Laboratory for Computational Proteomics

www.FenyoLab.org

E-mail: Info@FenyoLab.org

Facebook: *NYUMC Computational Proteomics Laboratory*

Twitter: *@CompProteomics*

Rapid, optimized interactomic screening

Zhanna Hakhverdyan[1], Michal Domanski[1,2], Loren E Hough[1,10], Asha A Oroskar[3], Anil R Oroskar[3], Sarah Keegan[4,5], David J Dilworth[6,7], Kelly R Molloy[8], Vadim Sherman[9], John D Aitchison[6,7], David Fenyö[4,5], Brian T Chait[8], Torben Heick Jensen[2], Michael P Rout[1] & John LaCava[1]

We must reliably map the interactomes of cellular macromolecular complexes in order to fully explore and understand biological systems. However, there are no methods to accurately predict how to capture a given macromolecular complex with its physiological binding partners. Here, we present a screening method that comprehensively explores the parameters affecting the stability of interactions in affinity-captured complexes, enabling the discovery of physiological binding partners in unparalleled detail. We have implemented this screen on several macromolecular complexes from a variety of organisms, revealing novel profiles for even well-studied proteins. Our approach is robust, economical and automatable, providing inroads to the rigorous, systematic dissection of cellular interactomes.

High-throughput DNA sequencing facilitates whole-genome characterization within weeks[1,2]. Likewise, advances in mass spectrometry (MS)[3,4] are enabling cellular proteomes to be defined. However, we have yet to exhaustively map any interactome—the cell's comprehensive biomolecular interaction network[5,6]; we may have identified less than 20% of the protein interactions in humans, not counting dynamic, tissue- or disease-specific interactions[7–9].

A main approach for interactomic exploration is affinity capture[10]. For this, cells are lysed and their contents extracted into a solution that ideally preserves each target macromolecular complex. Complexes are then specifically enriched from the cell extract using affinity reagents—usually antibodies—that recognize the target, either directly or through an epitope tag, thus permitting subsequent characterization of the complex. However, one of the foremost challenges with affinity capture studies is the precise optimization of the extraction conditions because no single condition is optimal for the preservation of the many different types of interactions found in macromolecular complexes[10]. As a result, affinity capture experiments either require time-consuming optimization on a case-by-case basis, or a compromise must be made by using unoptimized conditions; the latter is a common strategy but often results in sparse coverage of protein-protein interactions and error-prone data[11–13]. A variety of advanced bioinformatics tools[14,15] and databases of common contaminant proteomes[16,17] have attempted to mitigate this problem, but they cannot fully substitute for optimized sample preparation[11]. Because any given extraction solution influences the complement of copurifying proteins, multiple extractant formulations are required if one intends to broadly sample the interactome, as underscored by a recent high-throughput study of membrane-protein interactions in yeast[18].

The problem of maintaining post-extraction protein-complex stability is comparable to that which once hindered protein crystallographic efforts. For crystallography, the answer came with the development of massively parallel crystallization optimization screens[19,20] that allow hundreds of conditions to be simultaneously explored. Inspired by this, we have developed improved methods for the rapid processing of cellular material in conjunction with parallelized, multiparameter searches of extraction conditions. Our approach is compatible with both standard lab-scale investigations and high-throughput robotics, and it facilitates the systematic exploration of the interactome of any given protein in a cell.

RESULTS

Designing a large-scale interactomics screen

Our strategy (**Fig. 1**) starts with the distribution of cryo-milled cell material[21,22] to a multiwell plate. To enable the uniform delivery of frozen cell powder to each well in the plate, we designed dispensing manifolds (**Fig. 1** and **Supplementary Fig. 1a**,**b**). After being dispensed, the powder in the wells is thawed by addition of an array of distinct extractants. The resulting extracts are clarified of insoluble material using a clog-resistant filtration device (**Fig. 1** and **Supplementary Fig. 1c**,**d**) that provides a filtrate matching the quality of centrifugally clarified cell extract. The remainder of the procedure implements commercially available supplies and equipment (Online Methods and **Supplementary Protocol 1**). Our screen allowed us to thoroughly explore reagents commonly used in affinity capture experiments: salts, buffers and detergents (**Fig. 2a** and **Supplementary Table 1**).

[1]Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, New York, USA. [2]Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark. [3]Orochem Technologies Inc., Naperville, Illinois, USA. [4]Center for Health Informatics and Bioinformatics, New York University School of Medicine, New York, New York, USA. [5]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, New York, USA. [6]Institute for Systems Biology, Seattle, Washington, USA. [7]Seattle Biomedical Research Institute, Seattle, Washington, USA. [8]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York, USA. [9]High Energy Physics Instrument Shop, The Rockefeller University, New York, New York, USA. [10]Present address: Department of Physics and BioFrontiers Institute, University of Colorado, Boulder, Colorado, USA. Correspondence should be addressed to M.P.R. (rout@rockefeller.edu) or J.L. (jlacava@rockefeller.edu).

RECEIVED 23 JUNE 2014; ACCEPTED 3 MARCH 2015; PUBLISHED ONLINE 4 MAY 2015; DOI:10.1038/NMETH.3395

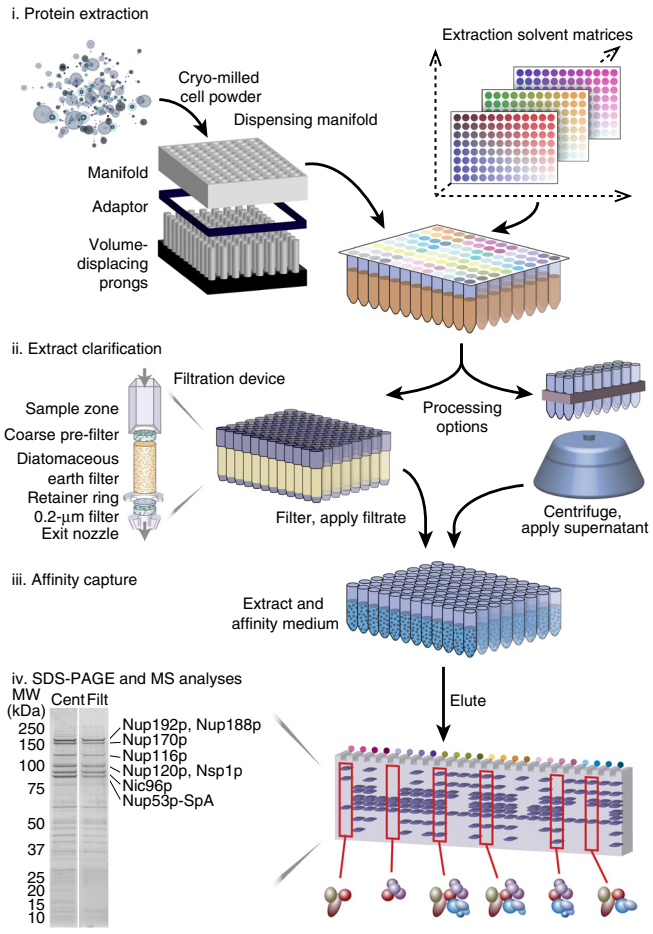

Figure 1 | Schematic representation of the parallelized affinity capture procedure. (i) Cells expressing a tagged protein of interest are cryo-milled, and precise aliquots of the resulting frozen powder are deposited into a multiwell plate using an in-house–made dispensing manifold of adjustable well volume. A diverse set of extraction solvents is rapidly added in parallel, and the powder is resuspended completely by mechanical agitation and brief sonication. (ii) Rapid removal of insoluble material is achieved either by centrifugation or using a novel deep-bed filtration device. (iii) Magnetic bead–based affinity capture is performed on the clarified extract: affinity medium is collected, washed and eluted with the aid of a 96-well magnetic separator. (iv) Elution of the complexes is followed by SDS-PAGE, Coomassie blue staining and (as desired) MS analysis. The resulting copurification profiles are cross-compared to infer interactomes and determine preparative conditions appropriate for further experiments. The gel on the left shows a comparison of SDS-PAGE analysis of yeast Nup53p-SpA immunopurification with both extract clarification methods: centrifugation at 20,000*g* for 10 min (Cent) and filtration at 3,000*g* for 5 min (Filt); MW, molecular weight standard. Duplicate experiments produced identical results (not shown). Proteins are labeled in accordance with **Figure 3a**. See Online Methods for extraction solvent composition.

