Journal of Chromatography A, 1217 (2010) 7661–7668

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Coupling protein complex analysis to peptide based proteomics

Qiang Gao[a], Ashraf G. Madian[a], Xiuping Liu[a], Jiri Adamec[b], Fred E. Regnier[a,*]

[a] Department of Chemistry, Purdue Univeristy, West Lafayette, IN 47907, USA
[b] Bindley Bioscience Center, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 22 June 2010
Received in revised form
23 September 2010
Accepted 27 September 2010
Available online 29 October 2010

Keywords:
Protein complex
Size-exclusion chromatography
Proteomics
Hub proteins

ABSTRACT

Proteolysis is a central component of most proteomics methods. Unfortunately much of the information relating to the structural diversity of proteins is lost during digestion. This paper describes a method in which the native proteome of yeast was subjected to preliminary fractionation by size exclusion chromatography (SEC) prior to trypsin digestion of SEC fractions and reversed phase chromatography-mass spectral analysis to identify tryptic peptides thus generated. Through this approach proteins associated with other proteins in high molecular mass complexes were recognized and identified. A focus of this work was on the identification of Hub proteins that associate with multiple interaction partners. A critical component of this strategy is to choose methods and conditions that maximize retention of native structure during the various stages of analysis prior to proteolysis, especially during cell lysis. Maximum survival of protein complexes during lysis was obtained with the French press and bead-beater methods of cell disruption at approximately pH 8 with 200 mM NaCl in the lysis buffer. Structure retention was favored by higher ionic strength, suggesting that hydrophobic effects are important in maintaining the structure of protein complexes. Recovery of protein complexes declined substantially with storage at any temperature, but storage at −20 °C was best when low temperature storage was necessary. Slightly lower recovery was obtained with storage at −80 °C while lowest recovery was achieved at 4 °C. It was concluded that initial fractionation of native proteins in cell lysates by SEC prior to RPC-MS/MS of tryptic digests can be used to recognize and identify proteins in complexes along with their interaction partners in known protein complexes.

1. Introduction

Genomics has aided proteomics enormously by predicting the sequence, size, and complexity of a proteome. We know from the *Saccharomyces cerevisiae* genome for example the probable sequence of proteins expressed by yeast and that the proteome is composed of approximately 6000 polypeptides [1,2]. DNA databases have become a cornerstone of proteomics for this reason. But unfortunately, there are still many features of a proteome that are not predictable from genomics. This is a major issue in proteomics today.

Proteome diversity is far greater than suggested by the genome [3], probably by an order of magnitude. Alterations in gene expression by microRNA species [4–6], alternative splicing [7], and more than a hundred different types of post-translational modification (PTM) produce a wide array of protein variants not suggested by genomics [8]. Quaternary structure is a similar case. Interaction partners of many proteins are not suggested directly by a genome. These features are of major importance in gaining a fundamental understanding of how biological systems are regulated and go awry in disease states.

This means that much of proteome diversity must be analyzed and elucidated experimentally. Unfortunately current methods are not well suited for the study of structural diversity, even in a case as simple as protein complexes. Quaternary structure analysis is not possible with most of the global proteomic methods except by direct mass spectral (MS) analysis. Unfortunately direct MS methods require prior purification of the complex and do not work well with complexes in the million dalton range.

Proteins complexes occur widely in cells where they are sometimes referred to as the cellular '*interactome*' [9]. These complexes are of distinct structure and function [10], being fundamental to cellular regulation and survival [11]. Up to 40% of the proteins in yeast are probably part of one or more protein complexes at some time [12]. Unfortunately, methods for rapidly isolating large numbers of complexes and elucidating their composition do not exist. Concerted efforts have been made during the past decade to study protein complexes on an individual basis along with attempts to understand how each is impacted by external stimuli [13–15], but this one-at-a-time strategy is both lengthy and labor intensive.

* Corresponding author. Tel.: +1 7654961132; fax: +1 7654940239.
E-mail address: fregnier@purdue.edu (F.E. Regnier).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.09.071

Methods for studying an interactome on a global scale lag far behind current proteomic methods in speed and universality.

Protein complexes are generally characterized by a combination of genetic, biochemical, or biophysical techniques. Phage display, synthetic genetic array (SGA) analysis [16], protein chips [17], and the yeast two hybrid (Y2H) methods have been widely used to study protein–protein interactions on a binary basis. Among these methods the Y2H method is the more widely used, the advantage of this method being that it allows weak or transient interaction partners to be recognized *in vivo*. A weakness is that it requires a substantial degree of genetic manipulation and it only sees binary interactions, not all the partners in a complex simultaneously. Moreover, the Y2H method requires that interactions take place in the nucleus. This reduces efficacy further. Interaction partners of proteins anchored in a membrane cannot be detected as well. Also, transcription factors that interact with DNA directly will generate "false positive" interactions [18].

Tandem affinity purification (TAP) [19] is a separation method in which affinity selectable polypeptide tags are added to putative members of a complex by genetic manipulation at either their amino- or carboxy-termini. The tagged proteins are then sequentially selected in two affinity chromatography steps, bringing with them other members of the complex [20]. Tags are designed to interact with an affinity column weakly so the complex can be eluted without dissociation of other members in the complex. The advantage of the TAP method over the Y2H method is that, under a given set of conditions, all high affinity interaction partners in a complex are pulled down in a single experiment. Members of a complex thus resolved can then be identified by modern proteomic methods. An unfortunate limitation of this method is that it suffers from poor reproducibility [21]. Additional issues are (i) the complexity of genetically constructing large numbers of tagged proteins, (ii) the possibility that tags will interfere with complex formation, (iii) difficulty of experimentation in human subjects, and (iv) the fact that weakly bound proteins dissociate from the complex during the multiple steps of affinity chromatography.

Free-flow electrophoresis (FFE) [22], blue native gel electrophoresis [23], and chromatographic methods other than TAP are more global in that they allow separation of multiple complexes simultaneously. It has recently been shown that more than a third of all proteins in yeast lysates elute from size exclusion chromatography (SEC), anion exchange chromatography (AEC), and hydrophobic interaction chromatography (HIC) columns in high molecular mass protein complexes [24]. Molecular mass in some cases extends to 10^6 Da. Simple fractionation methods are a requisite first step in developing global methods for protein complex analysis.

