors. Of these 16, 8 displayed an increased ratio with decreased incubation time in Tc-PAM-SILAC experiments, thus providing additional evidence of their dynamic nature. Half of the identified dynamic interactors are involved in the ubiquitin–proteasome degradation pathway with associated functions that are dynamic in nature.

Mousson *et al.* [25•] recently described a similar strategy for the identification of transient protein interactions. Flag tagged and HA tagged TATA-binding protein (TBP) complexes were purified from SILAC-labelled HeLa cells using a strategy comparable to MAP and PAM-SILAC [24•] (Figure 4). When tagged TATA-binding protein complexes were purified from mixed, tagged (heavy) and control (light) cell lysates, one protein (BTAF) contained 40% light peptides that had dissociated from the control purification and re-associated with the tagged TBP complex. This protein would have been considered a non-specific interactor from this strategy alone, but purification of tagged complexes and control samples separately (PAM-SILAC) revealed that BTAF was indeed a dynamic interactor. Additionally, purification of BTAF complexes from heavy isotope-labelled cells spiked with increasing amounts of unlabelled recombinant BTAF showed increasing amounts of light peptides from the recombinant protein, further demonstrating the dynamic nature of this interaction.

Figure 4



A quantitative MS strategy for the identification of dynamic protein interactions. Recently, it has been found that the classical PAM-SILAC strategy in which a TAP experiment is performed on mixed, tagged and control (heavy and light) cell lysates does not allow obvious distinction between dynamic and non-specific interactions because of mixing of heavy and light forms of proteins during the purification [24•,25•]. When TAP experiments are performed separately and eluted complexes mixed (MAP-SILAC) after purification, signal intensity differences for dynamic interactors are more discriminating. This approach allows enhanced resolution of dynamic and non-specific interactions and has been successfully applied to identify dynamic components in TATA-binding protein (TBP) transcription complexes [25•] and 26S proteasome complexes [24•].

Architecture of multiprotein complexes

Hernández *et al.* [26•] combined TAP and specialised MS analysis of intact non-covalent complexes to investigate subunit architecture of selected yeast complexes. Highly selective TAP experiments were performed to yield pure complexes that are necessary for top-down MS analysis. The TAP protocol was modified to remove detergents before the last purification step and buffer exchange and concentration steps were necessary to make purified complexes compatible for intact MS analysis. Additionally, protein complexes were disrupted to generate subcomplexes in solution (chemical additives) or by in-source collision-induced dissociation (CID). MS/MS analysis was performed on digested complexes to generate sequence data and thus protein identifications and MS was performed on intact complexes, subcomplexes and individual proteins from in-source CID to calculate molecular masses. This approach was applied to the exosome complex, which is composed of 10 subunits, with a predominant species of 398 kDa. Analysis of three different tagged components of the complex and of subcomplexes allowed the generation of a detailed structural interaction map of exosome subunits. This coupling of TAP tagging and intact MS analysis of protein complexes is an exciting development and it will be interesting to see how well this approach translates to other protein complexes, particularly larger and more heterogeneous complexes.

Alber *et al.* [27,28••] reported an integrative investigation of the constituents of the yeast nuclear pore complex (NPC), a 50 MDa protein assembly that regulates nucleocytoplasmic transport. This complex is composed of an assembly of 456 protein molecules that constitute around 30 distinct nucleoporins [29,30]. The authors took a six-layered hierarchical approach to interrogate and computationally infer the structure of the NPC, as no current single experimental method is sufficient to describe the structure of such a large protein complex. The NPC component list, component stoichiometry, approximate shape and size of each component, overall size shape and symmetry of the NPC, localisation of components in the NPC and finally subcomplex structure, were all used to derive the most complete picture of this protein assembly to till date. The complex substructure was interrogated using overlay assays to determine pairs of interacting proteins and affinity purification to identify proteins interacting directly and indirectly with bait proteins. Protein A tagged bait proteins were used to purify complexes from whole cell lysates as well as different cell fractions using 20 different variants of extraction buffers to generate different kinds of complexes. Only Coomassie-stained bands were analysed by MS to ensure that the focus was on the most abundant proteins in complexes and hence proteins that are most closely associated with bait proteins. This tagging approach generated 73 composites or sub complexes, varying in complexity from dimers to those containing 20 proteins. Purification of complexes under different conditions allowed insight into the hierarchical assembly of components with certain bait proteins found as dimers, trimers, tetramers, pentamers, hexamers and full septameric complexes [28••]. Data from affinity purifications allowed protein proximities to be inferred and when combined with other data types provided substructure resolution and refinement of the overall model of the NPC.