Evaluation of affinity capture profiles

Because large amounts of data were generated during screening, we developed a web portal to assist in affinity capture data management. Our software (described in **Supplementary Note 1** and publicly accessible at http://www.copurification.org/) accepts images of gels along with experimental metadata. Gel images are automatically sectored lane-by-lane and annotated with the conditions applied to each, respectively. The lanes are also clustered according to protein banding pattern similarity, to ease the discovery of lane-to-lane differences and trends (for example, **Fig. 2b**). This database provides a platform for the work of different experimentalists to be compared side by side, with instantaneous access to the respective experimental conditions for ease of reproduction.

There are two commonly used approaches for analyzing affinity-captured samples: (i) SDS-PAGE with dye-based visualization and subsequent MS of select conditions and (ii) direct MS of the samples. We compared these two approaches for the 96-well purification of Nup1p, a component of the yeast nuclear pore complex (NPC), exploring a diverse set of extraction conditions. Two replicates were carried out for the comparison: one set resolved by SDS-PAGE and stained with Coomassie blue (**Supplementary Fig. 2**), the other set processed for liquid chromatography–tandem MS (LC-MS/MS) (**Supplementary Table 2** and Online Methods). Dendrograms of the data sets were produced as follows: (i) for SDS-PAGE, we used our software to cluster the lanes on the basis of the intensity rank of bands exhibiting similar apparent molecular masses (**Fig. 2b**, **Supplementary Table 3** and **Supplementary Fig. 3**); (ii) for MS, we clustered the filtered data according to the presence of common proteins (**Fig. 2b** and **Supplementary Fig. 3**). The resulting dendrograms exhibited a statistically significant similarity ($P < 1 \times 10^{-7}$, permutation test; **Supplementary Fig. 4**). Three proteins were revealed in most conditions—Kap95p, Kap60p and Nup1p-SpA—and a larger number of proteins were observed in, for example, low acetate or high citrate/acetate (**Fig. 2b**). Our conclusions are twofold. First, SDS-PAGE provides a representative readout of the composition of affinity-captured samples, faithfully revealing the effects of changing affinity capture conditions in a rapid, robust and inexpensive fashion. Second, with sufficient resources, direct sample-to-MS analytical approaches may be used.

Exploring the molecular organization of a 50-MDa complex

The yeast NPC is ~50 MDa and consists of multiple copies of ~30 different proteins. It presents an excellent test bed for our screen because it comprises a diverse physicochemical landscape and has a modular architecture consisting of subcomplexes of different sizes. Moreover, an extensive catalog of already existing affinity capture results[23] enabled us to assess our findings and the quality of results produced by the screen. On the basis of the above-described initial results, we further modified our conditions matrix to test other reagents and applied three optimal conditions to SpA-tagged NPC proteins Nup1p, Nup53p and Pom152p (**Fig. 3a**).

Interestingly, all three proteins responded similarly to the extraction conditions (Online Methods), exhibiting improvements in yield, background and hierarchical coverage compared to the best results previously obtained[23]: condition (i) gave small complexes with few interactors for all three and no common components; condition (ii) gave more complicated profiles that partially overlapped in composition with each other; and condition (iii) gave the most complicated and highly similar profiles, representing almost the entirety of the NPC (**Supplementary Data**). Although distinctive, the profiles were in agreement with the established arrangement of proteins in the NPC, constituting interaction shells of increasing size and degree of overlap[23] (**Fig. 3b**).

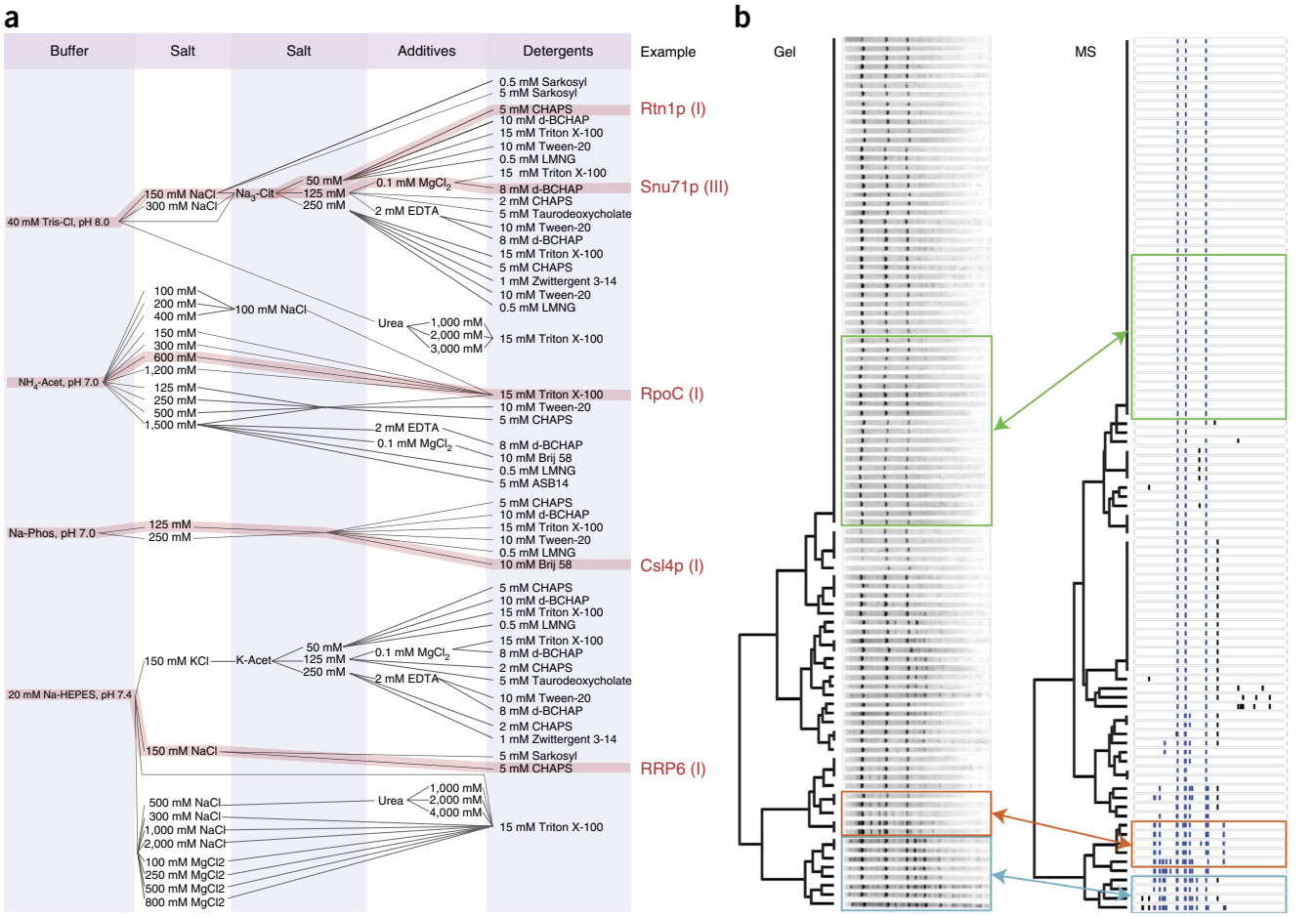

Figure 2 | Extraction condition design and copurification pattern analysis. (**a**) Diagram showing mixtures of components in select extraction solvent formulations. The main components are a pH buffer, one or two salts, an additive and a detergent. Some examples of useful formulations discovered through screening are indicated (right; refer to **Fig. 4** to view the associated copurification patterns, indicated by Roman numerals). (**b**) Comparison of SDS-PAGE (left) and LC-MS/MS (right) clustering analysis of Nup1p-SpA 96-well purification. For comparison, MS data are represented as a pseudo-gel, where each band corresponds to a protein above a certain intensity threshold (see Online Methods). Known Nup1p-interacting proteins are indicated with blue bands; the rest are labeled black. Coclustering conditions with identical or highly similar components producing distinct copurification profiles are highlighted in blue (low ammonium acetate or low potassium acetate), orange (high sodium citrate or high ammonium acetate with Triton X-100) and green (sodium citrate or potassium acetate with CHAPS) boxes. See **Supplementary Table 1** for extraction solvent compositions, **Supplementary Table 2** for MS data and **Supplementary Figure 3** for pseudo-gel and SDS-PAGE lane labels.

