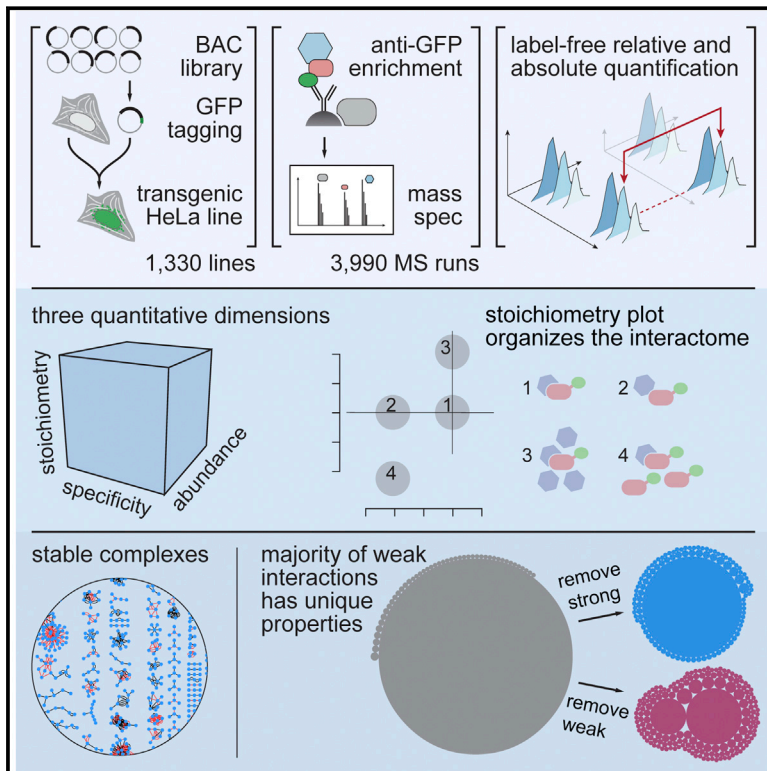# Cell

A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances

Graphical Abstract

Authors

Marco Y. Hein, Nina C. Hubner, Ina Poser, ..., Frank Buchholz, Anthony A. Hyman, Matthias Mann

Correspondence

hyman@mpi-cbg.de (A.A.H.), mmann@biochem.mpg.de (M.M.)

In Brief

Weak interactions shape the cellular protein interaction network as determined from proteomic measures of cellular interaction specificities, the strength of those interactions, and the cellular copy numbers of the proteins involved.

Highlights

- Human interactome dataset connecting 5,400 proteins with 28,500 interactions

- Three quantitative dimensions measure specificities, stoichiometries, and abundances

- Stable complexes are rare but stand out by a signature of balanced stoichiometries

- Weak interactions dominate the network and have critical topological properties

CrossMark

CellPress

A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances

Marco Y. Hein,[1,6,8] Nina C. Hubner,[1,6,9] Ina Poser,[2] Jürgen Cox,[1] Nagarjuna Nagaraj,[1] Yusuke Toyoda,[2,10] Igor A. Gak,[3] Ina Weisswange,[4,5] Jörg Mansfeld,[3] Frank Buchholz,[2,4] Anthony A. Hyman,[2,7,*] and Matthias Mann[1,7,*]

[1]Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[2]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[3]Cell Cycle, Biotechnology Center, TU Dresden, 01307 Dresden, Germany
[4]Medical Systems Biology, UCC, Medical Faculty Carl Gustav Carus, TU Dresden, 01307 Dresden, Germany
[5]Eupheria Biotech GmbH, 01307 Dresden, Germany
[6]Co-first author
[7]Co-senior author
[8]Present address: Howard Hughes Medical Institute, University of California San Francisco, San Francisco, CA 94143, USA
[9]Present address: Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, Nijmegen 6525 GA, the Netherlands
[10]Present address: Institute of Life Science, Kurume University, Kurume, Fukuoka 839-0864, Japan
*Correspondence: hyman@mpi-cbg.de (A.A.H.), mmann@biochem.mpg.de (M.M.)
http://dx.doi.org/10.1016/j.cell.2015.09.053

SUMMARY

The organization of a cell emerges from the interactions in protein networks. The interactome is critically dependent on the strengths of interactions and the cellular abundances of the connected proteins, both of which span orders of magnitude. However, these aspects have not yet been analyzed globally. Here, we have generated a library of HeLa cell lines expressing 1,125 GFP-tagged proteins under near-endogenous control, which we used as input for a next-generation interaction survey. Using quantitative proteomics, we detect specific interactions, estimate interaction stoichiometries, and measure cellular abundances of interacting proteins. These three quantitative dimensions reveal that the protein network is dominated by weak, substoichiometric interactions that play a pivotal role in defining network topology. The minority of stable complexes can be identified by their unique stoichiometry signature. This study provides a rich interaction dataset connecting thousands of proteins and introduces a framework for quantitative network analysis.

INTRODUCTION

Proteins are central protagonists of life at the molecular level. They interact for structural, regulatory, and catalytic purposes, forming macromolecular structures as well as stable or transient multi-protein complexes. Accordingly, protein interactions vary greatly in their biophysical properties, while protein abundances range from a few to millions of copies per cell. The interactome is therefore the product of two factors: binary affinities between protein interfaces (Rual et al., 2005; Stelzl et al., 2005; Rolland et al., 2014) and the cellular proteome, which itself is characterized by subcellular localization, post-translational modifications and protein concentrations (Hein et al., 2013; Mann et al., 2013).

Mapping the protein interactome landscape has been a longstanding goal of modern biology and a variety of methods have been developed to this end (Seebacher and Gavin, 2011). Affinity purification followed by mass spectrometry (AP-MS) can in principle determine the members of protein complexes in their cellular context in an unbiased manner (Gingras et al., 2007) and has enabled large-scale protein interaction studies of several model organisms, including human cells (Ewing et al., 2007; Malovannaya et al., 2011). Nanoscale liquid chromatography (LC) coupled to sensitive and fast mass spectrometers has boosted interaction proteomics technology in recent years, increasing coverage and minimizing false negative rates. It has also enabled a paradigm shift from identification to quantification of interacting proteins (Bantscheff et al., 2012). Quantitative approaches permit the use of mild immunoprecipitation (IP) protocols and allow specific binders to stand out by their quantitative signature even from very large backgrounds of unspecific proteins (Mellacheruvu et al., 2013; Keilhauer et al., 2015). Additionally, MS-based proteomics is now able to characterize entire cellular proteomes with increasingly complete coverage (Beck et al., 2011; Mann et al., 2013), providing abundances and copy-number estimates of the expressed proteins. This should now allow studying the quantitative interactome as a function of the underlying proteome. To generate model systems that closely recapitulate in vivo conditions, we have previously developed bacterial artificial chromosome (BAC) transgeneomics: GFP-tagged proteins are expressed in mammalian cell lines from BAC transgenes with near-endogenous expression patterns from human or orthologous mouse loci (Poser et al., 2008). GFP-based tags are dual-purpose in that they can be used for both imaging and as affinity handle. Combining these
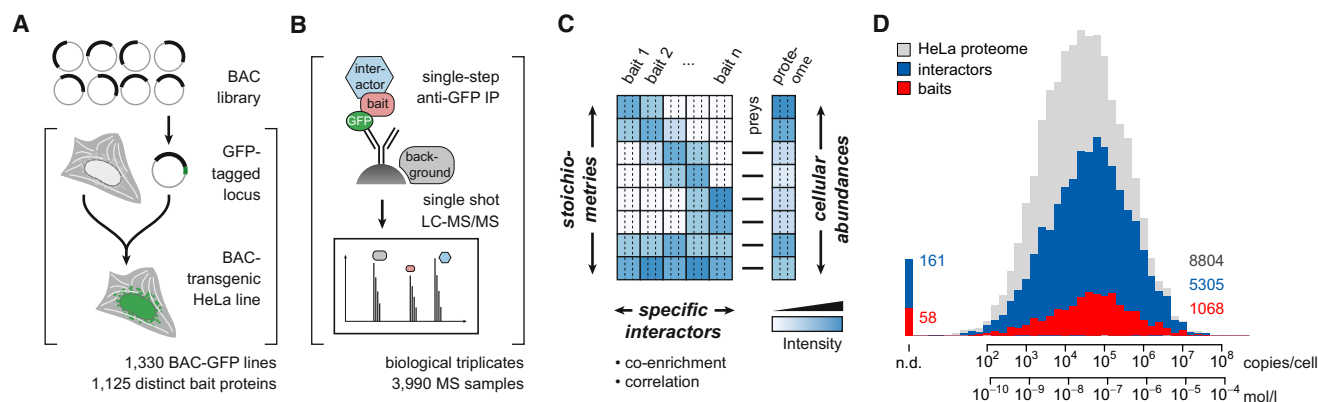
Figure 1. Quantitative BAC-GFP Interactomics
(A) BAC recombineering workflow for generating transgenic HeLa lines.
(B) Single-step affinity-purification, single-run liquid chromatography-tandem mass spectrometry (LC-MS/MS) workflow.
(C) Schematic protein quantification matrices in interactome and proteome samples with three dimensions of quantification.
(D) Proteome coverage and abundance distribution of the bait proteins and their interactors.
See also Tables S1, S2, and S3.

cell lines with the quantitative proteomics workflow resulted in a versatile and highly specific method that we termed quantitative BAC-GFP interactomics (QUBIC) (Hubner et al., 2010).