Conclusions

Although the analysis of protein complexes is still not routine, significant improvements in tagging, purification and analysis of protein complexes have been made. In particular, the ability to perform large-scale protein interaction mapping in mammalian cells should produce biological insights that are not achievable using other methods. We envisage that quantitative MS-based approaches and stable isotope labelling will be increasingly used for exploring dynamics of protein complexes involved in specific biological processes, enabling a transition from descriptive to functional pathway analysis.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM *et al.*: **Functional organization of the yeast proteome by systematic analysis of protein complexes**. *Nature* 2002, **415**:141-147.

2. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K *et al.*: **Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry**. *Nature* 2002, **415**:180-183.

3. •• Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B *et al.*: **Proteome survey reveals modularity of the yeast cell machinery**. *Nature* 2006, **440**:631-636.

TAP-MS strategy in which protein complexes associated with 1754 yeast bait proteins were characterised. 2760 protein identifications were produced and computational analysis on the basis of the raw protein identification data allowed a core set of 1483 proteins organised into 491 distinct complexes to be defined.

4. •• Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP *et al.*: **Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae***. *Nature* 2006, **440**:637-643.

TAP-MS strategy in which protein complexes associated with 2357 yeast bait proteins were characterised. 4087 protein identifications were produced and the use of a machine-learning algorithm on the basis of published protein interaction data allowed a core set of 2708 proteins organised into 547 distinct complexes to be defined.

5. Arifuzzaman M, Maeda M, Itoh A, Nishikata K, Takita C, Saito R, Ara T, Nakahigashi K, Huang HC, Hirai A *et al.*: **Large-scale identification of protein-protein interaction of *Escherichia coli* K-12**. *Genome Res* 2006, **16**:686-691.

6. Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N *et al.*: **Interaction network containing conserved and essential protein complexes in *Escherichia coli***. *Nature* 2005, **433**:531-537.

7. •• Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M *et al.*: **Large-scale mapping of human protein–protein interactions by mass spectrometry**. *Mol Syst Biol* 2007, **3**:89.

This is the first large-scale analysis of protein complexes in human cells. 338 Flag tagged human bait proteins of known or suspected disease associations were used to identify 24 540 PPIs amongst 2826 identified proteins. Further computational analysis produced a final dataset of 6463 PPIs between 2235 proteins and the development of a novel interaction confidence score allowed ranking of protein interactions.

8. Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S *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.

9. Major MB, Camp ND, Berndt JD, Yi X, Goldenberg SJ, Hubbert C, Biechele TL, Gingras AC, Zheng N, Maccoss MJ *et al.*: **Wilms tumor suppressor WTX negatively regulates WNT/beta-catenin signaling**. *Science* 2007, **316**:1043-1046.

10. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J *et al.*: **Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme**. *Mol Cell* 2007, **27**:262-274.

11. Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH: **A protein interaction network for pluripotency of embryonic stem cells**. *Nature* 2006, **444**:364-368.

12. 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**:1030-1032.

13. • Burckstummer T, Bennett KL, Preradovic A, Schutze G, Hantschel O, Superti-Furga G, Bauch A: **An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells**. *Nat Methods* 2006, **3**:1013-1019.

The efficiency of this yeast TAP tag was compared with that of three other TAP tags. A GS-TAP (Protein G and SBP combination) strategy produced a 10-fold increase in the efficiency compared to the conventional TAP tag.

14. Gloeckner CJ, Boldt K, Schumacher A, Roepman R, Ueffing M: **A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes**. *Proteomics* 2007, **7**:4228-4234.

15. Poser I, Sarov M, Hutchins JR, Heriche JK, Toyoda Y, Pozniakovsky A, Weigl D, Nitzsche A, Hegemann B, Bird AW *et al.*: **BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals**. *Nat Methods* 2008, **5(5)**:409-415.

