research articles Journal of proteome research

I-DIRT, A General Method for Distinguishing between Specific and Nonspecific Protein Interactions

Alan J. Tackett, Jeffrey A. DeGrasse, Matthew D. Sekedat, Marlene Oeffinger, Michael P. Rout, and Brian T. Chait*

The Rockefeller University, New York, New York 10021

Received July 19, 2005

Isolation of protein complexes via affinity-tagged proteins provides a powerful tool for studying biological systems, but the technique is often compromised by co-enrichment of nonspecifically interacting proteins. We describe a new technique (I-DIRT) that distinguishes contaminants from bona fide interactors in immunopurifications, overcoming this most challenging problem in defining protein complexes. I-DIRT will be of broad value for studying protein complexes in biological systems that can be metabolically labeled.

Keywords: immunoisolation • immunopurification • background • contamination • protein complex • affinity tag • pol epsilon

Introduction

Advances in rapid isolation techniques and mass spectrometric identification of protein complexes have produced a plethora of protein interaction data.[1−5] Researchers are now faced with the challenge of determining which of these interactions are specific or nonspecific to the complex under study. Protein complexes are now routinely immunoisolated from cell lysates via an affinity-tagged member.[4,6−10] These technologies have become widely utilized in *S. cerevisiae* due to the relative ease of incorporating a genomic tag by homologous recombination, and also to the commercially available TAP-tag collection of dual affinity-tagged proteins.[11] The immunoisolation technique is an exceptionally powerful method for rapidly and efficiently extracting a protein complex from cell lysate under conditions that preserve in vivo protein interactions. Nevertheless, a common problem is the co-enrichment of proteins that associate nonspecifically with the affinity-tagged complex. To attenuate nonspecific interactions, researchers often vary the stringency of the isolation conditions; however, increased stringency often results in the loss of specific protein−protein interactions (and does not guarantee complete removal of contaminants). Protein specificity has classically been monitored by enrichment during purification[12] and more recently with isotopic chemical-labeling,[13] but these procedures do not allow one to differentiate contaminants from true interactors prior to enrichment—and in practice, enrichment procedures are fraught with artifacts.

We have therefore developed a technique termed *I*sotopic *D*ifferentiation of *I*nteractions as *R*andom or *T*argeted (I-DIRT), which differentiates specific from nonspecific protein interactions prior to enrichment, allowing us to distinguish between them at the analytical step. We illustrate the utility of this approach for the characterization of a yeast DNA polymerase ϵ complex isolated under nonstringent conditions.

Experimental Section

Yeast Strains and Growth Conditions. *S. cerevisiae* strains are from the W303 background. One strain contained a Protein A (PrA) affinity tag at the C-terminal coding sequence of the *POL2* open reading frame.[7] The *POL2-PRA* strain was grown in synthetic complete medium to mid-log phase at 30°C, harvested by centrifugation and frozen as pellets in liquid nitrogen (isotopically light cells). A wild-type strain was grown in synthetic medium lacking lysine, which was supplemented with 0.1 mg/mL DL-lysine-4,4,5,5-*d4* 2HCl (C/D/N Isotopes Inc) to mid-log phase at 30°C, harvested by centrifugation, and frozen as pellets in liquid nitrogen (isotopically heavy cells). Isotopically light and heavy cells were mixed 1:1 (by weight), disrupted with a Retsch MM301 mixer mill that was maintained at liquid nitrogen temperature, and stored at −80 °C. A sample of isotopically light cells alone was also disrupted and stored at −80 °C.

Immunoisolation and Mass Spectrometric Identification of Protein Complexes. Immunoisolation of Pol2-PrA and associated proteins was performed as described except that the stringency of the purification buffer was lowered by decreasing the NaCl concentration to 100 mM rather than 300 mM.[7] These less stringent conditions were used to deliberately increase the level of nonspecific associations with the Pol2-PrA-containing complexes. We used 1 g of the mixture of lysed isotopically light (*POL2-PRA*) and heavy (wild-type) cells for the immunoisolation. As a control, we performed an immunoisolation on 0.5 g of lysate from isotopically light *POL2-PRA* cells alone.