Here, we applied QUBIC in a proteome-wide manner, using 1,125 bait proteins to assemble a large-scale map of the human interactome. We characterize individual interactions in three quantitative dimensions that address statistical significance, interaction stoichiometry, and cellular abundances of interactors. This concept provides a unique perspective on the interactome, enabling the discovery and characterization of stable and transient protein complexes, guiding their functional interpretation and shedding light on the topological architecture of the entire network.

RESULTS

Quantitative BAC-GFP Interactomics
Collections of strains or cell lines expressing tagged proteins are indispensable tools for many systems biology approaches (Huh et al., 2003). Expressing GFP-tagged proteins from engineered BAC transgenes maintains the endogenous promoters, intron-exon-structures and regulatory elements, ensuring near-endogenous expression levels and patterns (Poser et al., 2008) (Figure S1A). We have previously used this system to study chromosome segregation and the function of motor proteins (Hutchins et al., 2010; Maliga et al., 2013). To map the protein interactome globally, we generated a resource of 1,330 stable BAC-GFP HeLa cell lines (Figure 1A; Table S1). Mouse BACs are excellent surrogates for their human orthologs and offer additional options, such as resistance to RNAi against their endogenous counterparts, streamlining functional studies of the tagged proteins (Kittler et al., 2005). In 615 cell lines, we used mouse BACs with a median sequence identity of 94% with their respective orthologs (Figure S1B). Overall, our collection encompasses 1,125 distinct bait proteins across all protein classes (Figures S1C–S1E), some present as C- and N-terminally

tagged versions, or as mouse and human sequences (not counted as distinct).

We performed QUBIC in three biological replicate experiments, resulting in 3,990 LC-MS runs recorded on an Orbitrap mass spectrometer, taking about a year of net measuring time (Figure 1B). To define specific interactors, we employed MaxLFQ, the label-free quantification (LFQ) module of the MaxQuant software (Cox and Mann, 2008; Cox et al., 2014). Bait proteins and their interactors are characterized by quantitative co-enrichment compared to their intensity profiles across many samples (Figures 1B and 1C), and we used generic statistical testing to determine significantly enriched cases (Keilhauer et al., 2015). To set thresholds for accepting a given candidate as an interactor, we developed an entirely data-driven, false discovery rate (FDR)-controlled approach that harnesses the absence of "negative" interactions and the concomitant asymmetry of the outlier population (Figures S1F and S1G). This approach does not rely on reference datasets or prior knowledge for training, but nonetheless validates favorably against gold standards (Figures S1I–S1K).

In addition to local co-enrichment, we found the intensity profiles of interacting proteins to be closely correlated globally (Supplemental Experimental Procedures). Profile correlations alone can indicate protein interactions when proteome samples are subjected to extensive native fractionation (Havugimana et al., 2012; Kristensen et al., 2012). Here, we use them as additional classifiers (Keilhauer et al., 2015) and the combination of enrichment FDRs and profile correlation coefficients defines the confidence class of each interaction (Figure S1H).

Overall, using the information in this first dimension of proteomic quantification, our analysis resulted in 28,504 unique and statistically significant interactions involving 5,462 distinct proteins (Table S2).

Interaction Stoichiometries and Protein Abundances
A second dimension of quantification can in principle be applied to determine the stoichiometries of proteins within complexes.
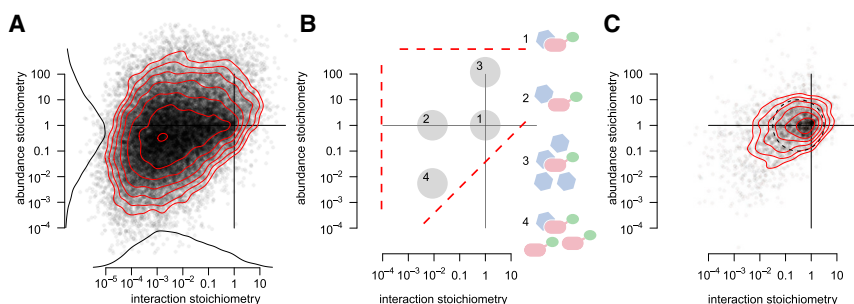
Figure 2. The Stoichiometry Plot

(A) Overlay of all interaction and abundance stoichiometry data for all interactions.

(B) The characteristic triangular shape is a consequence of the dynamic range limits in the interactome (left border), in the proteome (top border) and the stoichiometry limit imposed by the relative cellular protein abundances (diagonal). Schematic interaction scenarios: (1) equal cellular abundance, stable interaction; (2) equal cellular abundance, weak interaction; (3) stable interaction with greater cellular abundance of the prey; and (4) reciprocal case: quantitative recovery of a stably bound, less abundant prey.

(C) Stoichiometry plot of interactions between proteins annotated as CORUM complex members. The area of highest density can be approximated by a circle containing 58% of CORUM interactions.

See also Data S1.

These can be computationally extracted from label-free affinity purification data with accuracies reaching those of methods using isotopically labeled reference standards (Wepf et al., 2009; Smits et al., 2012). If a protein complex contained one copy of each subunit, one might expect them to be retrieved in equimolar amounts after immunoprecipitation (IP). However, in practice, measured stoichiometries between preys and baits span orders of magnitude (Collins et al., 2013; Hauri et al., 2013). This is because the observed stoichiometries depend on more than the initial composition of the individual complexes in the cell. For instance, limited kinetic and thermodynamic stability can result in substoichiometric recovery. Proteins may also reside in different alternative molecular assemblies with fractions of their total cellular pools. Hence, we hypothesized that globally, interaction stoichiometries might reflect the stability of a given protein-protein interaction and depend on the extent that interactors are engaged with each other. The cellular abundance of an interactor can be limiting for how much is recoverable after immunoprecipitation (IP), setting a lower bound for the interaction stoichiometry. We therefore reasoned that cellular copy numbers would provide a crucial third quantitative dimension.

For each pair of interacting proteins, we first quantified their stoichiometry in the immunoprecipitates using a sophisticated label-free strategy for absolute protein quantification (Supplemental Experimental Procedures). To determine the precision of our method, we systematically compared interaction stoichiometries from experiments where the same bait proteins were tagged on different termini, representing entirely separate experiments (Figures S2A and S2E). Stoichiometries showed high correlation, precision within a factor of three, and no systematic bias for a given terminus. This confirmed that our approach robustly delivers interaction stoichiometries in high throughput; however, these may not always be sufficiently accurate to reliably specify copy numbers of each subunit, even for stable complexes. We repeated the analysis for cases where either the mouse or human ortholog of a protein was used as bait, demonstrating the same level of reliability and no species bias (Figures S2B–S2D). This highlights the extraordinary degree of conservation of protein function in evolutionary time (Kachroo et al., 2015) and suggests that our human-centric dataset is representative not only of the human but other mammalian species.

To add a third dimension of proteomic quantification, we next performed a whole proteome quantification experiment on the parental HeLa cell line that all our BAC-transgenic lines are derived from, to a depth of about 9,000 proteins. To estimate cellular protein abundances, we applied our label-free approach and scaled the values to copies per cell using the "proteomic ruler" concept (Wiśniewski et al., 2014) (Supplemental Experimental Procedures). The proteome dataset provided cellular copy numbers for 5,305 proteins of the interactome dataset, covering 97% of all interactors (Figure 1D). The abundances of interacting proteins closely follow the distribution of bait abundances, covering the entire dynamic range of the proteome. This demonstrates that our BAC-based system recapitulates the in vivo situation, enabling us to probe the interactome as a function of the endogenous cellular proteome.