Among the separation methods examined to date, SEC is one of the most gentle. Separation of proteins complexes does not require surface association and no motive forces are involved that pull complexes apart. It has been shown with yeast that more than a third of all proteins in lysates elute from an SEC column at much higher apparent molecular mass than predicted from DNA databases [23]. This is interpreted to mean that proteins are migrating through the system in complexes. But even with SEC there could be some dissociation of proteins from complexes during elution. This probably results from dilution that occurs subsequent to cell lysis and during migration through SEC columns. Moreover, pH and ionic strength may contribute as well.

Recently, Paul et al., combined SEC chromatography with high accuracy mass spectrometry LTQ-Orbitrap to study the megadalton complexes with Mr up to 5 MDa in Arabidopsis [25]. The proteins eluted from SEC chromatography were quantified by spectral counting and further grouped to distinct protein functional assemblies by utilizing hierarchical clustering and protein heat maps. This method successfully extends the application of SEC chromatography in resolving large protein complexes like chloroplast ribosome associated complexes.

The extreme in interacting protein species is "hub proteins". These are proteins that may have a hundred or more interaction partners. Hub proteins play a critical role in organizing protein networks [26]. Their function is to arrange and hold numerous, less-connected proteins together in specific complexes. Finding hub proteins has generally involved sophistic statistical methods that require the analysis of large amounts of TAP and Y2H data. It would be desirable to have a method that rapidly recognizes a hub protein.

There is great interest today in how environmental stimuli impact an interactome. Changes in the concentration and distribution of proteins across multiple protein complexes as a function of stimulus induced alterations in expression and post-translation modification must be assessed to evaluate the biodynamics of an interactome. A goal of the work described here was to determine how SEC could be incorporated into global proteomics methods in such a way that proteins in complexes could be recognized, identified, and compared with interaction partners identified by other methods. Particular attention was focused on recognizing hub proteins and their interaction partners.

2. Materials and methods

2.1. Chemicals

Yeast nitrogen base (YNB) without amino acids and terrific broth, agar, yeast extract, and peptone were purchased from BD Bioscience (Franklin Lakes, NJ. USA). Sodium chloride, glycine, trifluoroacetic acid (TFA), HPLC grade acetonitrile, and D-glucose were obtained from Mallinckrodt Chemicals (Phillipsburg, NJ, USA). Iodoacetic amide (IAA), L-cysteine, α-cyano-4-hydroxy-cinnamic acid and Trizma base were supplied from Sigma Chemical Company (St. Louis, MO. USA). Sodium dodecyl sulfate (SDS), dithiothreitol (DTT) and urea were purchased from Bio-Rad Laboratories (Hercules, CA, USA). A proteomics analyzer calibration mixture (4700 Cal Mix, bradykinin, angiotensin I, glu[1]-fibrinopeptide B, ACTH fragment 1–17, ACTH fragment 18–39, and ACTH fragment 7–38) was purchased from Applied Biosystem Inc. (ABI) (Foster City, CA. USA). An SEC calibration standard (blue dextran, thyroglobulin, ferritin, catalase, aldolase, bovine serum albumin, ovalbumin, ribonuclease A, and chymotrypsinogen) was supplied by Amersham Biosciences (Uppsala, Sweden). Proteomic sequencing grade trypsin was obtained from Promega (Madison, WI, USA).

2.2. Yeast strains and culture conditions

S. cerevisiae strains BY4741 were obtained from Open Biosystems (Huntsville, AL. USA). Yeast used in all the experiments described in this paper was grown on YEPD media (1% yeast extract, 2% peptone, 2% glucose). Inoculate from a yeast culture grown on YPD plates for 36 h at 30 °C was transferred to 5 ml of YEPD medium and incubated on a 250 rpm shaker overnight. The culture was then transferred to 250 ml of medium and incubated 6–8 h until the culture reached exponential growth at O.D. 1.0–1.2. After harvesting by centrifugation for 10 min at 8000 rpm, the cells were washed twice with 1×PBS (pH 7.4) buffer and then stored in −80°.

2.3. Sample preparation

Yeast pellets were lysed in three ways: with a French press, by the bead-beater method, and with Roche chemical lysis buffer. For the French press lysis method, cells in lysis buffer (50 mM Tris pH 7.5, 100 mM NaCl, 0.5% NP40 and protease inhibitor from Roche Applied Science were held in the press for 3 min at 1300–1500 psi.

This lysis protocol was repeated 3 times on each sample. Bead-beater lysis was achieved by suspending cells in cold lysis buffer (4 °C) in a glass bead mill (Bead-beater, Biospec) during 10 alternating cycles of 1 min agitation and 2 min cooling. During lysis, cells were kept cold in an ice bath to avoid denaturation of protein complexes. With the Roche chemical lysis method, 1 ml of complete lysis medium from Roche Applied Science (Indianapolis, IN, USA) was added to the yeast pellets and kept on ice for 30 min during the lysis process. Cell debris was removed from lysates by centrifugation at 14,000 rpm (4 °C) for 15 min.

2.4. SEC calibration

The SEC column was calibrated by protein calibration standards divided into two groups. The group I standard contained thyroglobulin (Mr 670 kDa), catalase (Mr 220 kDa), bovine serum albumin (Mr 67 kDa), ribonuclease A (Mr 13.7 kDa), vitamin B_{12} (Mr 1355 kDa), and cytidine (Mr 243 kDa). The group II standard contained ferrintin (Mr 440 kDa), aldolase (Mr 160 kDa), ovalbumin (Mr 45 kDa), chymotrypsinogen (Mr 25 kDa), vitamin B_{12} (Mr 1355 kDa), and cytidine (Mr 243 kDa). Calibration standard samples were run on the SEC column with a Tris (50 mM, pH 7.5) mobile phase at a flow rate of 0.5 ml/min.