Protein complexes containing Pol2-PrA were resolved by SDS-PAGE on 4−20% polyacrylamide gels (Invitrogen) and visualized by Coomassie blue-staining.[7] The entire gel lane (5.2 cm long) was sliced into 17 sections and proteins in each gel

* To whom correspondence should be addressed. E-mail: chait@rockefeller.edu.

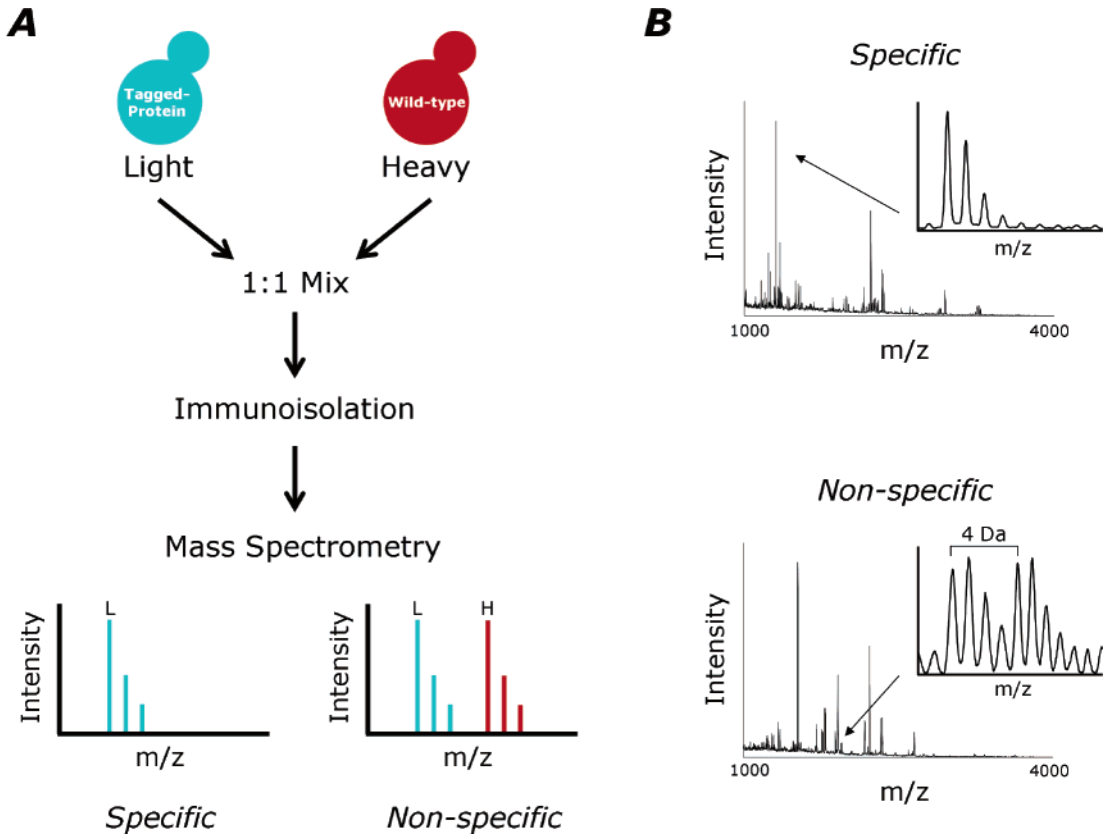

Figure 1. I-DIRT procedure for determining specific members of a protein complex. (A) Cells containing an affinity-tagged protein are grown in light isotopic medium, while wild-type cells are grown in heavy isotopic medium (*d4*-lysine). Equal quantities of these two cell preparations are mixed and the affinity-tagged protein complex is isolated. After isolation of the affinity-tagged protein complex, specific protein interactions are identified by mass spectrometry as isotopically light, while nonspecific interactions appear as a mixture of isotopically light and heavy. (B) Examples of single-stage mass spectra of tryptic peptides from either specific (top) or nonspecific (bottom) proteins. The 4 Da difference between the light and heavy peptides corresponds to *d4*-lysine labeling in the heavy cells.

slice were identified by MALDI-mass spectrometry.[14,15] Briefly, a mass spectrum of tryptic peptides derived from each gel slice was obtained with an in-house modified MALDI-Q-qTOF mass spectrometer (Sciex Centaur, Concord, ON).[14] All m/z values with a signal-to-noise of >2 were subjected to MALDI−MS/MS analysis with an in-house modified MALDI-ion trap mass spectrometer (Finnigan LCQ, ThermoElectron Corp., San Jose, CA).[15] Single-stage and multi-stage mass spectrometric data was used for protein identification with the programs XProteo (www.xproteo.com) and X! Tandem (www.thegpm.org),[16] and confirmed manually for all peptides used in the present study.