Quantifying the Interactome in Two Additional Dimensions

Having established a set of specific interactions with the first dimension of quantification, we next combined our second and third dimension of quantification, namely the interaction stoichiometries and relative cellular abundances of interactors. A plot of the stoichiometry landscape for each bait protein is a powerful tool to organize its interactome, because each region reflects a different scenario (Figures 2A and 2B): stable, one-to-one, and fully recovered complexes in which the partners have equal cellular abundance appear around the origin of the plot (case 1 in Figure 2B). Superstoichiometry, the recovery of more prey than bait, is only expected for stable complexes containing more prey than bait copies and indeed we find few of these. If interactions are weak and complexes dissociate partially during IP, or if interactions involve only part of the bait pool, interactors are recovered at substoichiometric levels (case 2), reflecting lower occupancy of interaction interfaces of the bait. A vast predominance of sub- over superstoichiometry confirms our initial hypothesis that stability and occupancy are the main determinants for most interactions.

We observed many cases of stable interactors (~1:1 interaction stoichiometry) that involved a more abundant prey (case 3), such as the interaction of the abundant GTP-binding protein RAN with its guanine-nucleotide releasing factor RCC1 or that of α-tubulin with the NEK9 kinase (see Table S2). The reciprocal
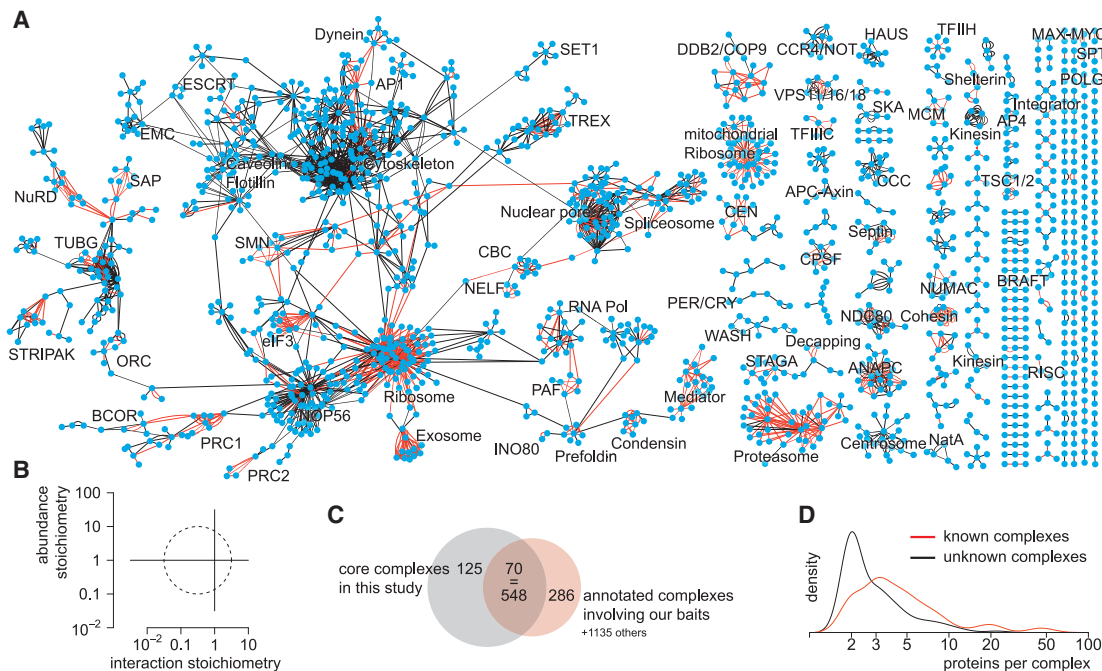
Figure 3. The Core-Complex Network

(A) Sub-networks of interactions matching the CORUM-characteristic core stoichiometry signature. Red edges are known interactions annotated in UniProt or CORUM.

(B) Graphical definition of the core stoichiometry signature. Center: −0.5, 0; radius: 1 ($\log_{10}$ units).

(C) Quantification of the CORUM overlap. A total of 125 isolated networks remain unannotated; 70 networks are annotated with 548 partially redundant CORUM terms; 286 terms assigned to our baits were not shared by any interactor.

(D) Size distribution of annotated versus unannotated networks.

See also Figure S3.

interaction stoichiometry readouts are necessarily smaller than one, because any higher abundant bait can maximally recover the entire pool of its lower abundant prey (case 4). (Note that this would be the default case for overexpressed baits.) We retrieved substoichiometric interactions over an estimated five orders of magnitude; for example, NEK9 was recovered at $6 \times 10^{-6} \times$ the amount of α-tubulin. The proteome-interactome relationship requires that interactors can only be recovered to the extent permitted by their abundance and translates into a diagonal cut-off in the plot, which results in a characteristic triangular shape of the "cloud" of interactions (Figure 2A).

Approximately 10% of our interactions connected members of well-characterized complexes annotated in the CORUM database (Ruepp et al., 2010). They populate a confined area characterized by a signature of balanced stoichiometries (case 1 in Figure 2B). Thus the prototypical case of a stable protein complex as typically described in the literature mostly consists of proteins of equal cellular abundances that are all constitutively bound to each other.

Extrapolating from the signature of known complexes, we reasoned that deduction of similar complexes should be possible solely from the stoichiometry signature of individual baits as opposed to analysis of the entire network (Collins et al., 2007; Hart et al., 2007). We filtered our data for those featuring the core stoichiometry signature (Figure 3B), yielding

a larger cluster connecting several molecular assemblies such as major cytoskeletal proteins, the nuclear pore complex and the ribosome as well as 194 isolated putative core complexes (Figure 3A). These recapitulated the majority of CORUM-annotated complexes that involve our bait proteins (Figure 3C). We confirmed the known tendency of large complexes to be well annotated (Havugimana et al., 2012), while smaller assemblies lacked previous description (Figure 3D). The largest of our 125 networks with no database annotation at the time is the recently discovered COMMD/CCDC22/CCDC93 (CCC) complex (Phillips-Krawczak et al., 2015).

The stoichiometry plot offers a unique opportunity for comparing the overlap of our dataset with published data (Figures S3A–S3C). For instance, the intersection with a recent co-fractionation interactome study (Havugimana et al., 2012) closely recapitulated the core-complex signature, with 26% of our core-interactions overlapping with that study. This indicates that the co-fractionation methodology offers an attractive short-cut to finding stable, obligate core complexes. Conversely, the overlap with iRefWeb, a portal of consolidated protein interactions from different sources (Turner et al., 2010), reached much further into the substoichiometric region, beyond stable complexes, but still only covered 16% of our dataset. Finally, the overlap with recent large-scale yeast-two-hybrid data (Rolland et al., 2014) was low (0.4%) and mostly limited to cases characterized by quantitative

prey recovery to the extent permitted by cellular abundance. Moreover, the stoichiometry plot quantitatively confirmed the intuitive notion that high-stoichiometry interactions are easier to detect as they are enriched in the 1% FDR compared to the 5% FDR cohort (Figures S3D and S3E). This is also reflected in the overlap of gene ontology (GO) annotations in pairs of interacting proteins (Figure S3F).

Interactions Explain Phenotypes and Genetic Associations

Our dataset provides an extensive resource that can be mined for new or poorly characterized protein interactions. For instance, among the interactors of SUCO, one other protein, TAPT1, stood out by its core stoichiometry signature, suggesting a novel, stable complex consisting of these two low abundant integral membrane proteins of the ER (Figures 4A, 4B, and S4A–S4C). Mutants of their murine orthologs exhibit severe defects during skeletal development: Truncation of TAPT1 causes transformations in the axial skeleton and perinatal lethality (Howell et al., 2007), whereas loss of the SUN domain-containing ossification factor SUCO (also known as OPT) impairs postnatal bone formation, causing fractures and neonatal death (Sohaskey et al., 2010). The latter study linked the phenotype to impaired rough ER expansion and consequent failure of osteoblasts to secrete collagen required for bone formation. Knockdown of human SUCO increased the cells' resistance against ricin, whose toxicity depends on endocytosis and retrograde trafficking to the ER (Bassik et al., 2013). Similarly, the yeast ortholog of TAPT1, EMP65 (YER140W), is involved in protein folding in the ER and shows buffering genetic and physical interactions with the SUN domain protein SLP1 (YOR154W) (Jonikas et al., 2009; Friederichs et al., 2012). We used our interaction methodology on GFP-tagged strains to confirm this complex (Figures S4D and S4E). Similarly, we validated the reciprocal interaction in the mammalian system using TAPT1 as bait (Figure S4B). Together, our findings establish TAPT1–SUCO as the higher eukaryote ortholog of SLP1-EMP65: a low abundant ER membrane complex that is required for normal skeletal development.