2.5. SEC and 1D SDS-gel analysis of proteins

Yeast cells were resuspended in 1.5 ml lysis buffer (150 mM NaCl, 1.9 mM NaH_2 PO_4, 8.1 mM Na_2H PO_4, 1% Triton-100, 1× protease inhibitor (Roche)). Five hundred microliters of acid washed glass beads were mixed with yeast and vortexed 7 times for 30 s at 2 min interval while maintaining the temperature at 4 °C. The yeast lysate was then transferred to another tube and centrifuged for 20 min at 15,000 rpm and 4 °C. The supernatant was collected and a 500 μl (~13 mg/ml) sample placed on a Superdex 200 (300×100) SEC column that was eluted with a 50 mM Tris mobile phase (pH 7.5) at 0.5 ml/min using a Beckman ProteomeLab PF 2D LC pumping system (Fullerton, CA). Protein elution was monitored by absorbance at 280 nm. Eluent fractions from SEC were concentrated with a micro spin column (Millipore Ultrafree) and separated by SDS 4–12% gradient gel electrophoresis (Bio-Rad, CA, USA) using the recommended Bio-Rad running buffer. The gel was stained with coomassie blue using a Colloidal Blue Staining Kit (Invitrogen). Each lane of SDS-gel was cut into 5–8 slices and trypsin (Promega, WI) was infused into the gel slices for proteolysis. The gel pieces were then cut into 1 mm^3 cubes and washed 3–5 times with 50 mM ammonium bicarbonate and 50% acetonitrile. The gel cubes were dried in a speedvac for 30 min and 10 mM DTT/25 mM ABC was added to cover the gel pieces before incubation at 56 °C for 1 h. DTT solution was removed and 55 mM iodoacetamide added to cover the gel cubes. After washing 5 times with 25 mM ammonium bicarbonate using constant vortexing for 10 min the gel cubes were dried in a speedvac for 30 min. The gel pieces were re-hydrated with 33.3 ng/μl trypsin in 50 mM ammonium bicarbonate and incubate for 16 h at 37 °C for protein digestion. The supernatant was transferred to a new tube and the remaining gel pieces extracted twice with 60% acetonitrile/5% trifluoroacetic acid under sonication on an ice bath for 30 min. The extracts were combined, desalted using a Vydac C18 silica column (the Nest Group, Inc) and stored for mass spectral analysis.

2.6. Protein identification by nano-LC separation offline couple with MALDI in MS/MS mode

An Agilent 1100 series HPLC was used to fractionate peptide digests prior to mass spectrometry identification. Tryptic digests were injected directly into a C18 column (Zorbax 300sB-C18, 3.5 μm, 100 μm i.d., 15 cm length, Agilent Technologies, Santa Clara, CA) and eluted using a linear gradient of 2% B to 40% B in 45 min at a flow of 800 nl/min. Mobile phase A contained 100% H_2O, 0.1% trifluoroacetic acid. Mobile phase B was composed of 100% acetonitrile, 0.1% trifluoroacetic acid. Peptides were spotted onto a stainless steel plate combined with a MALDI matrix (α-cyano-4-hydroxycinnamic acid, 4 mg/ml in 60% ACN/0.1% TFA) delivered at 1.2 μl/min by using a Agilent 1100 micro fraction collector and calibration standards (sigma) were spotted at six calibration positions on the MALDI plate. Peptides were analysis with an ABI 4800 MALDI analyzer in the positive mode using a laser intensity of 4900 in the MS mode and 5100 in MS/MS mode, a mass range (m/z) from 600 to 4000, and a signal-to-noise ratio above 30. Fifteen of the strongest first dimension MS peaks were examined by MS/MS. All spectra files obtained in the MS and MS/MS modes were acquired automatically with the 4000 Explorer software, including monoisotopic masses (m/z) of the parent ions and their corresponding fragment ions, charge state (z), and ion intensity in spectra. The data collected was submitted for identification and quantification using the Pro Group TM algorithm (ABI) search engine, Protein Pilot software 2.0, and the Swiss-prot database. Searches assumed a mass accuracy of 50 ppm and 99% confidence level in identification as the minimum acceptance criterion. Proteins were identified based on the presence of at least two unmodified peptides from the same protein identified by the Pro Group algorithm at the 99% confidence level. Proteins identified in this analysis are listed according to their Swiss-Prot entry names and accession numbers.

3. Results

The work presented here explores the possibility that when integrated with mass spectrometry based proteomics of tryptic digests, preliminary fractionation of native protein mixtures by SEC provides a means to recognize and confirm the presence of proteins in intermolecular complexes within the workflow used by many proteomics laboratories. Critical to this approach is finding conditions that allowed protein complexes to survive the various steps of the analytical process before proteolysis and mass spectrometry.

3.1. Analytical protocol

A concern in the analysis of protein complexes is that they stay intact during the requisite analytical operations. As a consequence studies were carried out to find the most suitable conditions for analysis. The experiments described below examine the gross effects of various conditions and methods on the size exclusion chromatography (SEC) of protein complexes.

3.2. Cell lysis effects

Previous studies have indicated that the way cells are lysed can impact the stability of protein complexes [24]. Three different methods were examined in these studies, the bead-beater method, the French press, and chemical lysis with the Roche lysing buffer. Yeast lysates were fractionated by SEC using a 10 mm × 330 mm Superdex 200 SEC column eluted isocratically at 0.5 ml/min with 50 mM Tris–HCl (pH 7.5). Elution was monitored by absorbance at 280 nm (Fig. 1). The calibration curve above the SEC chromatograms in Fig. 1 was produced using protein standards. Values on this curve indicate the molecular mass of protein standards in kilodaltons. The resolution of SEC with the particle size used produces baseline resolution between proteins that vary twofold in molecular weight. Peak capacity of a 25 cm column packed with particles of 5 μm or less is ~12. We integrated the areas of the peaks eluted from SEC 14 to 20 min by the software Beckman PF2D and defined it as segment 1. We integrated the areas of the peaks eluted from SEC