Measurement of Isotopic Ratios. A list of tryptic peptides containing at least one lysine residue was obtained for each of the proteins identified from the mass spectrometric data (see above). For each of these lysine-containing peptides, we used the program M-over-Z to visualize a range of the single-stage MALDI-mass spectrum that was large enough to display the isotopically light lysine-containing peptide as well as the heavy version of the peptide. The heavy version of the peptide will be shifted from the light version by 4 Da increments per *d4*-lysine. Using the M-over-Z program, we obtained monoisotopic peak areas for the isotopically light (A_L) and heavy (A_H) peptides. Peak areas were corrected by background subtraction. Since the natural isotopic distribution of the light peptide may overlap with the monoisotopic peak of the heavy peptide, we further corrected A_H with the theoretical isotopic overlap determined from the program Isotopident (http://haven.isb-sib.ch/tools/isotopident/htdocs/). We then used the corrected A_L and A_H values to determine the percent of the peptide that was isotopically light. These calculations are shown in eq 1:

$$\% \text{ light} = \frac{(A_L - A_B)}{(A_L - A_B) + ([A_H - A_B] - [I_O \times \{A_L - A_B\}])} x100 \quad (1)$$

The additional variables in eq 1 are as follows: A_B = background peak area; I_O = theoretical fraction of isotopic overlap of the natural isotope peak from the light peptide with the monoisotopic peak from the heavy peptide. In cases where more than one lysine-containing peptide was identified for a given protein, we report the average of the % light measurements together with the standard deviation.

Results and Discussion

The I-DIRT Technique. The I-DIRT (*I*sotopic *D*ifferentiation of *I*nteractions as *R*andom or *T*argeted) technique is outlined in Figure 1A. Cells containing an affinity-tagged protein are grown in light isotopic medium, while wild-type cells are grown in heavy isotopic medium (here *d4*-lysine). Cells from the isotopically light and heavy cultures are mixed at a 1:1 ratio by cell weight. The heavy:light cell mixture is thoroughly blended and lysed under cryogenic conditions. The isotopically light affinity-tagged protein is immunoisolated on affinity resin. Co-isolating with the affinity-tagged protein will be both specifically and nonspecifically interacting proteins. The specifically interacting proteins, which form their association with the tagged

protein prior to cell lysis, are exclusively light labeled. In contrast, the nonspecific proteins, which form their association with the tagged protein after thawing of the lysed mixture of light and heavy cells, are both isotopically heavy *and* light labeled. After isolation of the affinity-tagged protein complex and trypsin digestion of the co-enriching proteins, specifically interacting proteins are identified by mass spectrometry as isotopically light, while nonspecifically interacting proteins appear as a mixture of isotopically light and heavy (Figure 1B).

Preparation of Samples for I-DIRT. The I-DIRT procedure requires an isotopically light and heavy preparation of cells (Figure 1A). When incorporating a heavy isotope into cellular proteins, one should choose an isotopically heavy metabolite to produce a mass shift that will be readily resolved from a light peptide by mass spectrometric analysis. For example, choosing a heavy metabolite that yields a 1 Da mass difference between a heavy and light peptide will produce mass spectra that will be difficult to interpret due to the natural isotopic distribution of the light peptide. Thus, we used *d4*-lysine, which will produce a 4 Da difference between light and heavy single-lysine containing peptides (or an 8 Da difference for peptides containing two lysines). This 4 Da shift is readily distinguished by our MALDI-QqTOF mass spectrometer (Figure 1B). To produce an isotopically heavy yeast sample, we provided *d4*-lysine as the only source of lysine to wild-type *S. cerevisiae*. Wild-type yeast grown in the presence of *d4*-lysine only incorporated isotopically heavy lysine into polypeptides (AJT and BTC, unpublished observation). We cultured a sufficient quantity of the heavy-labeled reference yeast (2 L) to allow for multiple I-DIRT experiments with a variety of tagged strains. Cells isolated by centrifugation were frozen as small pellets in liquid nitrogen, and stored at −80 °C until needed for the I-DIRT procedure.