Going beyond stable complexes, we discovered an interaction between the anaphase promoting complex or cyclosome (APC/C) and the uncharacterized protein KIAA1430. The stoichiometry plot indicated that KIAA1430 is of lower cellular abundance and is not an obligate member of the APC/C, as the partners were recovered substoichiometrically at ~1% of the respective baits in reciprocal experiments (Figures 4C and 4D).

To independently test whether KIAA1430 was indeed a transient interactor of the APC/C, we performed a purify-after-mixing (PAM)-SILAC experiment (Wang and Huang, 2008) (Figures 4E and 4F; Supplemental Experimental Procedures). We mixed differentially SILAC-labeled lysates from tagged and control cell lines before the affinity step. Subsequently measured SILAC ratios are indicative of the stability of the interaction, because transient interactors exchange dynamically with unbound counterparts, shifting their ratio toward unity, whereas stable interactors maintain their label ratio. Our results confirmed that only known subunits of the APC/C are stably bound to the core subunit CDC23. Consistently, only some of them were recovered when assayed for binding to KIAA1430 and with ratios indicating

a high degree of dynamic exchange. Next, we tested whether KIAA1430 is a substrate of the APC/C by monitoring its levels during mitosis and early G1 phase. Unlike known substrates, KIAA1430 levels remained stable (Figure S4F).

In interphase, a fraction of GFP-tagged KIAA1430 localized to the centrosomes, in particular the centrioles, and was largely excluded from the nucleus (Figures 4G, 4H, S4G, and S4H), while the APC/C is known to be predominantly nuclear (Kraft et al., 2003; Hubner et al., 2010). During mitosis, after nuclear envelope breakdown (NEBD), APC/C accumulates on mitotic spindles, centromeres, and centrosomes (Kraft et al., 2003; Acquaviva et al., 2004), reflecting a partially common localization with KIAA1430. Consistently, we confirmed the APC/C–KIAA1430 interaction in mitotically arrested, but not in interphase cells (Figure 4I). To functionally investigate the mitotic interaction, we used time-lapse microscopy to determine the time cells require from NEBD to the onset of anaphase as a function of APC/C activity. KIAA1430 knockdown resulted in a mild delay that was sensitive to reversine, a small molecule inhibitor of the mitotic checkpoint kinase MPS1 (Figures 4J and 4K) (Santaguida et al., 2010). These findings suggest that the depletion of KIAA1430 activates the spindle assembly checkpoint, thereby postponing the activation of the APC/C. Recent reports identified the ciliary protein hemingway as the *Drosophila* ortholog of KIAA1430 (Soulavie et al., 2014) and implicated the APC/C in regulating ciliary length and polarity (Ganner et al., 2009; Wang et al., 2014). Given that centrioles are common features of cilia and centrosomes, our data suggest that in human cells, KIAA1430 recruits a sub-fraction of the APC/C to the centrosome to facilitate mitotic progression.

These examples illustrate how the combination of three quantitative dimensions offers a unique view on the interactions of individual proteins that extends beyond their identification and facilitates their functional investigation.

We have compiled this information into an easily usable resource, provided as Data S1 and available via the IntAct database. For each of the 1,330 tagged cell lines, we present a concise, one-page summary outlining the abundance of the bait protein, the co-enrichment and confidence classification of candidate interactors along with the stoichiometry plot and the predictions of the core complexes. A reading guide is presented in Figure S5.

The Relevance of Substoichiometric Interactions

Our study revealed that interactions within obligate complexes constitute only a small minority of the interactome. We reasoned that the majority of remaining interactions should be of a functionally and conceptually different nature, as indicated by our example of the KIAA1430-APC/C interaction.

To investigate the interplay of the different types of interactions, we interrogated the chaperonin TRiC (also called CCT), which is known to act on a large number of client proteins (Hartl et al., 2011). Its core machinery of eight subunits was clearly identified as an abundant obligate complex (Figure 5A) and represents a prominent hub in our interactome dataset. Virtually all interactors co-enriched with tagged TRiC core subunits were co-chaperones, regulatory proteins of the phosducin family or proteins containing known substrate motifs (Yam et al., 2008) (Table
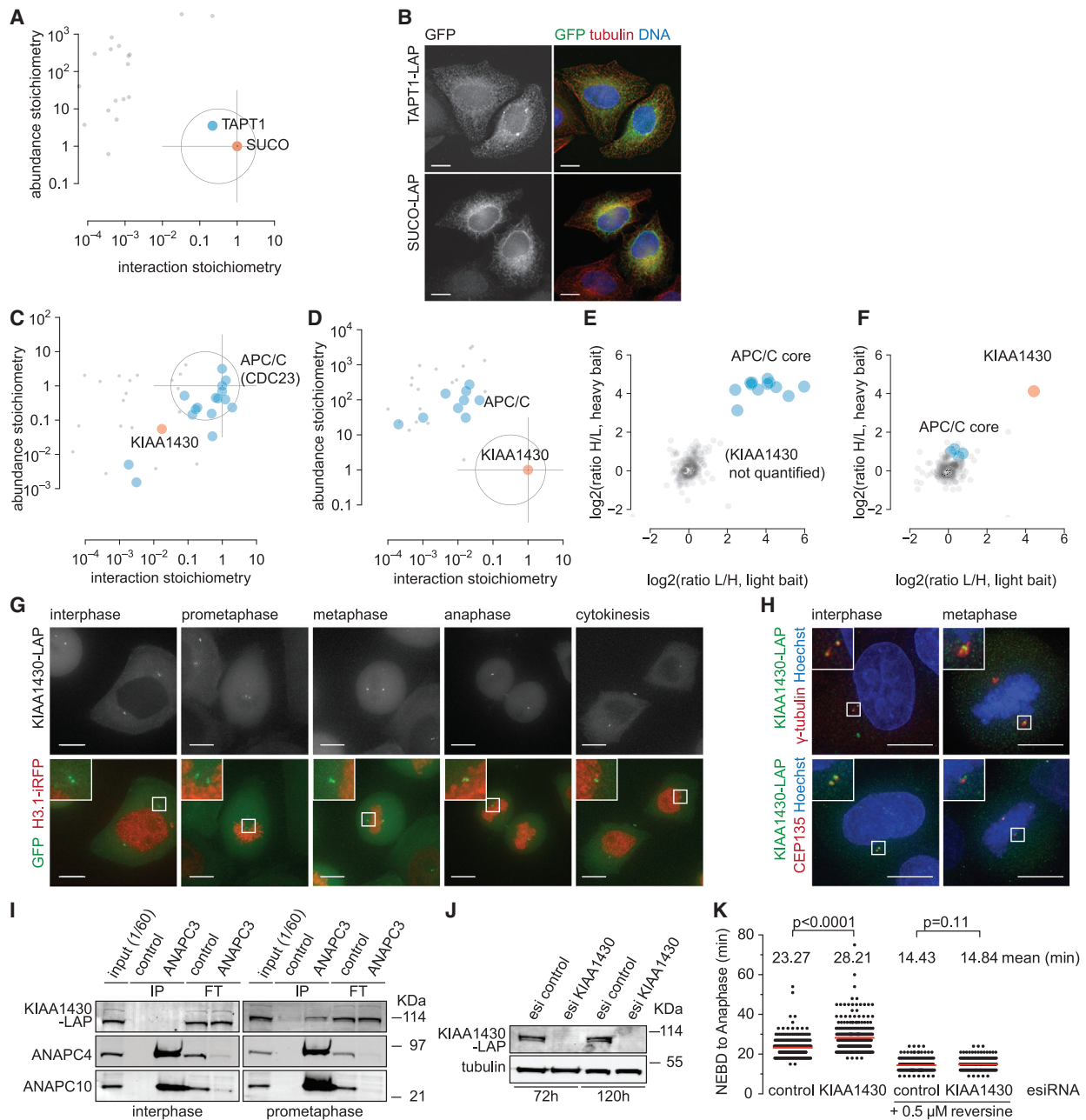
Figure 4. The TAPT1-SUCO Complex and the KIAA1430-APC/C Interaction

(A) Stoichiometry plot indicates stable TAPT1-SUCO complex.