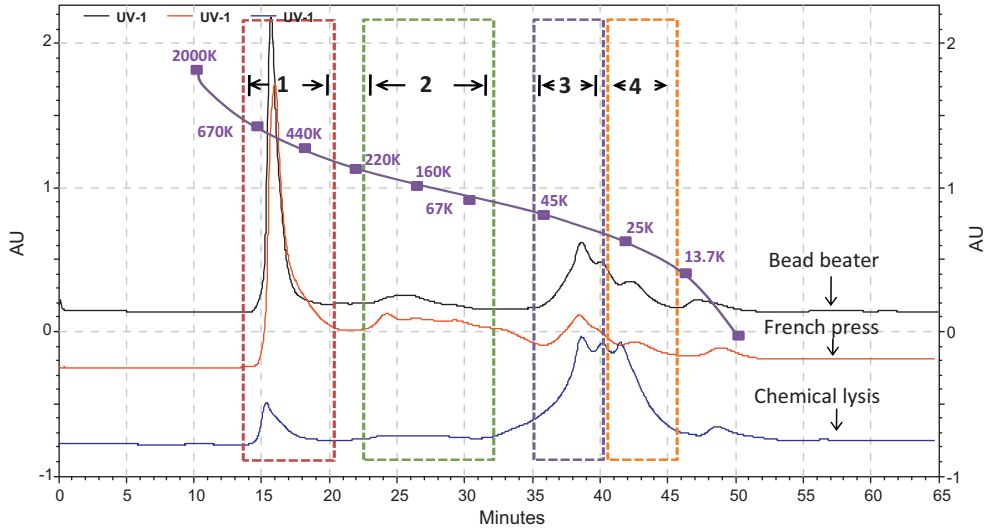

Fig. 1. SEC chromatograms of yeast lysates prepared by different lysis methods as labeled. Separations were achieved using a 10 mm × 330 mm Superdex 200 SEC column eluted isocratically at 0.5 ml/min with 50 mM Tris–HCl at pH 7.5. Protein elution was monitored by absorbance at 280 nm as seen on the left *y*-axis. The molecular mass curve above the SEC chromatograms is based on the elution time of standard proteins. Values on this curve are the molecular mass in kilodaltons of protein standards. Four segments, 1 through 4 as indicated in the SEC chromatogram were collected and used for further characterization.

22–32 min and defined it as segment 2. We integrated the areas of the peaks eluted from 35 min to 40 min and defined it as segment 3. We integrated the areas of the peaks eluted from SEC 40–46 min and defined it as segment 4.

It is clear from Figs. 1 and 2 that the chemical lysis agent from Roche allowed fewer high molecular mass complexes in the 300–700 kDa fraction (segment 1) to survive than the bead-beater method or the French press method of cell lysis. Among the three methods, the French press yielded slightly more proteins in the 60–200 kDa fraction (segment 2) while more proteins were found in the 13–45 kDa fractions (segments 3 and 4) with the bead-beater. Based on convenience of use the bead-beater was used in the remaining studies.

3.3. Mobile phase pH effects

Fortunately SEC can be carried out at almost any pH. The impact of mobile phase pH during SEC was examined at pH 7, 8 and 9 using 50 mM Tris and bead-beater lysed samples (Fig. 3). Essentially equivalent results were obtained with the pH 7 and 8 mobile phases except in the 20–40 kDa fraction. It is possible that the slightly lower amount of protein in the >300 kDa fraction is being seen in the larger 20–40 kDa fraction with the pH 7 mobile phase. Clearly pH 9 is not a good choice. Staying close to physiological pH is desirable in maintaining the structure of complexes as has been demonstrated with rhGCSF [27], insulin [28] and recombinant factor VIII SQ [29].

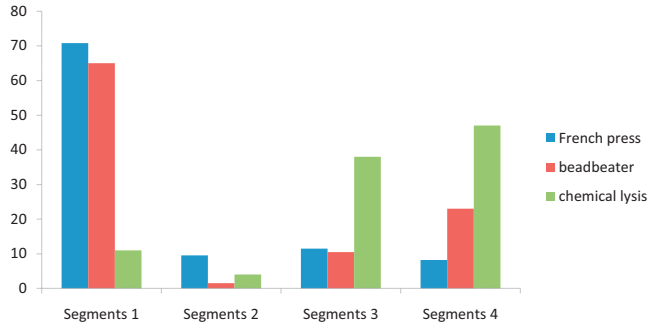

Fig. 2. Impact of cell lysis protocol on the protein complexes eluted from the SEC column. Samples and molecular mass segments being quantified are designated 1 through 4 in Fig. 1.

3.4. Ionic strength effects

Tertiary and quaternary structure of proteins is maintained by a balance between hydrophobic and electrostatic interactions. Electrostatic interactions are maximized at low ionic strength while hydrophobic effects are maximized at high ionic strength. The impact of ionic strength on protein complex stability was examined by SEC with mobile phases of different ionic strength. It was assumed in this approach that if ionic strength impacts protein structure, structural changes would be instantaneous and be reflected in SEC retention times.

Again SEC separations were carried out with a Superdex 200 column using a 50 mM Tris (pH 7.5) mobile phase at 0.5 ml/min. Mobile phase concentrations of NaCl ranging from 50 to 200 mM in 50 mM increments were examined. Clearly the 150 and 200 mM concentrations of NaCl were best in stabilizing high molecular mass complexes (Fig. 4). This suggests that hydrophobic interacts play a prominent role in maintaining the structure of these protein complexes. It also suggests that the high ionic strength mobile phases used in ion exchange and hydrophobic interaction chromatography

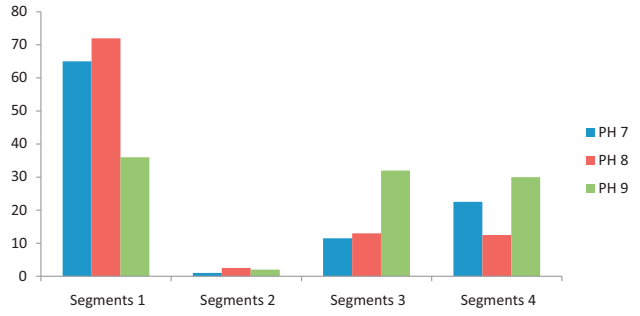

Fig. 3. Impact of mobile phase pH on protein complexes eluted from the SEC column. A 50 mM Tris–HCl mobile phase buffer adjusted to either pH 7, 8 or 9 was used at a velocity of 0.5 ml/min. The SEC column and chromatographic conditions were as in Fig. 1. Peak area integration was achieved with the software available on the Beckman LC system. Samples being examined in this case and the molecular mass segments being quantified are the same as those in Fig. 1.

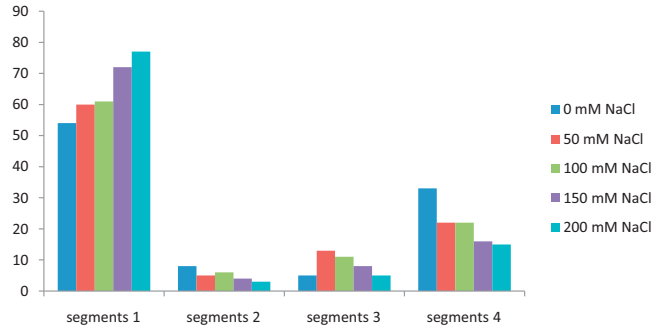

Fig. 4. Impact of salt effects on protein complexes eluted from the SEC column. The mobile phase was as in Fig. 1, but with added NaCl. The dark blue bar represents the relative areas of peaks in the SEC chromatogram with a mobile phase buffer containing no added NaCl. The red bar represents the relative area of peaks in the SEC chromatogram with 50 mM NaCl added to the mobile phase. The green bar is for 100 mM NaCl. The purple bar represents 150 mM NaCl addition and the light blue bar is for 200 mM NaCl. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

will minimally disturb the structure of many protein complexes if near neutral pH is used. Ion exchange and hydrophobic interaction separations of proteins has in fact been reported in a previous study [30].