Taken together, these results demonstrated that this screen could be used for the systematic interrogation of optimized purification conditions and comprehensive interactomic mapping of different macromolecular complexes.

Robotic automation for higher throughput

Translation to automation has several advantages, including increased throughput and reproducibility. Using a liquid-handling robot, we developed a version of the screen that includes automated production of extractant matrices and sample handling from the addition of affinity medium to clarified extracts through to the final wash. Given the intriguing results observed using trisodium citrate to purify nuclear pore subcomplexes (for example, **Fig. 3a**), we implemented automation to systematically explore the effect of this reagent on Nup53p-SpA affinity capture profiles over 48 increments from 50 to 300 mM. We observed three distinct profiles: copurification with Nup170p and Kap121p (iv); dimer with Nup170p (v); and a larger subcomplex of the NPC—Nup192p, Nup188p, Nup170p, Nup116p, Nup120p, Nsp1p and Nic96p (vi) (**Fig. 3c** and **Supplementary Note 2**). These results demonstrated that the automated procedure was precise and revealed systematic changes in the copurification pattern specific to trisodium citrate, which involved the loss of Kap121p and the increased retention of a large number of NPC components as the concentration increased.

Simultaneous mapping of distinct interaction networks

We tested the screen on many different types of macromolecular complexes. We examined four proteins with two different tags acquired from commercial collections. These proteins (Arp2p-GFP, Csl4p-TAP, Snu71p-TAP and Rtn1p-GFP) exhibit distinct subcellular localization patterns and functions. Each protein was subjected to a 32-condition screen (**Supplementary Figs. 5–8**), which allowed us to assay multiple proteins within the same 96-well plate. High-quality copurification profiles were obtained (**Fig. 4a** and **Supplementary Data**), including the observation of novel and distinctive copurification patterns for proteins already extensively subjected to affinity capture MS strategies.

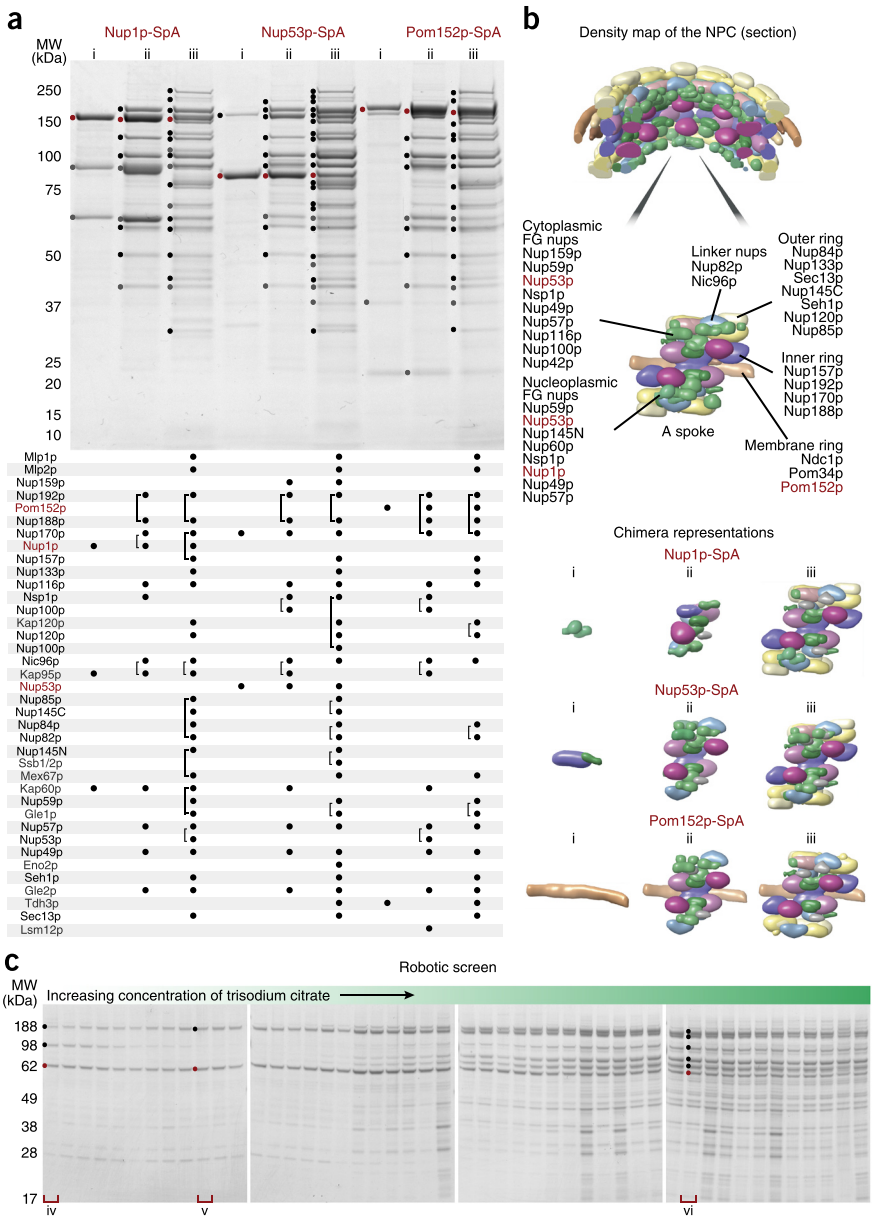

Figure 3 | Nuclear pore complex (NPC) purification from single proteins to macromolecular assemblies. (**a**) Representative SDS-PAGE image and MS analysis of affinity capture of three nucleoporins: Nup1p, Nup53p and Pom152p. The protein bands identified by MS are marked on the gel with dots. The table below contains the list of identified proteins. Affinity-tagged Nup proteins are labeled red, the remaining NPC constituents are black, and non-NPC proteins are gray. Each protein identified by MS is marked by a dot under the corresponding lane. The brackets indicate comigrating proteins identified in a single band. See Online Methods for extraction solvent composition and **Supplementary Data** for MS analysis. (**b**) Section through the density map of the NPC[23] with one spoke enlarged, and minimal Chimera representations of NPC subcomplexes in **a** for one spoke. (**c**) SDS-PAGE analysis of one Nup53p-SpA purification screen performed using a Hamilton STAR liquid handling workstation, testing 50–300 mM trisodium citrate. The purifications from lanes iv, v and vi were manually repeated and analyzed by MALDI-TOF-MS (**Supplementary Data**).

The Arp2/3 complex is a conserved actin nucleator that participates in multiple actin-dependent processes, including endocytosis[24]. Screening revealed a putative novel macromolecular assembly composed of clathrin (Chc1p) with its adaptor protein (Ent2p) and actin (Act1p) with its recruiting and activating proteins[25] (Pan1p, End3p and Arp2/3 complex; **Fig. 4a**, Arp2p-GFP, i). Of these, however, only Act1p and Pan1p are known to physically interact with Arp2p[26,27]. As End3p is linked to Arp2p genetically[28] and no direct links between Arp2p and Ent2p have been demonstrated, we chose to perform a secondary affinity capture for GFP-tagged versions of these proteins. The results (**Fig. 4a**, Ent2p-GFP, Ent3p-GFP and Arp2p-GFP, and **Supplementary Note 2**) supported the role of Pan1p as a core scaffold that couples actin and clathrin with the rest of the endocytic machinery in an interaction network involved in the early stages of actin-dependent clathrin-mediated endocytosis[27,29].

Csl4p is a component of the eukaryotic exosome, a modular multiprotein RNase with compartment-specific components[30,31]. Screening revealed conditions that selectively destabilized the compartment-specific components Rrp6p, Lrp1p, and Ski7p (**Fig. 4a**, Csl4p-TAP, iii and iv, compared with the canonical exosome, profile ii), while retaining the component Dis3p. Owing to the relative stabilities of these different components in established purification conditions[32,33], yeast strains with genetic deletions have been necessary to obtain comparable complexes in previous experiments (for example, ref. 33). In a separate profile, we observed components of the cytoplasmic exosome cofactor Ski complex Ski2p/Ski3p (**Fig. 4a**, Csl4p-TAP, i), which are considered recalcitrant to copurification[34].