(B) Immunofluorescence of TAPT1 and SUCO in HeLa shows ER localizations.

(C) Stoichiometry plot of APC/C interactors (bait: CDC23). Known core complex members (blue) and KIAA1430 (red) as novel substoichiometric interactor.

(D) Stoichiometry plot of KIAA1430 (red) interactors shows APC/C subunits (blue) as substoichiometric interactors.

(E) PAM-SILAC ratios plotted as medians of forward triplicate against label-swapped reverse triplicate (bait: CDC23).

(F) PAM-SILAC data using KIAA1430 as bait. Baits and stable interactors are recovered at ratios corresponding to label incorporation levels. Ratios of transient interactors are shifted toward 1:1.

(G) Maximum intensity projections of living interphase and mitotic cells expressing KIAA1430-LAP and histone 3.1-iRFP indicate that KIAA1430 localizes to centrioles.

(H) Co-localization of KIAA1430 with centrosomal marker γ-tubulin and the centriolar protein CEP135.

(I) Western blot analysis of ANAPC3 IPs and corresponding flow-throughs (FT) from interphase and mitotically arrested cells expressing KIAA1430-LAP.

(legend continued on next page)
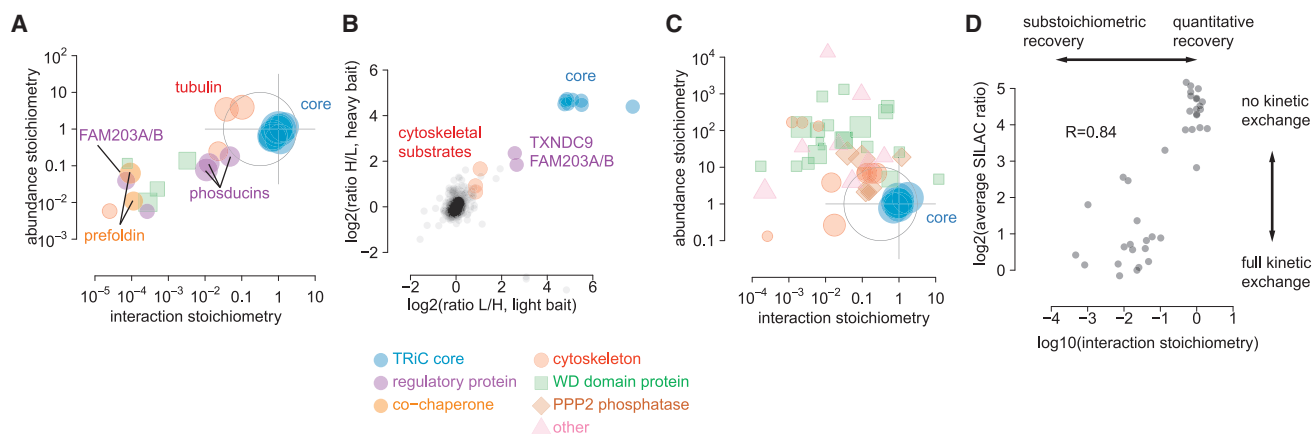
Figure 5. The TRiC Interactome Is Defined by Substoichiometric Links

(A) Stoichiometry plot of CCT3 interactors, representative of TRiC core subunits.

(B) PAM-SILAC results from the same bait protein.

(C) Reciprocal stoichiometry plot of averaged positions of the TRiC subunits from all bait pull-downs enriching at least three TRiC subunits. Symbol size indicates profile correlation. See also Tables S4 and S5.

(D) Systematic comparison of interaction stoichiometries and PAM-SILAC ratios for all interactions observed using CDC23, KIAA1430, and CCT3 as baits.

S4). Characteristic of all was a lower cellular abundance than TRiC (except for some cytoskeletal proteins) and substoichiometric recovery, classifying these interactors as distinct from the core subunits. When we performed a PAM-SILAC experiment to test for stable versus transient binding, the core complex composition that we had already established by the stoichiometry plot was confirmed, as these were all found to be stable binders (Figure 5B). Other interactors were transient, as their SILAC ratios indicated full dynamic exchange. Notable exceptions were some regulatory proteins and abundant cytoskeletal substrates, whose ratios lay between stable and fully dynamic binders. We consistently found the uncharacterized protein FAM203A/B as a substoichiometric interactor with intermediate dynamic exchange behavior. Its ortholog in *Caenorhabditis elegans* shows a cytoskeletal knockdown phenotype (Fievet et al., 2013). All of this was reminiscent of phosducin proteins, which TRiC requires to fold actin and tubulin (Hayes et al., 2011) and we therefore speculate that FAM203A/B might have a similar function.

In reciprocal interaction experiments, TRiC core complex members were co-enriched by ~5% of all bait proteins (Figure 5C; Table S5). This is in line with estimates of TRiC being involved in folding of 5%–10% of the proteome (Hartl et al., 2011). However, only some of these baits were also found in the reciprocal TRiC IPs. This asymmetry can be explained with knowledge of the underlying proteome: At 1.3 million copies of the hexadecameric complex, TRiC is much more abundant than most substrates, of which only a fraction will be in the process of folding at any given time (Table S3). Consequently, only a minute fraction of the TRiC pool will be acting on each substrate

and its recovery will be "diluted" to substoichiometric levels. In the reciprocal case, however, TRiC occupies a significant fraction of the client protein population—the fraction in the state of folding—rendering the interaction more readily detectable within the dynamic range (Figure 5C).

The stoichiometry of TRiC recovery in the substrate IPs ranges from less than 10^{-3} to above 10^{-1} (Figure 5B). With TRiC substrates comprising 5%–10% of all protein molecules (a HeLa cell contains an estimated at 6×10^9 protein molecules), our stoichiometry data imply that on average 0.2%–0.4% of them are bound to the chaperone at any time. While substoichiometry may be thought to be of lower biological relevance, these interactions fulfil important functions as they connect very diverse set of protein classes. Moreover, our data also illustrate how the proteome-interactome relationship balances the amount of TRiC with the cumulative amount of its substrates.

Extrapolating from our APC/C, KIAA1430, and TRiC case studies, we investigated whether the different stoichiometric classifications of interactions carry over to other characteristics: first, we systematically compared interaction stoichiometries with dynamic exchange data for all interactions for which both orthogonal pieces of information were available (Figure 5D). There was almost perfect congruence of stoichiometric interactors with kinetically stably bound proteins and a surprisingly good overall correlation of substoichiometric recovery and the extent of dynamic exchange. This indicates that interaction stoichiometries are globally predictive of the biophysical stability of an interaction. Next, we investigated whether interaction stoichiometry is indicative of co-expression across tissues or cell types. We extracted protein abundance correlation profiles across

(J) Western analyses showing the extent of KIAA1430 depletion before and after the time-lapse analyses presented in (K).

(K) Time KIAA1430-depleted cells require to proceed from NEBD to anaphase, compared to control cells (n = 300 each); 0.5 μM reversine rescues the delay (n = 200 each). Red lines, mean. Significance according to two-tailed Mann-Whitney test. Scale bars, 10 μm.