3.5. Temperature effects

Occasionally it is necessary to store samples before analysis. Subsequent to storage for 24 h at a specific, fixed temperature, samples were examined by SEC. Three conditions of storage were examined and compared to a fresh control sample; storage at temperatures of 4 °C, −20 °C, and −80 °C. The results show that use of fresh samples is best while storage at 4 °C is the least desirable. Storage at −20 °C and −80 °C were of intermediate preference (Fig. 5).

3.6. Hub protein analysis

Although a number of hub proteins have been identified, alcohol dehydrogenase 1 (ADH1) was chosen for study because of its role as a hub protein in directing the formation of protein complexes that channel pyruvate into either the TCA cycle or the production of ethanol. The alcohol dehydrogenase family of isoenzymes has been

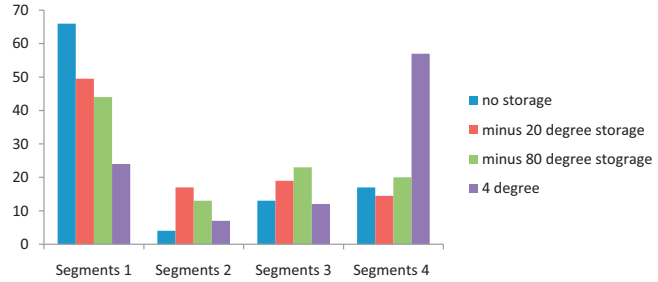

Fig. 5. Impact of storage temperature on protein complexes eluted from the SEC column. The blue bar represents the relative areas of peaks in SEC chromatogram of yeast lysate without any storage. The red bar represents the relative areas of peaks in SEC chromatogram of yeast lysate stored in −20 degree refrigerator and thawed in room temperature. The green bar represents the relative areas of peaks in the SEC chromatogram of yeast lysate stored in −80 degree refrigerator and thawed in room temperature after the cell is lysised. The purple bar represents the relative areas of peaks in SEC chromatogram of yeast lysate stored in 4 degree refrigerator after the cells were lysised. The SEC chromatography is done in Tris–HCl buffer pH 7.5 at a velocity of 0.5 ml/min. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

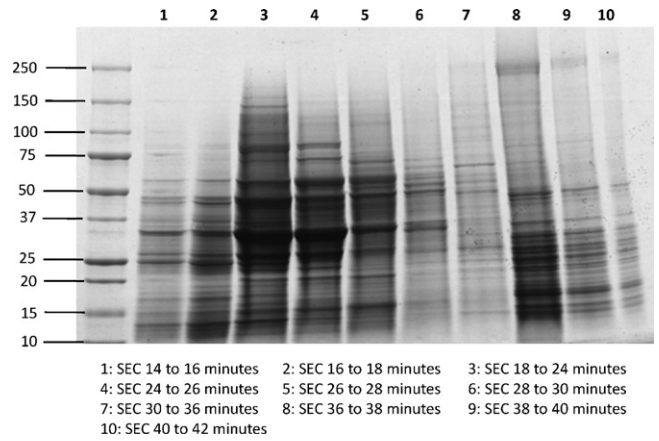

Fig. 6. SDS-gel picture of proteins eluted from SEC chromatography. The elution from SEC chromatography was concentrated and loaded to Bio-Rad 4–12% Tris–HCl gradient gel. The gel was stained by coomassie blue R250 and in gel tryptic digestion was performed. The tryptic peptides were eluted from the gel and further analyzed by orbitrap mass spectrometry.

reported to be hub proteins in a wide variety of organisms, ranging from yeast to humans. In yeast, five genes have been reported to encode enzymes responsible for ethanol metabolism, ADH1–ADH5 [31–34]. ADH1 is the major enzyme converting acetaldehyde to ethanol during glucose fermentation [35]. Overexpression of ADH1 can enhance the formaldehyde resistance of the yeast cells [36]. ADH1 is a tetramer, composed of four identical domains. Each domain is made up of a 347 amino acid polypeptide with a Mr of 36 kDa. Each domain contains two zinc ions. One zinc is primarily responsible for maintaining the protein structure and is coordinated with the sulfhydryl groups on cysteine residues 97, 100, 103 and 111, respectively [37]. The other zinc is essential for catalysis [38].

Fresh yeast lysates were used in these studies. Immediately after removal of particulates from lysates by centrifugation, supernatants were introduced into SEC columns. Fractions collected from the SEC column were concentrated and further fractionated by SDS-gel electrophoresis using a 4–12% gel gradient (Fig. 6). After coomassie blue staining, separation lanes from the gel were excised, cut into 5–8 sections and subjected to in-gel trypsin digestion. Peptide fragments were extracted from the gel slices and further fractionated by capillary reversed phase chromatography (RPC). Fractions of the RPC column effluent were spotted onto MALDI plates along with MALDI matrix. Peptide fractions thus derived were examined by matrix assisted laser desorption ionization (MALDI) mass spectrometry using an ABI 4800 TOF/TOF instrument. Peptide ions from the first dimension of TOF analysis were subjected to collision induced dissociation (CID) and the fragment ions formed in this process further resolved in a second dimension of TOF mass analysis. Amino acid sequences derived from these two dimensions of mass analysis were used to identify peptides and their protein parents. Approximately 900 proteins were identified in each of the SDS-gel fractions, accounting for a large portion of the yeast proteome. In total, 1830 proteins were identified from all the SEC fractions combined. In SEC fraction 14–16 min, 97 proteins were identified. In SEC fraction 16–18 min, 268 proteins were identified. In SEC fraction 18–24 min, 385 proteins were identified. In SEC fraction 24–26 min, 219 proteins were identified. In SEC fraction 26–28 min, 154 proteins were identified. In SEC fraction 28–30 min, 159 proteins were identified. In SEC fraction 30–36 min, 215 proteins were identified. In SEC fraction 36–38 min, 113 proteins were identified. In SEC fraction 38–40 min, 113 proteins were identified. In SEC 40–42 min, 107 proteins were identified. Essentially the same number of proteins were identified