Snu71p is a component of the U1 small nuclear ribonucleoprotein (snRNP) complex, a constituent of the spliceosome[35,36]. Of interest, we purified U1 snRNP (**Fig. 4a**, Snu71p-TAP, ii) with nuclear mRNA–associated proteins Sto1p and Pab1p and the major coat protein of ScVLA virus, which is known to covalently bind the mRNA cap[37] (**Fig. 4a**, Snu71p-TAP, i). We also noted a profile demonstrating a direct interaction between Snu71p and Prp40p (**Fig. 4a**, Snu71p-TAP, iii). Despite more than a decade of research on the composition of this complex, this dimer was unambiguously shown only recently by the introduction of a deletion mutation to a third constituent of this 17-component ribonucleoprotein[38]. Here, it was obtained within a single screen.

Rtn1p, an integral endoplasmic reticulum (ER) membrane component, embodies many of the challenges to affinity capture approaches: it is spread between multiple localizations and functions, is expected to form particularly dynamic or transient

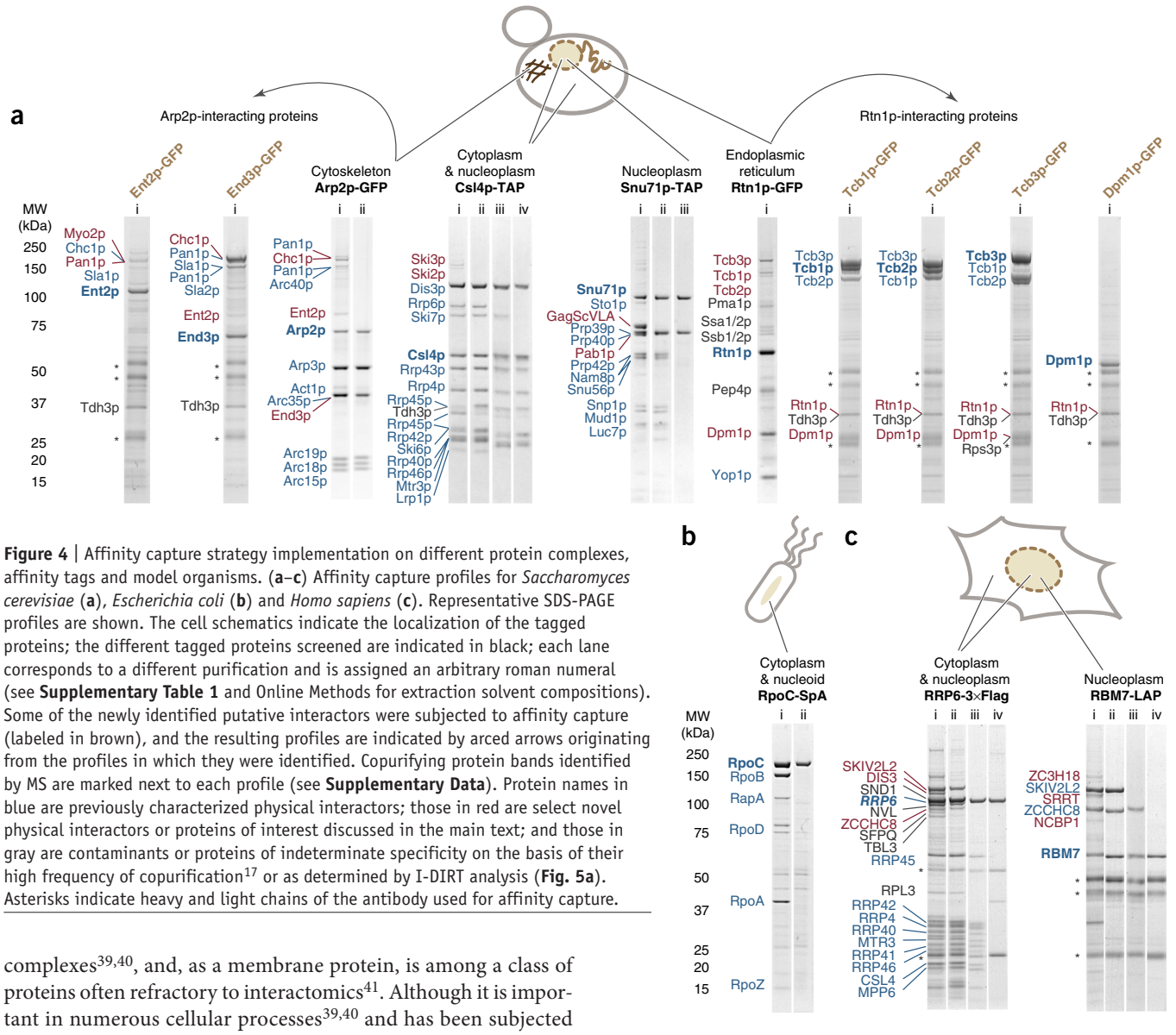

Figure 4 | Affinity capture strategy implementation on different protein complexes, affinity tags and model organisms. (**a**–**c**) Affinity capture profiles for *Saccharomyces cerevisiae* (**a**), *Escherichia coli* (**b**) and *Homo sapiens* (**c**). Representative SDS-PAGE profiles are shown. The cell schematics indicate the localization of the tagged proteins; the different tagged proteins screened are indicated in black; each lane corresponds to a different purification and is assigned an arbitrary roman numeral (see **Supplementary Table 1** and Online Methods for extraction solvent compositions). Some of the newly identified putative interactors were subjected to affinity capture (labeled in brown), and the resulting profiles are indicated by arced arrows originating from the profiles in which they were identified. Copurifying protein bands identified by MS are marked next to each profile (see **Supplementary Data**). Protein names in blue are previously characterized physical interactors; those in red are select novel physical interactors or proteins of interest discussed in the main text; and those in gray are contaminants or proteins of indeterminate specificity on the basis of their high frequency of copurification[17] or as determined by I-DIRT analysis (**Fig. 5a**). Asterisks indicate heavy and light chains of the antibody used for affinity capture.

complexes[39,40], and, as a membrane protein, is among a class of proteins often refractory to interactomics[41]. Although it is important in numerous cellular processes[39,40] and has been subjected to an affinity capture screen intended for membrane proteins[18], there are comparatively few validated physical interaction data available for Rtn1p. Lacking prior knowledge of the expected interactions, we selected a condition from our screen using our experience-based SDS-PAGE profile criteria (see Discussion). The copurifying proteins (**Fig. 4**, Rtn1p-GFP) included known and uncharacterized putative Rtn1p-interacting partners. Among these was the ER membrane–associated protein Dpm1p, known to have a negative genetic interaction with Rtn1p[42]. We used secondary affinity capture to validate this interaction with Dpm1p as well as with the ER membrane tricalbins (Tcb1p–Tcb3p), which also copurified but have no previously demonstrated physical links to Rtn1p; all four GFP-tagged proteins copurified Rtn1p, and the Tcb proteins each copurified one another[43] and also yielded Dpm1p.

We repeated the Rtn1p-GFP affinity capture experiment, implementing isotopic differentiation of interactions as random or targeted (I-DIRT) analysis to distinguish interactions formed *in vivo* from those likely to be *in vitro* artifacts[44]. I-DIRT analysis (**Supplementary Table 4**) indeed confirmed that most of the strong bands in our optimized affinity capture represent protein interactions with Rtn1p that were present *in vivo* (**Fig. 5a**), with a few prominent bands corresponding to common contaminants (see also **Supplementary Note 2**). Among putative Rtn1p interactors were five out of the six proteins known to tether ER to the plasma membrane, namely Tcb1p–Tcb3p, Ist2p and Scs2p. Notably, Rtn1p, Tcb1p, and Tcb3p were identified in Sac1p (lipid phosphatase) and Scs2p (ER–plasma membrane tether) purifications[45]. Given its proposed function of stabilizing curved membranes[39,46], we suggest that Rtn1p may help stabilize membrane curvature at ER–plasma membrane contact sites where lipid transfer and modification occurs.

To address whether a standard affinity capture approach could have reproduced these results, we executed a side-by-side comparison of the popular tandem affinity purification procedure, recently tuned for membrane proteins[18], with an optimized procedure emerging from our screen (see **Supplementary Table 4** for MS analysis). The results illustrate that the classic approach cannot compete with our screening strategy in terms of either quality or efficiency (**Fig. 5b**).