See also Figure S4.
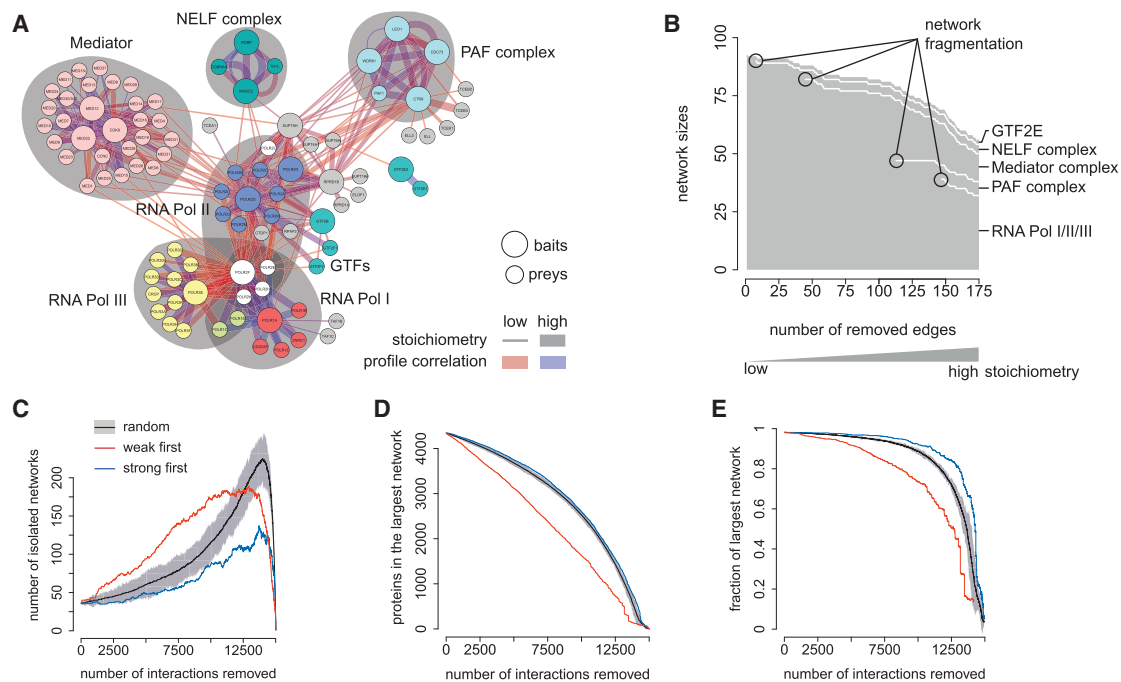
Figure 6. Strong and Weak Interactions Have Different Global Properties

(A) Sub-network of complexes surrounding RNA polymerases I/II/III. Proteins are colored by complex, edges by profile correlation, edge widths represent interaction stoichiometries.

(B) Effect of sequential removal of substoichiometric interactions on network sizes. Indicated are points where edge removal results in two fragmented sub-networks.

(C) Global network effect of random or targeted removal of interactions on the total number of isolated sub-networks.

(D) Effect on the number of proteins present in the largest entirely connected sub-network.

(E) Effect on the fraction of total connected proteins that are part of this largest sub-network.

See also Figure S6.

many tissues from a recent human proteome draft dataset (Kim et al., 2014). While co-expression coefficients scattered widely, there was still a notable relationship with interaction stoichiometry, with high-stoichiometry interactors more likely to be coherently expressed (Figure S6A). This is in agreement with earlier findings in yeast showing that members of stable complexes are enriched in co-regulated modules (Simonis et al., 2006). Conversely, substoichiometric interactions involve proteins that are not necessarily tightly co-regulated.

Finally, we tested whether interaction stoichiometry is predictive of the role of an interaction in network topology. We analyzed a sub-network of interactors surrounding RNA polymerases I, II, and III, recapitulating shared subunits and interactions with other complexes, such as general transcription factor complexes, the negative elongation factor (NELF) complex, the mediator complex, and the polymerase-associated factor (PAF) complex (Figure 6A). Sequential in silico removal of the most substoichiometric interactions from the network leads to fragmentation events, in which the individual complexes gradually lose their interconnections and emerge as individual modules (Figure 6B). Finally, the three polymerases remain internally connected via their shared subunits. Removing interactions in the reverse order does not lead to any network fragmentation, but rather results in roughly linear shrinkage of the network (Figure S6A).

Taking this approach to a global level, we probed the response of our entire network to the removal of edges according to their stoichiometry characteristics. Seminal studies on the topology of networks have shown that scale-free networks are resilient to random removal of edges, but sensitive to targeted attacks (Albert et al., 2000). Specifically, analysis of the network structure identifies the topologically most critical edges, removal of which leads to rapid network fragmentation.

In our case, we targeted edges for removal solely by their "local" interaction stoichiometry readout, agnostic to their global network roles. We removed edges sequentially, starting at either the lowest or highest interaction stoichiometry, comparing this with random removal of edges.

This revealed vastly different network responses (Figure 6C). The most substoichiometric interactions turned out to be most critical for network topology: Their preferential removal led to a rapid increase of the number of isolated network fragments, whereas removing the strongest 50% of edges hardly resulted in any network fragmentation (Figure 6C). The largest connected component, which causes the typical "hairball" appearance of large-scale networks, shrunk about linearly with removal of weak interactions (Figure 6D) and also left more proteins entirely unconnected (Figure S6A). Conversely, preferential removal of edges from the other end of the stoichiometry scale led to a

network response that increased its small-world characteristics: the largest network encompasses the vast majority of connected proteins (Figure 6E), fewer proteins are left without connections (Figure S6B) and isolated network fragments are smaller (Figure S6C). Similar patterns of network response were observed in a study analyzing mobile phone communication networks by removing the strongest versus the weakest interactions (Onnela et al., 2007). In analogy to that study, and based on our findings of the relationship between interaction stoichiometries and kinetic stabilities (Figure 5D), we propose a strong/weak terminology for interaction stoichiometries and term interactions with near-stoichiometric recovery of the prey "strong" and substoichiometric interactors "weak."

Together, our analyses show that interaction stoichiometries, which are local properties derived from single interaction experiments, predict the global behavior of the proteins involved: strong interactions are indicative of proteins that are co-regulated across cell types. In the network, they form modules of high interconnectivity, rendering the network topologically resilient to their removal. Weak interactions, on the other hand, dominate the network both in numbers and by their topologically critical role as long-range interactions between more diverse sets of proteins. As a consequence, interaction networks can be fragmented into individual, defined modules, by identifying and removing weak links. In summary, availability of interaction stoichiometries on a global scale effectively allows us to "comb" the interactome hairball, to identify modules, and visualize their interconnectedness.

DISCUSSION

Here, we have introduced a novel concept of interactome analysis. Using an efficient, low-stringent IP protocol, accurate label-free quantification of both the IPs and the complete proteome, we extracted three quantitative dimensions, all of which proved critical for characterizing protein interactions. While the first dimension identifies statistically significant interactions, the second and third dimension define their stoichiometric contexts. Earlier large-scale studies did not include or interpret these additional dimensions, in part because of the challenges involved in extracting accurate quantitative values. Moreover, past studies often employed overexpression of bait proteins, precluding meaningful stoichiometry readout (Gibson et al., 2013), and near-complete proteome coverage was also often not attainable.

Finding stable protein complexes is usually a major goal of interactomics studies. We showed that obligate protein complexes feature a unique signature of balanced stoichiometries—an infrequent occurrence among the multitude of interactions. Such a signature led us to discover the TAPT1-SUCO complex in the ER membrane. This complex ties together a body of available evidence, including knockout phenotypes of both TAPT1 and SUCO and genetic interactions of their yeast orthologs. As a representative of the majority of weaker, non-obligate interactions, we characterized the binding of KIAA1430 to the APC/C, suggesting that low interaction stoichiometries are the result of an interaction that is limited to centrioles in mitotic cells and biophysically weaker than interactions between APC/C core members. Furthermore, our stoichiometry-based classification subdivided the interactome of the TRiC chaperonin into obligate core complex subunits, regulatory interactors, and a large number of substrates. We found that lack of reciprocal verification can be indicative of an inherently asymmetric nature of biologically relevant interactions, particularly outside obligate core complexes. This example also illustrates how the observed interactome is shaped by protein abundances and, conversely, implies overall regulation of protein abundances by protein interactions. Therefore, the interactome always has to be interpreted as a function of the underlying proteome.

We have shown that interaction stoichiometries generally correlate with the biophysical stability of an interaction. Weak interactions have frequently gone undetected in interactome studies and may be thought to be less important; nevertheless they are crucial features of networks in general and social networks in particular (Granovetter, 1973; Csermely, 2006). Our study directly and quantitatively demonstrates the predominance of weak interactions in the protein interactome. MS-based methods cover more than four orders of magnitude of interaction stoichiometry (Collins et al., 2013), and our low-stringency biochemical workflow ideally harnesses this sensitivity. However, substoichiometric interactions involving low abundance preys can still be challenging to detect (Figures S3D and S3E). Therefore, the prevalence of weak interactions is likely to be even more pronounced and their relevance vastly underappreciated.

Previous studies typically counted all interactions as equal, once they had been accepted based on their statistical parameters or scores. Therefore, the roles of individual interactions had to be predicted from prior knowledge or from global network properties. Highly connected proteins were described as interaction hubs, regions of high clustering coefficients with many shared pathway annotations were characterized as complexes (Collins et al., 2007; Hart et al., 2007), and weak interactions were inferred from weaker connectivity patterns (Malovannaya et al., 2011). However, limited coverage of the interactome is a confounding factor for such strategies.