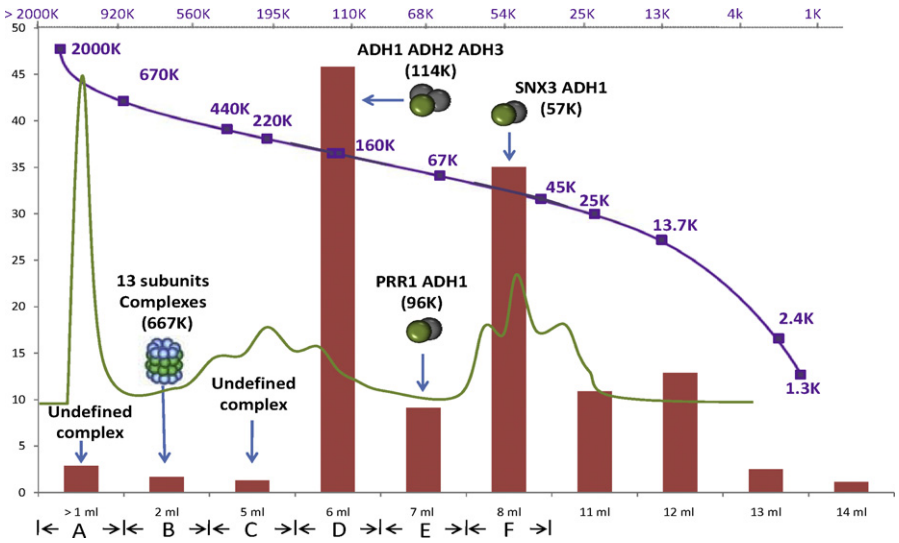

Fig. 7. Alcohol dehydrogenase 1 (ADH1) associated protein complexes identified from different fractions from SEC. Proteins eluted from the specified SEC fractions were tryptic digested and the peptides separated with a C18 column (Zorbax 300sB-C18, 3.5 μm, 100 μm i.d., 15 cm length, Agilent Technologies, Santa Clara, CA) using an Agilent 1100 HPLC and MALDI plate spotter. Elution from the RPC column was achieved using a linear gradient of 2% B to 40% B in 45 min. Spotted plates were analyzed by MALDI mass spectrometry using the ABI 4800. The intensities of peptides from ADH1 were averaged and plotted across all the SEC fractions. ADH1 binding partners were validated using the *Saccharomyces* Genome Database (SGD).

as in the shotgun method [39]. Around 539 proteins were overlapped between these two studies.

Elution of proteins from the SEC column spanned an 8 ml (16 min) volume window from 7 to 15 ml (14–30 min) (Fig. 1) with an exclusion limit slightly greater than a million daltons. ADH1 was found in a number of fractions, one being a very high molecular mass fraction (theoretical Mr >2000 kDa). ADH1 has not been previously reported to be in a complex of this size. The function of this complex is unknown.

ADH1 was also found in the 16–18 min fraction (theoretical Mr ~600 kDa) along with its interaction partners AOS1, ATG18, ECM1, ESA1, FAR11, FTH1, FUN19, ISW2, LEU1, MED7 and MYO4 (Fig. 7). This complex has previously been reported using the TAP method of structure analysis [40]. However, some of the interaction partners reported with the TAP method were not seen in this work. This could have occurred for either biological reasons or suppression of ionization. It is common that with some proteins insufficient numbers of peptides are ionized to allow unequivocal identification of the protein. Also, the abundance of some interaction partners was higher in the TAP based studies. The reason for this discrepancy is unclear.

The 24–26 min fraction also contained ADH1 along with interaction partners ADH2 and ADH3 (theoretical Mr ~177 kDa). This complex too has been reported using the TAP method [41]. These three proteins are part of a protein complex that channels pyruvate from glycolysis into energy production pathways.

ADH1 along with the interaction partner PRR1 was found in the 26–28 min fraction as well. PRR1 is a serine/threonine kinase that when complexed with ADH1 has a molecular mass of approximately 96 kDa according to the TAP method [42]. The detailed function of this complex is unknown, but it is known that ADH1 is co-localized with PRR1 *in vivo* and they might be involved in signaling transduction downstream in the MAPK kinase pathway [43].

SNX3 is still another binding partner of ADH1. Members of this complex were observed in the fraction eluting between 28 and 30 min. This interaction has been reported previously using the yeast two hybrid approach [44]. SNX3 has the function of sorting nexin, which is required to maintain Golgi enzymes in the proper location [45].

The isoenzymes pyruvate decarboxylase 1 (PDC1), enolase 2 (ENO2) and phosphoglycerate mutase 1 (PMG1) are all associated with glycolysis in yeast in addition to having been reported to interact with ADH1 [46–48]. One of the more surprising outcomes of this work was that all of these enzymes, including ADH1 elute together in at least 5–6 fractions. For example, PDC1 was found in the 14–16 min, 18–24 min, 24–26 min, 26–28 min and 28–30 min fractions along with ADH1. Eno2 showed similar behavior in eluting from the SEC column at 14–16 min, 16–18 min, 24–26 min, 26–28 min, and 28–30 min. PMG1 mimicked PDC1 and ADH1 in eluting at 14–16 min, 16–18 min, 18–24 min, 24–26 min, 26–28 min fraction 28–30 min. Whether PDC1, ENO2, and PMG1 are all members of different complexes formed around ADH1, or hub proteins in their own right is unclear and needs further study.

Glyceraldehyde-3-phosphate dehydrogenase 2 and glucose-6-phosphate isomerase [49,50] co-eluted in the 14–16 min, 18–24 min, and 24–26 min SEC fractions. Although these enzymes have different functions, they are both involved in glycogenesis and energy production. Although they may be interaction partners, this has not been previously reported.

4. Discussion

Even though analysis of native protein mixtures by mass spectrometry is evolving rapidly, the bulk of proteomics is still achieved at the peptide level with some variant of the shotgun proteomics method. An unfortunate limitation of this approach is that proteolysis of a proteome as the first step of analysis destroys a great deal of critical structural information. It is much more difficult, or even impossible in many cases to determine qualitative and quantitative differences between multiple forms of a protein along with the number of isoforms of the protein subsequent to proteolysis.