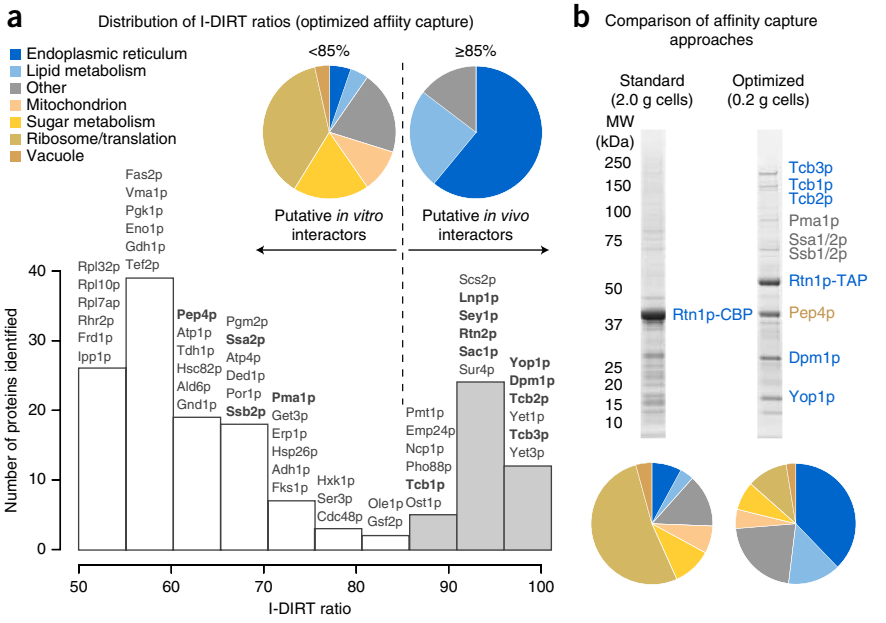

Figure 5 | In-depth analysis of Rtn1p affinity capture. (**a**) Bottom, frequency distribution of I-DIRT ratios (light protein intensity/total protein intensity, normalized to 100%) from the Rtn1p-GFP affinity capture experiment (extraction condition as in **Fig. 4a**, Rtn1p-GFP). Representative proteins are listed above each bar; proteins labeled in **Figure 4a** (Rtn1p-GPF) and known interactors of Rtn1p are boldface. Top, subcellular localizations and molecular functions for the putative *in vivo* and *in vitro* interactors (see Online Methods). (**b**) SDS-PAGE and MS comparison of standard and optimized affinity capture. 4 g of Rtn1p-TAP powder was processed essentially as previously described[18] using Triton X-100 as a detergent in a two-step affinity capture experiment, and 0.4 g of Rtn1p-TAP was processed in a one-step affinity capture experiment using conditions revealed in the present study (**Fig. 4a**, Rtn1p-GFP). Half of the elution was analyzed by SDS-PAGE followed by Coomassie staining, and the other half was analyzed by LC-MS/MS. The distribution of subcellular localizations and molecular functions was analyzed as in **a** and is plotted below the corresponding gel lanes.

Adaptability of the screen for diverse model organisms

Different model organisms often exhibit idiosyncrasies associated with affinity capture experiments. Our screen allows alternatives to be explored at each step. For example, an issue found with *Escherichia coli* was the high viscosity of cell extracts due to high concentrations of released genomic DNA. We therefore modified our procedure to include a short low-energy sonication using a multitip probe, sufficient to resuspend the frozen cell powder during extraction and reduce viscosity to levels compatible with affinity capture. Similarly, low quantities of starting material may present an issue when working with tissue culture cells. Thus, we also modified the screen for a 24-well format in order to economize on cell usage. In our procedure, the mass of yeast cell pellet required per purification was reduced from the typical order of grams[18,47–49] to the range of tens to hundreds of milligrams; similarly, we consumed only 50 mg of cryo-milled human cells per profile, an approximately eightfold reduction from the amounts used in contemporary high-throughput studies[50,51]. These modifications therefore enabled economical interactomic screens in diverse model organisms.

From *E. coli* we purified the RNA polymerase[52] complex corresponding to the σ70 containing holoenzyme in complex with RapA[53] and RpoC-SpA isolated away from RNA polymerase (**Supplementary Data** and **Fig. 4b**, RpoC-SpA, i and ii), demonstrating that the implemented modifications provide affinity capture results comparable in quality to those from yeast (**Supplementary Fig. 9**). Using human cells, we revisited the RNA exosome, conducting purifications via a 3×Flag-tagged hRRP6 (EXOSC10) and thereby adding another common tag variety to those tested thus far (**Supplementary Fig. 10**). Among our observations, we noted the stable retention of SKIV2L2 (hMTR4) in numerous interaction profiles (**Supplementary Data**; see also **Supplementary Note 2**). SKIV2L2 is also a member of the nuclear-specific human exosome cofactor NEXT complex[54], along with ZCCHC8 and RBM7. We also observed ZCCHC8, in human exosome profiles (for example, **Fig. 4c** RRP6-3×Flag, i and ii), raising the question as to whether a single copy or multiple copies of SKIV2L2 are present in the combined exosome/NEXT-containing fractions. To extend our exploration, we applied this screen to the NEXT complex itself, purified via LAP-tagged RBM7 (**Supplementary Fig. 11**). We observed NEXT in association with NCBP1 (CBP80), SRRT (ARS2) and ZC3H18 (NHN1) (**Supplementary Data** and **Fig. 4c**, RBM7-LAP, i; see also **Supplementary Note 2**). A parallel study demonstrated the physiological relevance of these interactors via a suite of functional assays[55]. We also made the novel observation of a direct interaction between RBM7 and ZCCHC8 (**Fig. 4c**, RBM7-LAP, ii vs. iii), demonstrating that these interactors form a stable dimer.

DISCUSSION

The solvent environment of the extractant plays a crucial role in dictating the stability of both real and artifactual protein-protein interactions during affinity capture; and finding extractants that maximally explore the real interactome while minimizing artifacts can be difficult. Our approach addresses this limitation, providing a fast, efficient and cost-effective means to scan many conditions for their ability to preserve physiological interactions and minimize noise. This is particularly important for researchers that hope to go beyond protein identifications and obtain high-quality protein preparations for biochemical and structural studies.

Doubtless because of the huge diversity of interaction types, we do not observe one set of conditions that works well for all the protein interactomes we have studied, thus underscoring the need for our screen. However, encouragingly, our results suggest that the optimal set of extraction conditions determined for a subset of constituents in a given complex will suffice for the interactomic exploration of all the components in that complex (for example, **Fig. 3a**). Moreover, we observed that during the secondary affinity capture of putative interactors identified in a primary screen (i.e., biochemical validation), copurification profiles containing both overlapping and distinct proteins were frequently obtained (see, for example, **Figs. 3a** and **4a**, Arp2p-GFP and Rtn1p-GFP; and **Fig. 4c**, RRP6-3×Flag and RBM7-LAP)—highlighting

the potential of this screen to uncover local (sub)complexes as well as the broader interactome. These combined attributes are particularly important given the current efforts to create a "human proteome encyclopedia[56]," which will undoubtedly require rigorous attention by any investigator to the preparation of the highest-quality samples for affinity capture MS analyses.

For general purposes, we favor SDS-PAGE with protein staining for sample quality assessment followed by band excision and MS to determine protein identities. Our findings, time and again, reinforced the notion that high-quality affinity capture experiments are typified in SDS-PAGE profiles by a discrete pattern of sharp, abundant and roughly stoichiometric bands as well as a paucity of background staining from other fainter bands (see, for example, refs. 23,36 and **Figs. 3a** and **4**). The existence of increasingly sensitive general protein stains provides gel-based visualization options even for very low–abundance samples[57]. These criteria in turn allow for the judicious application of MS analyses to only the most potentially informative samples.

When tens to hundreds of SDS-PAGE protein copurification profiles are viewed in parallel, patterns of common and changing proteins and their solvent dependencies typically become readily apparent, and several promising conditions reveal themselves. A typical 96-well screen, as a consequence of being thorough, may yield many gel lanes with comparable banding profiles (**Fig. 2b**) and many that do not meet the criteria for further analysis. To modulate between throughput and screen breadth, one can adjust the total number of conditions (presented here in multiples of 96, 32 and 24).

A promising protein copurification profile accompanied by high-quality MS-based protein identifications provides a strong basis for a hypothesis regarding the existence of a physically associated protein complex *in vivo*. Such data should encourage the design of orthogonal experiments intended to rigorously test this hypothesis. Hence, in one sense, the presented screen can be considered a rapid and efficient hypothesis-generation machine for physical interactions.