A culture of yeast containing a C-terminal PrA tag on the *POL2* gene (*POL2-PRA*) was grown in synthetic complete medium. These isotopically light cells, incorporating natural lysine into its polypeptides, were collected by centrifugation, frozen as small pellets in liquid nitrogen and stored at −80 °C.

To prepare a cellular lysate for immunoisolation with I-DIRT analysis, we mixed wild-type (heavy) and affinity-tagged (light) frozen cell pellets at a 1:1 ratio (by weight). This mixture of light and heavy cells was cryogenically lysed in a mixer mill maintained at liquid nitrogen temperature (Retsch MM301). Care was taken to ensure that the cells did not thaw at any stage of the grinding process. The resulting frozen mixture of lysed heavy and light cells can be stored at −80 °C, and utilized in aliquots for multiple immunoisolations with I-DIRT analysis. A control lysate was also prepared from isotopically light *POL2-PRA* cells alone.

Immunoisolation of Pol2-PrA with I-DIRT Analysis. To test the I-DIRT method, we immunoisolated the DNA polymerase epsilon (pol ϵ) catalytic subunit, Pol2, which was genomically tagged with PrA.[7] We and others have extensively studied this polymerase and found that it is a stable complex of four proteins: Pol2, Dpb2, Dpb3, and Dpb4.[7,17] We have found that immunoisolation of Pol2-PrA under stringent conditions (i.e., 300 mM NaCl) yields the three other members of pol ϵ, but also co-enriched lower levels of potentially nonspecific proteins (including ribosomal and heat shock proteins).[7] Here, we deliberately decreased the stringency (100 mM NaCl) of the Pol2-PrA immunoisolation to increase the level of potential nonspecific proteins and assayed with I-DIRT for specific protein interactions (Figure 2).

Two types of Pol2-PrA immunoisolations were performed: (1) *POL2-PRA* (light) mixed with wild-type (heavy) cells and (2) *POL2-PRA* (light) cells alone as a control. The proteins that co-purified with Pol2-PrA were resolved by SDS-PAGE and visualized by Coomassie blue-staining (Figure 2A). The entire gel lane was sliced from top to bottom and proteins in each gel slice were digested with trypsin. A mass spectrum of tryptic peptides derived from each gel slice was obtained by MALDI-QqTOF mass spectrometry (Figure 2B). All ion peaks with a signal-to-noise of >2 were subjected to MALDI-ion trap MS^2. The resulting MS and MS^2 data was used for protein identification.

A list of tryptic peptides containing at least one lysine residue was obtained for each of the proteins identified by mass spectrometry. For each of these lysine-containing peptides, we used the program M-over-Z to visualize a range of the single-stage mass spectrum that was large enough to display the *h4*-lysine-containing peptide as well as the corresponding *d4*-lysine-containing peptide, should it be present (Figure 2B). For each gel slice, we compared the mass spectra obtained in the heavy:light and that from the light only immunoisolations. This comparison allowed us to determine whether there were any light peptide peaks that could potentially interfere with the observation of a given putative heavy peptide component.

In the mass spectrum from the heavy:light immunoisolation, the heavy peptide is shifted by 4 Da increments per *d4*-lysine. We obtained monoisotopic peak areas for the isotopically light and heavy peptides, and determined the fraction of the peptide that was isotopically light (eq 1). If more than one lysine-containing peptide was identified from a given protein, then the average of the multiple measurements of the light fraction was reported together with the standard deviation of these measurements (Figure 2C).

The results of the I-DIRT analysis of the Pol2-PrA immunoisolation are shown in Figure 2C. If the percentage light for a protein approaches 100%, then that protein has maintained its interaction with Pol2-PrA from the isotopically light cells and is therefore a specific interaction. If the percentage light for a protein is ∼50%, then that protein has nonspecifically associated with the Pol2-PrA complex during the immunoisolation procedure (because there is a 50% chance that either an isotopically light or heavy version of the protein in the mixed tagged and reference cell lysate will nonspecifically bind). As predicted, we observed specific interactions for the four known components of pol ϵ (Pol2, Dpb2, Dpb3, Dpb4), while the remaining proteins were found to associate nonspecifically. Thus, we were able to identify all members of the pol ϵ protein complex and distinguish them from nonspecific proteins under nonstringent conditions.