It is also clear that among the requisite multiple dimensions of chromatography prior to mass spectrometry, reversed phase chromatography (RPC) is the separation method of choice immediately before mass spectrometry in peptide base identification of proteins. RPC is by far the highest resolution technique for peptide fractionation and the mobile phases used in this separation method are ideal for coupling to ESI-MS and MALDI-MS.

Putting these two sets of facts together suggests that preliminary fractionation of the native proteome followed by proteolysis and subsequent RPC-MS analysis would provide more structural information than can be obtained by conventional shotgun proteomics. The methods described here do that in the case of protein complexes. The fact that protein monomers and their complexes vary substantially in size and in a few min SEC columns are capable of fractionating a proteome into 8–10 fractions, each of which vary approximately twofold in hydrodynamic volume allows a course, but rapid global analysis of a cellular interactome. Moreover, SEC is easily achieved under non-denaturing conditions. SEC provides roughly the same number of fractions for RPC-MS analysis as the strong cation exchange (SCX) chromatography used with peptides in shotgun proteomics while gaining important information on protein complexes. Finally, the number of proteins identified in the SEC-RPC approach is comparable to the SCX-RPC method.

Beyond the fact that SEC is a very gentle fractionation method, it is also highly reproducible and relatively predictable. As a "*rule of thumb*" when the apparent molecular mass of a protein eluting from an SEC column exceeds that predicted by the genome by twofold or more, it is highly likely to be a member of a protein complex. The exception would be a glycoprotein with a high level of glycosylation, such as the mucins. Glycan content of these glycoproteins can equal or exceed that of the polypeptide backbone. The fact that a protein is a member of a complex can often be further established by the literature. It is interesting however that the apparent Mr of a complex by SEC does not always match that predicted from Y2H and TAP data. Whether this is due to differences in organisms and growth conditions, variations in sample handling, or deviations in the methods themselves is unknown.

A second rule of thumb is that when a protein appears in multiple fractions above its molecular mass it is likely to be a hub protein. Compared with previous study done by Olinares et al., in characterizing the MDa-sized macromolecular assemblies at cellular organelles [25], our work focused on the separation of hub proteins in complexes. The ADH1, ENO2, FRA11, APL6, CCA1, LST8, MED7, PWP1, PNA15, SPF1, PGK1, PDC1, SSB1 and PMG1 fell into this class. The exception to this rule would be the case where the protein is dissociating from a high molecular mass complex as it is migrating through the column and is spread across the column effluent. But this possibility is easily eliminated. Because dissociation rates are constant, varying migration time by altering the mobile phase velocity during SEC would cause changes in the elution profile. The elution position of complex members associated with high affinity in contrast will be independent of mobile phase velocity. Elution ratios were found to be constant in the case of the proteins cited above.

It was noted in these studies that proteins sometimes appear in the size exclusion chromatogram at less than their expected Mr. These are generally degradation products of sufficient stability they accumulate in cells and can be isolated. Partially degraded proteins are difficult to recognize with shotgun proteomics methods where identification is based on as few as two peptides.

Still higher resolution of complexes can be obtained by further fractionation of proteins collected from the SEC column. It has been noted that many yeast proteins remain in protein complexes during ion exchange and hydrophobic interaction chromatography [23]. Co-migration through multiple dimensions of chromatography provides even stronger evidence of residence in the same protein complex.

5. Conclusions

It is conclude that preliminary fractionation of a proteome by SEC under non-denaturing conditions prior to proteolysis and RPC-MS/MS identification of peptide cleavage products is a useful tool in identifying proteins that are members of a protein complex. Moreover, information thus obtained supports data obtained by the more laborious, but more definitive Y2H and TAP methods. Preliminary SEC fractionation can even suggest the need for Y2H and TAP studies. Information relating to the cellular interactome derived from the SEC method cannot be obtained from shotgun proteomics methods alone. An advantage of the SEC method described here is that it requires little additional work beyond that currently required in the shotgun proteomics workflow.

It is further concluded that the success of this method depends on judicious selection of cell lysis methods and mobile phases that maintain the native structure of protein complexes prior to and during chromatographic fractionation at the protein level. Protein complexes are most stable at near physiological pH and buffer concentrations in the range of 200 mM.

This method may be particularly useful in studying the distribution of proteins among protein complexes as a function of regulatory stimuli, genetic variation, environmental factors, and disease progression, especially when coupled with newly emerging methods for protein quantification. The method will be especially useful in determining the ratio between proteins in the monomer and complex states of association in biological systems.

Acknowledgement

The authors gratefully acknowledge support of this work through a grant (U24-CA-126480) from the U.S. National Cancer Institute.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2010.09.071.

References

[1] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S.G. Oliver, Science 274 (1996) 546.
[2] P. Kahn, Science 270 (1995) 369.
[3] S. Fields, Science 291 (2001) 1221.
[4] R.C. Lee, R.L. Feinbaum, V. Ambros, Cell 75 (1993) 843.
[5] B.J. Reinhart, F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, A.E. Rougvie, H.R. Horvitz, G. Ruvkun, Nature 403 (2000) 901.
[6] B. Wightman, I. Ha, G. Ruvkun, Cell 75 (1993) 855.
[7] B.R. Graveley, Trends Genet. 17 (2001) 100.
[8] E. Zamir, B. Geiger, J. Cell Sci. 114 (2001) 3583.
[9] M. Gerstein, N. Lan, R. Jansen, Science 295 (2002) 284.
[10] C.S. Goh, D. Milburn, M. Gerstein, Curr. Opin. Struct. Biol. 14 (2004) 104.
[11] B. Alberts, Cell 92 (1998) 291.
[12] N. Bertin, N. Simonis, D. Dupuy, M.E. Cusick, J.D.J. Han, H.B. Fraser, F.P. Roth, M. Vidal, Plos Biol. 5 (2007) 1206.
[13] F. Golebiowski, I. Matic, M.H. Tatham, C. Cole, Y.L. Yin, A. Nakamura, J. Cox, G.J. Barton, M. Mann, R.T. Hay, Sci. Signal. 2 (2009).
[14] B. Blagoev, I. Kratchmarova, S.E. Ong, M. Nielsen, L.J. Foster, M. Mann, Nat. Biotechnol. 21 (2003) 315.
[15] T. Bouwmeester, A. Bauch, H. Ruffner, P.O. Angrand, G. Bergamini, K. Croughton, C. Cruciat, D. Eberhard, J. Gagneur, S. Ghidelli, C. Hopf, B. Huhse, R. Mangano, A.M. Michon, M. Schirle, J. Schlegl, M. Schwab, M.A. Stein, A. Bauer, G. Casari, G. Drewes, A.C. Gavin, D.B. Jackson, G. Joberty, G. Neubauer, J. Rick, B. Kuster, G. Superti-Furga, Nat. Cell Biol. 6 (2004) 97.
[16] A.H.Y. Tong, G. Lesage, G.D. Bader, H.M. Ding, H. Xu, X.F. Xin, J. Young, G.F. Berriz, R.L. Brost, M. Chang, Y.Q. Chen, X. Cheng, G. Chua, H. Friesen, D.S. Goldberg, J. Haynes, C. Humphries, G. He, S. Hussein, L.Z. Ke, N. Krogan, Z.J. Li, J.N. Levinson, H. Lu, P. Menard, C. Munyana, A.B. Parsons, O. Ryan, R. Tonikian, T. Roberts, A.M. Sdicu, J. Shapiro, B. Sheikh, B. Suter, S.L. Wong, L.V. Zhang, H.W. Zhu, C.G. Burd, S. Munro, C. Sander, J. Rine, J. Greenblatt, M. Peter, A. Bretscher, G. Bell, F.P. Roth, G.W. Brown, B. Andrews, H. Bussey, C. Boone, Science 303 (2004) 808.
[17] H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R.A. Dean, M. Gerstein, M. Snyder, Science 293 (2001) 2101.
[18] D. Auerbach, S. Thaminy, M.O. Hottiger, I. Stagljar, Proteomics 2 (2002) 611.