In contrast, we here have shown directly that local stoichiometry data reflect global network topological properties of interactions, setting the stage for quantitative network analysis from the ground up.

Substoichiometric interactions form the "glue" that holds the cellular network together—as shown specifically for the RNA polymerase network and globally for the entire network—and are hence critical for network structure. This property, which may seem counterintuitive at first, prompted us to propose interaction stoichiometry as a measure of interaction strength. Of note, a range of underlying mechanisms can cause a weak interaction according to this terminology, for instance low biophysical affinity, high kinetic exchange rates, limited spatiotemporal overlap of interactors, or indirect interactions that are bridged via other biomolecules, all of which may result in a substoichiometric readout. If such weak links are removed from the network, it collapses into defined modules that are tightly interconnected by the remaining strong links. Translated into biological terms, stable complexes would remain in isolation, but without weak links, they would not be able to connect to each other or to transient, dynamic regulators.

A major contribution of this study lies in the characterization of the interactomes surrounding more than 1,100 different baits, which together cover a large part of the expressed proteome with more than 28,000 interactions. We present our results in an accessible format that can be easily mined and interpreted by non-specialists. Our resource of mammalian cell lines expressing GFP-tagged proteins under endogenous control can be employed for other studies (e.g., focusing on subcellular localization or functional characterization of individual proteins). The interaction data validate these cell lines for such uses and the use of mouse orthologs as surrogates. They also imply remarkably similar protein interactomes between human and mouse.

We approach saturation with respect to the number of proteins that can be covered (Figure S6D), but observe only part of the entire interactome directly, which our data predicts to encompass between 80,000 and 180,000 detectable interactions in HeLa (Figure S6E). Our additional quantitative dimensions may prove helpful for increasing interactome coverage in silico, for example, by selective matrix expansion (Seebacher and Gavin, 2011). Given its usefulness in interpreting interaction data, the stoichiometry readout developed here can become a general basis for future interactome studies and for the analysis of interactome dynamics, which will manifest foremost as quantitative alteration of occupancies rather than qualitative gain or loss of interactors.

EXPERIMENTAL PROCEDURES

Cell Culture

HeLa Kyoto cell lines expressing N- or C-terminally tagged proteins from BAC transgenes were generated, cultured, and imaged as previously described (Poser et al., 2008). Tags are based on the "localization and affinity purification" (LAP) tag, consisting of GFP and a functionalized linker. All BAC cell lines and tag sequences are listed in Table S1 along with proteome and interactome metadata on the bait proteins. Cells were grown to near-confluency on two 15-cm cell culture dishes per interaction experiment, detached with Accutase, and snap frozen. Three replicates were harvested in at least two different passages.

Affinity Purification and Mass Spectrometry

Cell pellets were lysed and subjected to affinity purification on a robotic system, followed by single-shot mass spectrometric analysis on an Orbitrap instrument (Hubner et al., 2010). We processed triplicates separately on different days and carried out MS-analyses in randomized order over the course of weeks to months.

Whole Proteome Measurements

HeLa cells were lysed in guanidinium chloride lysis buffer and digested sequentially with LysC and trypsin as described (Kulak et al., 2014). Peptides were desalted on stacked C18 reverse phase (Waters Sep-Pak) and strong cation exchange cartridges and eluted using 70% acetonitrile. Pooled eluates were separated into six fractions on strong anion exchange (SAX) StageTips (Wiśniewski et al., 2010). MS measurements were performed in three replicates on a quadrupole Orbitrap mass spectrometer (Kulak et al., 2014).

Data Processing

Raw files were processed with MaxQuant (Cox and Mann, 2008) (version 1.3.9.10) in several sets, each containing ∼600 randomly assigned AP-MS runs and the HeLa proteome fractions. Tandem mass spectrometry (MS/MS) spectra were searched against a modified version of the November 2012 release of the UniProt complete human proteome sequence database. For each bait protein expressed from a mouse BAC locus, the human sequence in the fasta file was concatenated with the mouse sequence (unless identical). Human identifiers were used for mapping purposes. We used MaxLFQ, MaxQuant's label-free quantification (LFQ) algorithm to calculate protein intensity profiles across samples (Cox et al., 2014). We required one ratio count for each pairwise comparison step and activated the FastLFQ setting with two minimum and two average comparisons to enable the normalization of large datasets in manageable computing time.

Detection of Protein Interactions

Protein identifications were filtered, removing hits to the reverse decoy database as well as proteins only identified by modified peptides. We required that each protein be quantified in all replicates from the AP-MS samples of at least one cell line. Protein LFQ intensities were logarithmized and missing values imputed by values simulating noise around the detection limit. For each protein, a non-parametric method was used to select a subset of samples that provide a distribution of background intensities for this protein (Supplemental Experimental Procedures). This subset was used first to normalize all protein intensities to represent relative enrichment and then to serve as the control group for a two-tailed Welch's t test. Specific outliers in the volcano plots of logarithmized p values against enrichments were determined by an approach making use of the asymmetry in the outlier population (Figures S1E and S1F). We used two cut-offs of different stringencies, representing 1% and 5% of enrichment false discovery rate (FDR), respectively. Correlation coefficients between the intensity profiles of interacting proteins were calculated as additional quality parameters (Keilhauer et al., 2015). Enrichment FDR (classes A–C) and profile correlation (modifier + or –) define the confidence class of an interaction (Figure S1G).

Interaction Stoichiometries and Cellular Copy Numbers

Estimating interaction stoichiometries requires the comparison of the amounts of different proteins relative to each other in one IP. We first subtracted the median intensity across all samples to account for background binding. We then divided LFQ intensities by the number of theoretically observable peptides for this protein (Schwanhäusser et al., 2011). Finally, we expressed stoichiometries relative to the bait protein. Cellular copy numbers and abundances were calculated using a similar approach (Wiśniewski et al., 2014) on the whole proteome data and brought to absolute scale by normalization to a total protein amount of 200 pg in a cell volume of 1 pl for a HeLa cell.

Network Analyses

Network analyses were performed based on the data listed in Table S2. For the purpose of counting unique interactions and for the histogram of the numbers of interactors, we regarded interactions as non-directional, flattened multiple protein groups mapping to the same gene name and considered interactions found multiple times only once. For network perturbation analyses, we selected all non-self-interactions of confidence classes A+, A, and B+ and assembled them into graphs. We then removed edges sequentially according to their interaction stoichiometry readout. Prey-bait combinations discovered multiple times were treated as separate edges. Once a protein had lost all its edges, it was removed. As control, we deleted edges randomly and represented the median of 100 random repetitions and represent the scatter as the first or third quartile ±1.5 interquartile ranges.

ACCESSION NUMBERS

The accession number for the protein interaction data reported in this paper, submitted through IntAct (Orchard et al., 2014), to the IMEx Consortium (http://www.imexconsortium.org): IM-24272. The accession numbers for the raw mass spectrometric data, deposited via PRIDE (Vizcaino et al., 2013) to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org): PXD002815 and PXD002829.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, five tables, and one data set and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.09.053.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Acquaviva, C., Herzog, F., Kraft, C., and Pines, J. (2004). The anaphase promoting complex/cyclosome is recruited to centromeres by the spindle assembly checkpoint. Nat. Cell Biol. 6, 892–898.

Albert, R., Jeong, H., and Barabasi, A.L. (2000). Error and attack tolerance of complex networks. Nature 406, 378–382.

Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal. Bioanal. Chem. 404, 939–965.

Bassik, M.C., Kampmann, M., Lebbink, R.J., Wang, S., Hein, M.Y., Poser, I., Weibezahn, J., Horlbeck, M.A., Chen, S., Mann, M., et al. (2013). A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell 152, 909–922.

Beck, M., Claassen, M., and Aebersold, R. (2011). Comprehensive proteomics. Curr. Opin. Biotechnol. 22, 3–8.

Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., and Krogan, N.J. (2007). Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol. Cell. Proteomics 6, 439–450.

Collins, B.C., Gillet, L.C., Rosenberger, G., Röst, H.L., Vichalkovski, A., Gstaiger, M., and Aebersold, R. (2013). Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. Nat. Methods 10, 1246–1253.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics 13, 2513–2526.