Our procedure is also compatible with direct sample-to-MS analyses (**Fig. 2b**) and can be implemented using robotic automation (**Fig. 3c**), which greatly enhances throughput. Future data-mining opportunities will include the development of unsupervised, machine-based classification schemes to further improve our ability to identify promising samples, thus augmenting high-throughput interactomic studies.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank The Rockefeller University High Energy Physics Instrument Shop for diligence in custom apparatus design and fabrication; X. Wang for assistance with MS data analysis; and members of the Chait, Jensen and Rout laboratories for help and discussion. I. Poser and A. Hyman (Max Planck Institute of Molecular Biology and Genetics, Dresden) provided the RBM7-LAP cell line. This work was funded by the US National Institutes of Health (NIH) grant nos. U54 GM103511 and P41 GM109824 (J.D.A., B.T.C. and M.P.R.), P50 GM076547 (J.D.A.) and P41 GM103314 (B.T.C.); the Lundbeck Foundation (to T.H.J. and J.L.) and the Danish National Research Foundations (to T.H.J.).

AUTHOR CONTRIBUTIONS

J.L. and M.P.R. conceived the screening strategy; J.L. carried out proof-of-concept experiments, assisted by L.E.H.; L.E.H. and V.S. designed manifolds, which were fabricated by V.S. and tested by L.E.H., J.L. and Z.H.; filters were designed by A.A.O., A.R.O. and J.L., fabricated by A.A.O. and A.R.O., and tested by J.L. and Z.H.; J.L., Z.H. and M.D. designed experiments, executed screens and further developed procedures—with yeast work primarily carried out by Z.H. and human cell line work primarily carried out by M.D.; MS analyses were carried out by J.L., Z.H. and K.R.M., with I-DIRT done by Z.H.; transposing the procedure to robotic automation was carried out by D.J.D. assisted by J.L.; J.L., Z.H. and D.F. conceived of the protein copurification gel database and software, which was built by S.K. and D.F. with testing and feedback from J.L. and Z.H.; J.D.A., D.F., B.T.C., T.H.J., M.P.R. and J.L. supervised the project; Z.H., B.T.C., M.P.R. and J.L. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Brockhurst, M.A., Colegrave, N. & Rozen, D.E. Next-generation sequencing as a tool to study microbial evolution. *Mol. Ecol.* **20**, 972–980 (2011).
2. Ross, J.S. & Cronin, M. Whole cancer genome sequencing by next-generation methods. *Am. J. Clin. Pathol.* **136**, 527–539 (2011).
3. Charbonnier, S., Gallego, O. & Gavin, A.-C. The social network of a cell: recent advances in interactome mapping. *Biotechnol. Annu. Rev.* **14**, 1–28 (2008).
4. Collins, M.O. & Choudhary, J.S. Mapping multiprotein complexes by affinity purification and mass spectrometry. *Curr. Opin. Biotechnol.* **19**, 324–330 (2008).
5. Kiemer, L. & Cesareni, G. Comparative interactomics: comparing apples and pears? *Trends Biotechnol.* **25**, 448–454 (2007).
6. Williamson, M.P. & Sutcliffe, M.J. Protein-protein interactions. *Biochem. Soc. Trans.* **38**, 875–878 (2010).
7. del Sol, A., Balling, R., Hood, L. & Galas, D. Diseases as network perturbations. *Curr. Opin. Biotechnol.* **21**, 566–571 (2010).
8. Stumpf, M.P.H. *et al.* Estimating the size of the human interactome. *Proc. Natl. Acad. Sci. USA* **105**, 6959–6964 (2008).
9. Menche, J. *et al.* Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015).
10. LaCava, J. *et al.* Affinity proteomics to study endogenous protein complexes: pointers, pitfalls, preferences and perspectives. *Biotechniques* **58**, 103–119 (2015).
11. Bell, A.W., Nilsson, T., Kearney, R.E. & Bergeron, J.J.M. The protein microscope: incorporating mass spectrometry into cell biology. *Nat. Methods* **4**, 783–784 (2007).
12. Devos, D. & Russell, R.B. A more complete, complexed and structured interactome. *Curr. Opin. Struct. Biol.* **17**, 370–377 (2007).
13. Breitkreutz, B.-J. *et al.* The BioGRID interaction database: 2008 update. *Nucleic Acids Res.* **36**, D637–D640 (2008).
14. Armean, I.M., Lilley, K.S. & Trotter, M.W.B. Popular computational methods to assess multiprotein complexes derived from label-free affinity purification and mass spectrometry (AP-MS) experiments. *Mol. Cell. Proteomics* **12**, 1–13 (2013).
15. Pardo, M. & Choudhary, J.S. Assignment of protein interactions from affinity purification/mass spectrometry data. *J. Proteome Res.* **11**, 1462–1474 (2012).
16. Trinkle-Mulcahy, L. *et al.* Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* **183**, 223–239 (2008).
17. Mellacheruvu, D. *et al.* The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).
18. Babu, M. *et al.* Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*. *Nature* **489**, 585–589 (2012).
19. Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24**, 409–411 (1991).
20. Chayen, N.E. High-throughput protein crystallization. *Adv. Protein Chem. Struct. Biol.* **77**, 1–22 (2009).
21. Oeffinger, M. *et al.* Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat. Methods* **4**, 951–956 (2007).

22. Domanski, M. *et al.* Improved methodology for the affinity isolation of human protein complexes expressed at near endogenous levels. *Biotechniques* **0**, 1–6 (2012).
23. Alber, F. *et al.* The molecular architecture of the nuclear pore complex. *Nature* **450**, 695–701 (2007).
24. Machesky, L.M. & Gould, K.L. The Arp2/3 complex: a multifunctional actin organizer. *Curr. Opin. Cell Biol.* **11**, 117–121 (1999).
25. Moseley, J.B. & Goode, B.L. The yeast actin cytoskeleton: from cellular function to biochemical mechanism. *Microbiol. Mol. Biol. Rev.* **70**, 605–645 (2006).
26. Liu, S.-L., Needham, K.M., May, J.R. & Nolen, B.J. Mechanism of a concentration-dependent switch between activation and inhibition of Arp2/3 complex by coronin. *J. Biol. Chem.* **286**, 17039–17046 (2011).
27. Duncan, M.C., Cope, M.J., Goode, B.L., Wendland, B. & Drubin, D.G. Yeast Eps15-like endocytic protein, Pan1p, activates the Arp2/3 complex. *Nat. Cell Biol.* **3**, 687–690 (2001).
28. Moreau, V., Galan, J.M., Devilliers, G., Haguenauer-Tsapis, R. & Winsor, B. The yeast actin-related protein Arp2p is required for the internalization step of endocytosis. *Mol. Biol. Cell* **8**, 1361–1375 (1997).
29. Tang, H.Y., Munn, A. & Cai, M. EH domain proteins Pan1p and End3p are components of a complex that plays a dual role in organization of the cortical actin cytoskeleton and endocytosis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **17**, 4294–4304 (1997).
30. Mitchell, P. *et al.* Rrp47p is an exosome-associated protein required for the 3′ processing of stable RNAs. *Mol. Cell. Biol.* **23**, 6982–6992 (2003).
31. Wasmuth, E.V. & Lima, C.D. Structure and activities of the eukaryotic RNA exosome. *Enzymes* **31**, 53–75 (2012).
32. Allmang, C. *et al.* The yeast exosome and human PM-Scl are related complexes of 3′→5′ exonucleases. *Genes Dev.* **13**, 2148–2158 (1999).
33. Dziembowski, A., Lorentzen, E., Conti, E. & Séraphin, B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.* **14**, 15–22 (2007).
34. Synowsky, S.A., van den Heuvel, R.H.H., Mohammed, S., Pijnappel, P.W.W.M. & Heck, A.J.R. Probing genuine strong interactions and post-translational modifications in the heterogeneous yeast exosome protein complex. *Mol. Cell. Proteomics* **5**, 1581–1592 (2006).
35. Gottschalk, A. *et al.* A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA* **4**, 374–393 (1998).
36. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032 (1999).
37. Blanc, A., Goyer, C. & Sonenberg, N. The coat protein of the yeast double-stranded-RNA virus L-A attaches covalently to the cap structure of eukaryotic messenger-RNA. *Mol. Cell. Biol.* **12**, 3390–3398 (1992).
38. Görnemann, J. *et al.* Cotranscriptional spliceosome assembly and splicing are independent of the Prp40p WW domain. *RNA* **17**, 2119–2129 (2011).
39. De Craene, J.-O. *et al.* Rtn1p is involved in structuring the cortical endoplasmic reticulum. *Mol. Biol. Cell* **17**, 3009–3020 (2006).
40. Dawson, T.R., Lazarus, M.D., Hetzer, M.W. & Wente, S.R. ER membrane-bending proteins are necessary for *de novo* nuclear pore formation. *J. Cell Biol.* **184**, 659–675 (2009).
41. Helbig, A.O., Heck, A.J.R. & Slijper, M. Exploring the membrane proteome—challenges and analytical strategies. *J. Proteomics* **73**, 868–878 (2010).
42. Schuldiner, M. *et al.* Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**, 507–519 (2005).
43. Creutz, C.E., Snyder, S.L. & Schulz, T.A. Characterization of the yeast tricalbins: membrane-bound multi-C2-domain proteins that form complexes involved in membrane trafficking. *Cell. Mol. Life Sci.* **61**, 1208–1220 (2004).
44. Tackett, A.J. *et al.* I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J. Proteome Res.* **4**, 1752–1756 (2005).
45. Manford, A.G., Stefan, C.J., Yuan, H.L., Macgurn, J.A. & Emr, S.D. ER-to-plasma membrane tethering proteins regulate cell signaling and ER morphology. *Dev. Cell* **23**, 1129–1140 (2012).
46. Voeltz, G.K., Prinz, W.A., Shibata, Y., Rist, J.M. & Rapoport, T.A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586 (2006).
47. Puig, O. *et al.* The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**, 218–229 (2001).
48. Gavin, A.-C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006).
49. Krogan, N.J. *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637–643 (2006).
50. Sowa, M.E., Bennett, E.J., Gygi, S.P. & Harper, J.W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009).
51. Behrends, C., Sowa, M.E., Gygi, S.P. & Harper, J.W. Network organization of the human autophagy system. *Nature* **466**, 68–76 (2010).
52. Ebright, R.H. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J. Mol. Biol.* **304**, 687–698 (2000).
53. Sukhodolets, M.V., Cabrera, J.E., Zhi, H.J. & Jin, D.J. RapA, a bacterial homolog of SWI2/SNF2, stimulates RNA polymerase recycling in transcription. *Genes Dev.* **15**, 3330–3341 (2001).
54. Lubas, M. *et al.* Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637 (2011).
55. Andersen, P.R. *et al.* The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.* **20**, 1367–1376 (2013).
56. Anonymous. The call of the human proteome. *Nat. Methods* **7**, 661 (2010).
57. Gauci, V.J., Wright, E.P. & Coorssen, J.R. Quantitative proteomics: assessing the spectrum of in-gel protein detection methods. *J. Chem. Biol.* **4**, 3–29 (2011).