We envision that there will also be intermediate cases, in which a specifically interacting protein exchanges with a soluble pool during immunoisolation. In such cases, the %light (eq 1) will lie somewhere between 100% and 50% depending on the equilibrium dissociation time constant; at 50% light a very rapidly exchanging protein would be indistinguishable from a contaminant. However, such occasional complications are minimized by using short immunoisolation incubation times (I. M. Cristea, R. Williams, B. T. Chait, M. P. Rout, submitted).

We also utilized the I-DIRT technique on the well-defined Nup84 subcomplex of the nuclear pore complex,[3,18] the Dpb4-chromatin remodeling complex,[7] and the GINS DNA replication complex.[19] As with the pol ϵ complex, we were able to distinguish nonspecific proteins from the published complex

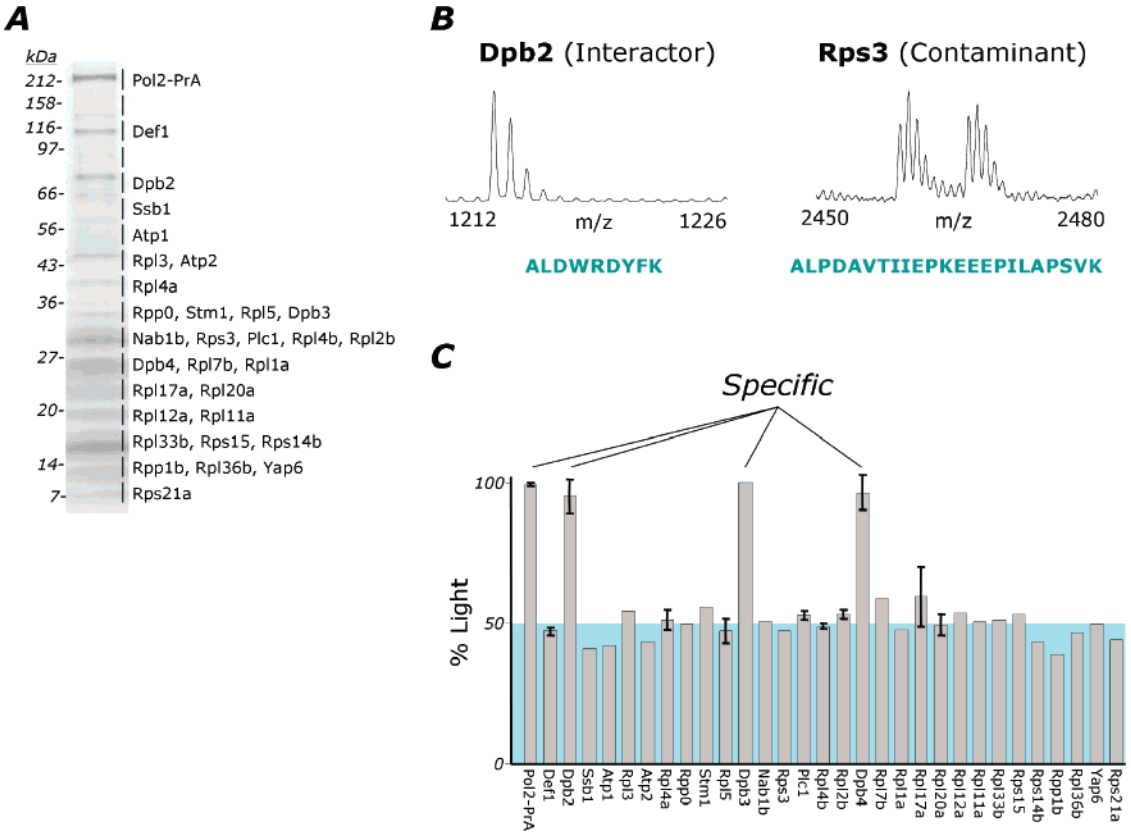

Figure 2. Immunoisolation of Pol2-PrA with I-DIRT analysis. (A) Coomassie blue-stained gel of proteins co-enriching with Pol2-PrA from a mixture of isotopically light *POL2-PRA* and isotopically heavy wild-type cells. The gel lane was sliced as indicated, and the proteins were identified by mass spectrometry. (B) Representative mass spectra of tryptic peptides from a specific interactor with Pol2-PrA (i.e., Dpb2) and a nonspecific contaminant (i.e., Rps3). C) Results of I-DIRT analysis for all the proteins observed in the gel. The contaminants are revealed by virtue of their containing a 1:1 mixture of light and heavy isotopes; real interactors contain only light isotopes.

members (data not shown), thereby validating the technique on distinct protein complexes.

Conclusions

To elucidate the function of a given protein in a cell, it is useful to determine its interacting partners. To probe such interactions, researchers have increasingly turned to the isolation of protein complexes via affinity-tagged proteins. Despite the great power of the immunoisolation approach, the technique is compromised by the presence of nonspecifically interacting proteins that co-enrich during purification. Here we describe a straightforward technique termed I-DIRT (*I*sotopic *D*ifferentiation of *I*nteractions as *R*andom or *T*argeted) to discriminate between specific and nonspecific proteins in an immunoisolated protein complex (Figure 1A). The I-DIRT procedure promises to become an essential component of many immunoisolation protocols because of the following: (1) it provides definitive identification of specifically interacting proteins, (2) it does not compromise the immunoisolation or mass spectrometric identification steps, (3) it uses a single reference material applicable to multiple different immunoisolations, and (4) it is straightforward to implement. Key to the success of the present strategy is mixing and breakage of the heavy and light cells prior to thawing, in contrast to strategies that incorporate mixing and binding post-lysis.[20−22]