16. Angrand PO, Segura I, Volkel P, Ghidelli S, Terry R, Brajenovic M, Vintersten K, Klein R, Superti-Furga G, Drewes G *et al.*: **Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling**. *Mol Cell Proteomics* 2006, **5**:2211-2227.

17. Brajenovic M, Joberty G, Kuster B, Bouwmeester T, Drewes G: **Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network**. *J Biol Chem* 2004, **279**:12804-12811.

18. Goll J, Uetz P: **The elusive yeast interactome**. *Genome Biol* 2006, **7**:223.

19. Orchard S, Salwinski L, Kerrien S, Montecchi-Palazzi L, Oesterheld M, Stumpflen V, Ceol A, Chatr-aryamontri A, Armstrong J, Woollard P *et al.*: **The minimum information required for reporting a molecular interaction experiment (MIMIx)**. *Nat Biotechnol* 2007, **25**:894-898.

20. Rinner O, Mueller LN, Hubalek M, Muller M, Gstaiger M, Aebersold R: **An integrated mass spectrometric and computational framework for the analysis of protein interaction networks**. *Nat Biotechnol* 2007, **25**:345-352.

21. Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP: **Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics**. *Proc Natl Acad Sci U S A* 2008, **105**:1454-1459.

22. Guerrero C, Tagwerker C, Kaiser P, Huang L: **An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network**. *Mol Cell Proteomics* 2006, **5**:366-378.

23. Tagwerker C, Flick K, Cui M, Guerrero C, Dou Y, Auer B, Baldi P, Huang L, Kaiser P: **A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking**. *Mol Cell Proteomics* 2006, **5**:737-748.

24. • Wang X, Huang L: **Identifying dynamic interactors of protein complexes by quantitative mass spectrometry**. *Mol Cell Proteomics* 2008, **7**:46-57.

Quantitative MS analysis of the 26S proteasome was performed using SILAC to identify stable and dynamic interactions. Time course experiments revealed proteins with increased ratios in the bait purifications compared to controls with decreased incubation time. In addition, performing bait and control purifications separately and mixing after purification allowed dynamic interactions to be identified more readily compared to traditional SILAC experiments where the bait and control cell lysates are mixed before purification.

25. • Mousson F, Kolkman A, Pijnappel WW, Timmers HT, Heck AJ: **Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes**. *Mol Cell Proteomics* 2008, **7(5)**:845-852.

Quantitative MS analysis of TATA-binding protein (TBP) complexes was performed using SILAC to identify stable and dynamic interactions. In addition, performing bait and control purifications separately and mixing and after purification allowed dynamic interactions to be identified more readily compared to traditional SILAC experiments where the bait and control cell lysates are mixed before purification.

26. • Hernandez H, Dziembowski A, Taverner T, Seraphin B, Robinson CV: **Subunit architecture of multimeric complexes isolated directly from cells**. *EMBO Rep* 2006, **7**:605-610.

Direct MS analysis of intact TAP-tagged non-covalent complexes, subcomplexes and MS/MS analysis of individual proteins allowed the subunit architecture of the yeast exosome to be described.

27. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT *et al.*: **The molecular architecture of the nuclear pore complex**. *Nature* 2007, **450**:695-701.

28. •• Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT *et al.*: **Determining the architectures of macromolecular assemblies**. *Nature* 2007, **450**:683-694.

Integrative investigation of the constituents and structure of the yeast nuclear pore complex (NPC). The complex substructure was analysed by affinity purification to identify proteins interacting directly and indirectly with bait proteins. This approach generated 73 composites or subcomplexes, varying in complexity from dimers to those containing 20 proteins and these data were combined with other data sources to describe the architecture of the NPC.

29. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT: **The yeast nuclear pore complex: composition, architecture, and transport mechanism**. *J Cell Biol* 2000, **148**:635-651.

30. Cronshaw JM, Krutchinsky AN, Zhang W, Chait BT, Matunis MJ: **Proteomic analysis of the mammalian nuclear pore complex**. *J Cell Biol* 2002, **158**:915-927.