[19] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Seraphin, Nat. Biotechnol. 17 (1999) 1030.
[20] A.C. Gavin, P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L.J. Jensen, S. Bastuck, B. Dumpelfeld, A. Edelmann, M.A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A.M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J.M. Rick, B. Kuster, P. Bork, R.B. Russell, G. Superti-Furga, Nature 440 (2006) 631.
[21] J. Goll, P. Uetz, Genome Biol. 7 (2006).
[22] M. Hamdan, P.G. Righetti, Proteomics Today: Protein Assessment and Biomarkers Using Mass Spectrometry, 2-D Electrophoresis and Microarray Technology, Wiley-VCH, 2005.
[23] H. Schagger, G. Vonjagow, Anal. Biochem. 199 (1991) 223.
[24] X.P. Liu, W.C. Yang, Q. Gao, F. Regnier, J. Chromatogr. A 1178 (2008) 24.
[25] P.D.B. Olinares, L. Ponnala, K.J. van Wijk, Mol. Cell. Proteom. 9 (2010) 1594.
[26] H. Jeong, S.P. Mason, A.L. Barabasi, Z.N. Oltvai, Nature 411 (2001) 41.
[27] S. Krishnan, E.Y. Chi, J.N. Webb, B.S. Chang, D.X. Shan, M. Goldenberg, M.C. Manning, T.W. Randolph, J.F. Carpenter, Biochemistry 41 (2002) 6422.
[28] L. Nielsen, R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V.N. Uversky, A.L. Fink, Biochemistry 40 (2001) 6036.
[29] A. Fatouros, T. Osterberg, M. Mikaelsson, Int. J. Pharma. 155 (1997) 121.
[30] J. Withka, P. Moncuse, A. Baziotis, R. Maskiewicz, J. Chromatogr. 398 (1987) 175.
[31] E.T. Young, D. Pilgrim, Mol. Cell. Biol. 5 (1985) 3024.
[32] D.W. Russell, M. Smith, V.M. Williamson, E.T. Young, J. Biol. Chem. 258 (1983) 2674.
[33] S.V. Shlyapnikov, V. Both, V.A. Kulikov, A.A. Dementiev, J. Zelinka, Bioorg. Khim. 13 (1987) 760.
[34] C. Drewke, M. Ciriacy, Biochim. Biophys. Acta 950 (1988) 54.
[35] V. Leskovac, S. Trivic, D. Pericin, FEMS Yeast Res. 2 (2002) 481.
[36] M. Grey, M. Schmidt, M. Brendel, Curr. Genet. 29 (1996) 437.
[37] E. Magonet, P. Hayen, D. Delforge, E. Delaive, J. Remacle, Biochem. J. 287 (1992) 361.
[38] J.P. Klinman, K. Welsh, Biochem. Biophys. Res. Commun. 70 (1976) 874.
[39] M.P. Washburn, D. Wolters, J.R. Yates, Nat. Biotechnol. 19 (2001) 242.
[40] A.C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M.A. Heurtier, R.R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, G. Superti-Furga, Nature 415 (2002) 141.
[41] S.R. Collins, P. Kemmeren, X.C. Zhao, J.F. Greenblatt, F. Spencer, F.C.P. Holstege, J.S. Weissman, N.J. Krogan, Mol. Cell. Proteom. 6 (2007) 439.
[42] J. Ptacek, G. Devgan, G. Michaud, H. Zhu, X.W. Zhu, J. Fasolo, H. Guo, G. Jona, A. Breitkreutz, R. Sopko, R.R. McCartney, M.C. Schmidt, N. Rachidi, S.J. Lee, A.S. Mah, L.H. Meng, M.J.R. Stark, D.F. Stern, C. De Virgilio, M. Tyers, B. Andrews, M. Gerstein, B. Schweitzer, P.F. Predki, M. Snyder, FASEB J. 20 (2006) A1308.
[43] H. Zhu, J.F. Klemic, S. Chang, P. Bertone, A. Casamayor, K.G. Klemic, D. Smith, M. Gerstein, M.A. Reed, M. Snyder, Nat. Genet. 26 (2000) 283.
[44] C.S. Vollert, P. Uetz, Mol. Cell. Proteom. 3 (2004) 1053.
[45] C.R. Haft, M.D. Sierra, V.A. Barr, D.H. Haft, S.I. Taylor, Mol. Cell. Biol. 18 (1998) 7278.
[46] T.G. Watson, J.S. Hough, J. Inst. Brew. 75 (1969) 359.
[47] V.M. Williamson, J. Bennetzen, E.T. Young, K. Nasmyth, B.D. Hall, Nature 283 (1980) 214.
[48] L.A. Fothergillgilmore, H.C. Watson, Biochem. Soc. Trans. 18 (1990) 190.
[49] S. Milewski, Biochim. Biophys. Acta—Protein Struct. Mol. Enzymol. 1597 (2002) 173.
[50] M.A. Sirover, J. Cell. Biochem. 66 (1997) 133.