Another key advantage of I-DIRT is the ability to use nonstringent immunopurification conditions to isolate a protein complex of interest (Figure 2). Immunoisolations under nonstringent conditions (e.g., lower than physiological salt concentrations and low levels of detergent) can result in co-enrichment of nonspecific proteins (e.g., heat-shock or metabolic proteins). To avoid these nonspecific protein associations, researchers typically increase the stringency of the isolation conditions until the nonspecific proteins are no longer present; however, one cannot be certain whether the increased stringency has resulted in the loss of specific protein−protein interactions. I-DIRT allows for immunoisolations under nonstringent conditions and therefore permits a researcher to probe for more weakly bound or transient members of a protein complex.

In the present study, we describe the addition of deuterium-labeled lysine to medium used to culture *S.* cerevisiae; however, multiple types of isotopically heavy additives (e.g., ^{13}C, ^{15}N) can be utilized, depending on the organism under study and whether identical chromatographic behavior is required for labeled versus unlabeled peptides. Moreover, the method should not be dependent on the manner of protein isolation—the complex can be purified either through a tag or an antibody raised against an untagged member of the complex. Thus, the I-DIRT procedure should be applicable to any biological system that allows for the incorporation of a stable isotope into its food supply, and the isolation of the protein complex of interest.

Note Added after ASAP Publication. This manuscript was originally published on the Web (08/25/2005) missing a sen-

tence at the end of the first paragraph of the Conclusions section (and its related references). The version published 09/20/2005 and in print is correct.

Acknowledgment. This work was supported by grants from the NIH to B.T.C. (RR00862), M.P.R. (GM062427), B.T.C. and M.P.R. (CA89810), and A.J.T. (GM066496), and from the ACS to M.P.R. (RGS-0404251).

References

(1) Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.; Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edelmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. *Nature* **2002**, *415*, 141−147.

(2) Ho, Y.; Gruhler, A.; Heilbut, A.; Bader, G. D.; Moore, L.; Adams, S. L.; Millar, A.; Taylor, P.; Bennett, K.; Boutilier, K.; Yang, L.; Wolting, C.; Donaldson, I.; Schandorff, S.; Shewnarane, J.; Vo, M.; Taggart, J.; Goudreault, M.; Muskat, B.; Alfarano, C.; Dewar, D.; Lin, Z.; Michalickova, K.; Willems, A. R.; Sassi, H.; Nielsen, P. A.; Rasmussen, K. J.; Andersen, J. R.; Johansen, L. E.; Hansen, L. H.; Jespersen, H.; Podtelejnikov, A.; Nielsen, E.; Crawford, J.; Poulsen, V.; Sorensen, B. D.; Matthiesen, J.; Hendrickson, R. C.; Gleeson, F.; Pawson, T.; Moran, M. F.; Durocher, D.; Mann, M.; Hogue, C. W.; Figeys, D.; Tyers, M. *Nature* **2002**, *415*, 180−183.