ONLINE METHODS

Affinity capture. All cell lines and strains used in this study are listed in **Supplementary Table 5**. Yeast, *E. coli* and human cell lines were cultured using standard procedures and were cryo-milled and affinity captured essentially as previously described[21,22], except we adapted the experiments for 96-well plates as described in the text and elaborated stepwise in **Supplementary Protocol 1**. Human cell lines were not subjected to mycoplasma testing during the course of the study. Rabbit IgG used for purifying TAP and SpA tagged proteins was purchased from Innovative Research. Anti-GFP polyclonal antibodies were prepared and conjugated as previously described[22], except the concentration of ammonium sulfate used during the conjugation was 1.5 M. In all cases, proteins were eluted from the affinity medium by the addition of 1× NuPAGE LDS sample loading solution (Life Technologies); elution of GFP-tagged proteins included incubation at 70 °C for 10 min. Extraction solvent working solutions were mixed from concentrated stock solutions in 2.5-ml deep-well plates (VWR) manually, using a Formulator (Formulatrix) or using a Hamilton STAR liquid handling workstation (program files provided in **Supplementary Protocol 2**). **Supplementary Figure 12** contains an engineering diagram of the powder-dispensing manifold. Resuspension of powders in extraction solvents included sonication with an ice water–chilled microplate horn (yeast) or eight-tip microprobe (bacteria and human) (Qsonica). Yeast lysates were also vortexed with steel beads to aid rapid homogenization. **Supplementary Figure 13** shows the bead-dispensing manifold used in yeast affinity capture experiments. Custom-manufactured filters (**Fig. 1**, **ii** and **Supplementary Fig. 1c**,**d** Orochem Technologies) were used to clarify yeast cell extracts for screens; otherwise, extracts were clarified by centrifugation at 14,000 r.p.m. and 4 °C for 10 min in a benchtop microcentrifuge. To ensure reproducibility, we repeated all purifications presented (and processed for MS) individually in microcentrifuge tubes using an otherwise comparable procedure except that extracts were clarified via centrifugation. Polyacrylamide gels were stained either with a homemade colloidal Coomassie brilliant blue G-250 solution[58] or with Imperial protein stain (Thermo Fisher Scientific). Gel images were recorded in TIFF using a Fujifilm LAS-3000 or an Epson Photo v700. In addition to the cited publications, detailed protocols for many preparatory procedures used in this study can be obtained at http://www.ncdir.org/public-resources/protocols/.

Affinity capture solvent compositions for immunopurifications in the main figures. Extractant formulations referenced in the figure legends are provided below. The remaining extractant formulations are presented in **Supplementary Table 1**.

In **Figure 1**, **iv**, Nup53p-SpA, the composition was 40 mM Tris-Cl, pH 8, 250 mM trisodium citrate, 150 mM NaCl, 1% (v/v) Triton X-100. In **Figure 3a**, Nup1p-SpA, Nup53p-SpA and Pom152p-SpA, the composition for (i) was 50 mM trisodium citrate, 300 mM NaCl, 0.1% (v/v) Tween-20, 2 mM EDTA, 40 mM Tris, pH 8; for (ii) was 250 mM trisodium citrate, 10 mM Brij58, 0.3 mM Sarkosyl, 40 mM Tris, pH 8; and for (iii) was 1.5 M ammonium acetate, 1% (v/v) Triton X-100. In **Figure 3c**, 40 mM Tris-Cl, pH 8, 1% (v/v) Triton X-100 are common to all lanes; trisodium citrate was present at 50–300 mM concentration in 5 mM increments. In **Figure 4a**, Ent2p and End3p, the composition was 40 mM Tris-Cl, pH 8, 150 mM NaCl, 250 mM trisodium citrate, 10 mM deoxy-BigCHAP; and for Tcb1p, Tcb2p, Tcb3p and Dpm1p was 40 mM Tris-Cl, pH 8, 150 mM NaCl, 50 mM trisodium citrate, 5 mM CHAPS.

Standard mass spectrometric identification of proteins. The major bands observed in SDS-polyacrylamide gels were excised and analyzed either by MALDI-TOF-MS essentially as previously described[22] or by nanoLC-ESI-MS/MS on an LTQ Orbitrap XL, Orbitrap Velos, Q Exactive Plus or Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). For analysis of excised protein bands using the LTQ Orbitrap XL or Orbitrap Velos, the dry peptides were resuspended in 0.5% (v/v) acetic acid, pressure loaded on a self-packed C18 column and subjected to a 10-min gradient: 8 min 0–43%, 2 min 43–100% solvent B (solvent A = 0.1 M acetic acid in water; solvent B = 0.1 M acetic acid, 70% (v/v) acetonitrile in water, flow rate 200 nl/min). As peptides eluted, the top ten peaks were selected for fragmentation, without dynamic exclusion. For analysis on the Q Exactive Plus and Orbitrap Fusion, peptides were resuspended in 0.1% (v/v) formic acid and separated using a 10-min gradient (8 min 0–30%, 2 min 30–100%) on an EASY-Spray column (Thermo Fisher Scientific) using an EASY-nLC 1000 (Thermo Fisher Scientific; solvent A = 0.1% (v/v) formic acid in water, solvent B = 0.1% (v/v) formic acid in acetonitrile, flow rate 300 nl/min). The three most abundant ions in each full scan were fragmented by HCD on the Q Exactive Plus. On the Orbitrap Fusion, a fixed duty cycle of 3 s was used. The RAW files were converted to MZXML format with the MM File Conversion tool (http://www.massmatrix.net/mm-cgi/downloads.py/) or MGF format by ProteoWizard[59] and searched against the yeast protein database with X! Tandem[60].