(3) Rout, M. P.; Aitchison, J. D.; Suprapto, A.; Hjertaas, K.; Zhao, Y.; Chait, B. T. *J. Cell. Biol.* **2000**, *148*, 635−651.

(4) Archambault, V.; Chang, E. J.; Drapkin, B. J.; Cross, F. R.; Chait, B. T.; Rout, M. P. *Mol. Cell.* **2004**, *14*, 699−711.

(5) Sanders, S. L.; Jennings, J.; Canutescu, A.; Link, A. J.; Weil, P. A. *Mol. Cell. Biol.* **2002**, *22*, 4723−4738.

(6) Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. *Nat. Biotechnol.* **1999**, *17*, 1030−1032.

(7) Tackett, A. J.; Dilworth, D. J.; Davey, M. J.; O'Donnell, M.; Aitchison, J. D.; Rout, M. P.; Chait, B. T. *J. Cell. Biol.* **2005**, *169*, 35−47.

(8) Aitchison, J. D.; Rout, M. P.; Marelli, M.; Blobel, G.; Wozniak, R. W. *J. Cell. Biol.* **1995**, *131*, 1133−1148.

(9) Aitchison, J. D.; Blobel, G.; Rout, M. P. *Science* **1996**, *274*, 624−627.

(10) Dou, Y.; Milne, T. A.; Tackett, A. J.; Smith, E. R.; Fukuda, A.; Wysocka, J.; Allis, C. D.; Chait, B. T.; Hess, J. L.; Roeder, R. G. *Cell* **2005**, *121*, 873−885.

(11) Ghaemmaghami, S.; Huh, W. K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. *Nature* **2003**, *425*, 737−741.

(12) de Duve, C. *Exploring Cells with A Centrifuge. Nobel Lectures in Physiology 1971−1980*; Lindsten, J., Ed.; World Scientific Publishing Co., London, 152−172., 1992.

(13) Marelli, M.; Smith, J. J.; Jung, S.; Yi, E.; Nesvizhskii, A. I.; Christmas, R. H.; Saleem, R. A.; Tam, Y. Y.; Fagarasanu, A.; Goodlett, D. R.; Aebersold, R.; Rachubinski, R. A.; Aitchison, J. D. *J. Cell. Biol.* **2004**, *167*, 1099−1112.

(14) Krutchinsky, A. N.; Zhang, W.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 493−504.

(15) Krutchinsky, A. N.; Kalkum, M.; Chait, B. T. *Anal. Chem.* **2001**, *73*, 5066−5077.

(16) Craig, R.; Cortens, J. P.; Beavis, R. C. *J. Proteome Res.* **2004**, *3*, 1234−1242.

(17) Dua, R.; Edwards, S.; Levy, D. L.; Campbell, J. L. *J. Biol. Chem.* **2000**, *275*, 28816−28825.

(18) Siniossoglou, S.; Lutzmann, M.; Santos-Rosa, H.; Leonard, K.; Mueller, S.; Aebi, U.; Hurt, E. *J. Cell. Biol.* **2000**, *149*, 41−54.

(19) Takayama, Y.; Kamimura, Y.; Okawa, M.; Muramatsu, S.; Sugino, A.; Araki, H. *Genes Dev.* **2003**, *17*, 1153−1165.

(20) Blagoev, B.; Kratchmarova, I.; Ong, S.-E.; Nielsen, M.; Foster, L. J.; Mann, M. *Nat. Biotechnol.* **2003**, *21*, 315−318.

(21) Schulze, W. X.; Mann, M. *J. Biol. Chem.* **2004**, *279*, 10756−10764.

(22) Schulze, W. X.; Deng, L.; Mann, M. *Mol. Systems Biol.* **2005**, msb4100012.

PR050225E