For the analysis of whole affinity captured fractions in **Figures 2b** and **5b**, protein samples were run ~4–6 mm into an SDS-polyacrylamide gel (gel plug), and gels were stained with Coomassie blue. Stained gel regions were excised, cut into 1-mm cubes, destained and digested for 6 h with 120 µl of 3.1 ng/µl trypsin (Promega) in 25 mM ammonium bicarbonate. An equal volume of 2.5 mg/ml POROS R2 20-µm beads (Life Technologies) in 5% (v/v) formic acid, 0.2% (v/v) trifluoroacetic acid was added, and the mixture was incubated on a shaker at 4 °C for 24 h. Digests were desalted on C18 tips, eluted, dried by vacuum centrifugation, resuspended in 0.1% formic acid and separated using a 10-min gradient on an EASY-Spray column (as above). The 12 most abundant ions were selected from each full scan and fragmented by HCD; dynamic exclusion was enabled. RAW files were converted and searched as above. In order to determine the molecular functions of constituent proteins (**Fig. 5b**), we searched the gene descriptions for key words (as for I-DIRT analysis; see below).

Gel and MS data clustering and correlation analysis. The details of the gel image analysis are provided in the **Supplementary Note 1**; the source code is publicly available at https://github.com/FenyoLab/copurification/. Once the lanes were sectored and bands identified, quantified and assigned an apparent molecular weight, they were categorized as dark, light or not observed (**Fig. 2b** and **Supplementary Fig. 3**).

Supplementary Table 2 contains the unfiltered search results of 96 LC-MS/MS runs. For each sample, we extracted the intensities of all the hits, filtered out exogenous and endogenous contaminants (**Supplementary Table 3**) and considered the hits that

were at least 10% as intense as the most intense hit (after initial contaminant filtering). We used a modified version of a source code available at the GPM repository ftp://ftp.thegpm.org/ to output the resulting protein sets as a pseudo-gel (**Fig. 2b** and **Supplementary Fig. 3**).

For both data sets, the Ward method was used for hierarchical clustering with distance between data points calculated as Euclidean distance[61]. To perform the correlation, we calculated the cophenetic distance for the gel and MS dendrograms[62]. The cophenetic correlation was then calculated, which is defined as the Pearson correlation between the cophenetic distance matrices of the dendrograms[62]. A *P* value was obtained by a permutation test: the labels of the MS dendrogram were shuffled 10 million times, and a correlation calculated between the MS and gel dendrograms for each random shuffle (see **Supplementary Fig. 4** for frequency distribution). A *P* value of $<1 \times 10^{-7}$ was calculated as a proportion of the random distribution equal to or greater than the actual correlation (0.53).

Graphical representation of NPC subcomplexes. We used the density maps for individual nucleoporins available at http://salilab.org/npc/ and the UCSF Chimera package[63] to graphically represent the NPC subcomplexes.

I-DIRT data analysis. I-DIRT was carried out essentially as previously described[44]: the Rtn1-GFP strain was grown in synthetic complete minimal medium lacking lysine and supplemented with 50 mg/l of isotopically light lysine, and a wild-type DF5α strain was grown in the same medium but supplemented with 50 mg/l of isotopically heavy lysine (l-lysine:2HCl $^{13}C_6$, Cambridge Isotopes). Both were frozen, mixed and cryo-milled. Rtn1-GFP was affinity captured from the mixed powder extracted in 40 mM Tris-Cl pH 8, 50 mM trisodium citrate, 150 mM NaCl and 5 mM CHAPS. The eluted sample was reduced, alkylated and precipitated with methanol and chloroform[64]. The precipitate was resuspended in 50 mM ammonium bicarbonate, 0.1% (w/v) RapiGest (Waters) via bath sonication with heating (20 min at 70 °C, followed by 2 min at 95 °C). The proteins were digested with trypsin (Promega) overnight. RapiGest was depleted following the manufacturer's instructions, and the digest was desalted over C18 Omix tips (Agilent Technologies). The eluted fractions were analyzed on LTQ Orbitrap XL as described before except with 1-h gradient and dynamic exclusion enabled. The output data were processed with MaxQuant[65] (http://maxquant.org/) using essentially the default parameters (for the light sample, all amino acids were set to light; for the heavy sample, 6-Da heavy lysine was selected; the yeast translated ORF sequences (http://www.yeastgenome.org/), reversed sequences and contaminants database were searched) to identify and measure the intensity of heavy and light peptides. The "Evidence.txt" file containing all the peptide identifications and heavy/light measurements was used in the final analysis. Peptides mapped to contaminants, constituents of a reversed sequence database or those containing no lysine were excluded from the analysis. We further excluded nonunique peptides and peptides with a single MS/MS fragmentation event. For the remaining peptides, I-DIRT ratios were calculated by dividing the intensity of the peptide with light lysine by the total intensity (light and heavy). To calculate the I-DIRT ratio of proteins, we averaged the I-DIRT ratios of its constituent peptides. Proteins with a single peptide contributing to the I-DIRT ratio measurement were excluded from the analysis as unreliable. For proteins with four or more peptides contributing to the I-DIRT ratio measurement, those peptides with outlying I-DIRT ratios were filtered out (no more than one peptide removed per protein) using the following criteria: if the calculated I-DIRT ratio for a peptide was <Q1 (first quartile) to 1.5 × IQR (interquartile range) or >Q3 (third quartile) + (1.5 × IQR), that peptide was excluded. All statistical calculations and plotting were done with R (http://www.r-project.org/)[66]. To assess the normality of protein I-DIRT ratio distribution, we constructed a Q-Q plot, which revealed a notable deviation from the $y = x$ line, implying that the data were not normally distributed (**Supplementary Fig. 14**). To assess the shape of the distribution, we binned the data in 5% intervals (**Supplementary Fig. 15**). Despite a low bimodality coefficient[67,68] (0.2956), the distribution deviates significantly from unimodal by the Hartigans' dip test[69] (*P* value = 0) and has a positive Akaike's information criterion difference[70] (0.1061)—suggesting bimodality. We used the Mixtools package in R[71] to fit two normal distributions to the data (**Supplementary Fig. 15**). We accepted a cutoff of mean ± 2 s.d. of the second distribution as stable interactors of Rtn1p (≥85%). We considered anything below 85% to constitute interactions indistinguishable from those formed post-extraction. In order to determine the molecular functions of constituent proteins, we searched the gene descriptions for key words. The following were the categories and key words searched: Endoplasmic reticulum—"er"/"endoplasmic reticulum"; Sugar metabolism—"glycolysis"/"gluconeogenesis"/"glucose"/"glycolytic"/"pentose"; Vacuole—"vacuol"; Ribosome/translation – "ribosom"/" translat"; Lipid metabolism—"lipid"/"fatty acid"/"choline"/"sterol"/"ceramide"; Mitochondrion—"mitochond"; Other—everything that didn't match. All searches were case insensitive. Once a gene description matched a key word, the protein encoded by the gene was put into the corresponding category, allowing us to count the number of proteins in each category and construct a pie chart of the distribution of molecular functions and localization for a given protein set (**Fig. 5**).

58. Candiano, G. *et al.* Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**, 1327–1333 (2004).
59. Chambers, M.C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
60. Craig, R. & Beavis, R.C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20**, 1466–1467 (2004).
61. Ward, J.H. Jr. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244 (1963).
62. Sokal, R.R. & Rohlf, F.J. The comparison of dendrograms by objective methods. *Taxon* **11**, 33–40 (1962).
63. Pettersen, E.F. *et al.* UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
64. Wessel, D. & Flügge, U.I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138**, 141–143 (1984).
65. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
66. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2012).
67. Ellison, A.M. Effect of seed dimorphism on the density-dependent dynamics of experimental populations of *Atriplex triangularis* (Chenopodiaceae). *Am. J. Bot.* **74**, 1280–1288 (1987).
68. Statistical Analysis System Institute. *SAS/STAT User's Guide* (SAS, 1990).
69. Hartigan, J.A. & Hartigan, P.M. The dip test of unimodality. *Ann. Stat.* **13**, 70–84 (1985).
70. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).
71. Benaglia, T., Chauveau, D., Hunter, D.R. & Young, D.S. mixtools: an R package for analyzing finite mixture models. *J. Stat. Softw.* **32**, 1–29 (2009).

doi:10.1038/nmeth.3395