Csermely, P. (2006). Weak Links: Stabilizers of Complex Systems from Proteins to Social Networks, First Edition (Springer).

Ewing, R.M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M., et al. (2007). Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol. Syst. Biol. 3, 89.

Fievet, B.T., Rodriguez, J., Naganathan, S., Lee, C., Zeiser, E., Ishidate, T., Shirayama, M., Grill, S., and Ahringer, J. (2013). Systematic genetic interaction screens uncover cell polarity regulators and functional redundancy. Nat. Cell Biol. 15, 103–112.

Friederichs, J.M., Gardner, J.M., Smoyer, C.J., Whetstine, C.R., Gogol, M., Slaughter, B.D., and Jaspersen, S.L. (2012). Genetic analysis of Mps3 SUN domain mutants in Saccharomyces cerevisiae reveals an interaction with the SUN-like protein Slp1. G3 (Bethesda) 2, 1703–1718.

Ganner, A., Lienkamp, S., Schäfer, T., Romaker, D., Wegierski, T., Park, T.J., Spreitzer, S., Simons, M., Gloy, J., Kim, E., et al. (2009). Regulation of ciliary polarity by the APC/C. Proc. Natl. Acad. Sci. USA 106, 17799–17804.

Gibson, T.J., Seiler, M., and Veitia, R.A. (2013). The transience of transient overexpression. Nat. Methods 10, 715–721.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nat. Rev. Mol. Cell Biol. 8, 645–654.

Granovetter, M.S. (1973). The Strength of Weak Ties. Am. J. Sociol. 78, 1360–1380.

Hart, G.T., Lee, I., and Marcotte, E.R. (2007). A high-accuracy consensus map of yeast protein complexes reveals modular nature of gene essentiality. BMC Bioinformatics 8, 236.

Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. Nature 475, 324–332.

Hauri, S., Wepf, A., van Drogen, A., Varjosalo, M., Tapon, N., Aebersold, R., and Gstaiger, M. (2013). Interaction proteome of human Hippo signaling: modular control of the co-activator YAP1. Mol. Syst. Biol. 9, 713.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. Cell 150, 1068–1081.

Hayes, N.V.L., Jossé, L., Smales, C.M., and Carden, M.J. (2011). Modulation of phosducin-like protein 3 (PhLP3) levels promotes cytoskeletal remodelling in a MAPK and RhoA-dependent manner. PLoS ONE 6, e28271.

Hein, M.Y., Sharma, K., Cox, J., and Mann, M. (2013). Proteomic Analysis of Cellular Systems. In Handbook of Systems Biology, Chapter 1, Walhout A.J.M., Vidal M., and Dekker J., eds. (Academic Press), pp. 3–25.

Howell, G.R., Shindo, M., Murray, S., Gridley, T., Wilson, L.A., and Schimenti, J.C. (2007). Mutation of a ubiquitously expressed mouse transmembrane protein (Tapt1) causes specific skeletal homeotic transformations. Genetics 175, 699–707.

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. J. Cell Biol. 189, 739–754.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature 425, 686–691.

Hutchins, J.R., Toyoda, Y., Hegemann, B., Poser, I., Hériché, J.K., Sykora, M.M., Augsburg, M., Hudecz, O., Buschhorn, B.A., Bulkescher, J., et al. (2010). Systematic analysis of human protein complexes identifies chromosome segregation proteins. Science 328, 593–599.

Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S., and Schuldiner, M. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science 323, 1693–1697.

Kachroo, A.H., Laurent, J.M., Yellman, C.M., Meyer, A.G., Wilke, C.O., and Marcotte, E.M. (2015). Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science *348*, 921–925.

Keilhauer, E.C., Hein, M.Y., and Mann, M. (2015). Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). Mol. Cell. Proteomics *14*, 120–135.

Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., et al. (2014). A draft map of the human proteome. Nature *509*, 575–581.

Kittler, R., Pelletier, L., Ma, C., Poser, I., Fischer, S., Hyman, A.A., and Buchholz, F. (2005). RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. Proc. Natl. Acad. Sci. USA *102*, 2396–2401.

Kraft, C., Herzog, F., Gieffers, C., Mechtler, K., Hagting, A., Pines, J., and Peters, J.M. (2003). Mitotic regulation of the human anaphase-promoting complex by phosphorylation. EMBO J. *22*, 6598–6609.

Kristensen, A.R., Gsponer, J., and Foster, L.J. (2012). A high-throughput approach for measuring temporal changes in the interactome. Nat. Methods *9*, 907–909.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat. Methods *11*, 319–324.

Maliga, Z., Junqueira, M., Toyoda, Y., Ettinger, A., Mora-Bermúdez, F., Klemm, R.W., Vasilj, A., Guhr, E., Ibarlucea-Benitez, I., Poser, I., et al. (2013). A genomic toolkit to investigate kinesin and myosin motor function in cells. Nat. Cell Biol. *15*, 325–334.

Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W., Ding, C., Shi, Y., Yucer, N., Krenciute, G., et al. (2011). Analysis of the human endogenous coregulator complexome. Cell *145*, 787–799.

Mann, M., Kulak, N.A., Nagaraj, N., and Cox, J. (2013). The coming age of complete, accurate, and ubiquitous proteomes. Mol. Cell *49*, 583–590.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods *10*, 730–736.

Onnela, J.P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A.L. (2007). Structure and tie strengths in mobile communication networks. Proc. Natl. Acad. Sci. USA *104*, 7332–7336.

Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., del-Toro, N., et al. (2014). The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. *42*, D358–D363.

Phillips-Krawczak, C.A., Singla, A., Starokadomskyy, P., Deng, Z., Osborne, D.G., Li, H., Dick, C.J., Gomez, T.S., Koenecke, M., Zhang, J.S., et al. (2015). COMMD1 is linked to the WASH complex and regulates endosomal trafficking of the copper transporter ATP7A. Mol. Biol. Cell *26*, 91–103.

Poser, I., Sarov, M., Hutchins, J.R., Hériché, J.K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A.W., et al. (2008). BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat. Methods *5*, 409–415.

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., et al. (2014). A proteome-scale map of the human interactome network. Cell *159*, 1212–1226.

Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. Nature *437*, 1173–1178.

Ruepp, A., Waegele, B., Lechner, M., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Mewes, H.W. (2010). CORUM: the comprehensive resource of mammalian protein complexes–2009. Nucleic Acids Res. *38*, D497–D501.

Santaguida, S., Tighe, A., D'Alise, A.M., Taylor, S.S., and Musacchio, A. (2010). Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. J. Cell Biol. *190*, 73–87.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Seebacher, J., and Gavin, A.C. (2011). SnapShot: Protein-protein interaction networks. Cell *144*, 1000.

Simonis, N., Gonze, D., Orsi, C., van Helden, J., and Wodak, S.J. (2006). Modularity of the transcriptional response of protein complexes in yeast. J. Mol. Biol. *363*, 589–610.

Smits, A.H., Jansen, P.W., Poser, I., Hyman, A.A., and Vermeulen, M. (2012). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. *41*, e28.

Sohaskey, M.L., Jiang, Y., Zhao, J.J., Mohr, A., Roemer, F., and Harland, R.M. (2010). Osteopotentia regulates osteoblast maturation, bone formation, and skeletal integrity in mice. J. Cell Biol. *189*, 511–525.

Soulavie, F., Piepenbrock, D., Thomas, J., Vieillard, J., Duteyrat, J.L., Cortier, E., Laurençon, A., Göpfert, M.C., and Durand, B. (2014). hemingway is required for sperm flagella assembly and ciliary motility in Drosophila. Mol. Biol. Cell *25*, 1276–1286.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell *122*, 957–968.

Turner, B., Razick, S., Turinsky, A.L., Vlasblom, J., Crowdy, E.K., Cho, E., Morrison, K., Donaldson, I.M., and Wodak, S.J. (2010). iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. Database (Oxford) *2010*, baq023.

Vizcaino, J.A., Cote, R.G., Csordas, A., Dianes, J.A., Fabregat, A., Foster, J.M., Griss, J., Alpi, E., Birim, M., Contell, J., et al. (2013). The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res. *41*, D1063–D1069.

Wang, X., and Huang, L. (2008). Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. Mol. Cell. Proteomics *7*, 46–57.

Wang, W., Wu, T., and Kirschner, M.W. (2014). The master cell cycle regulator APC-Cdc20 regulates ciliary length and disassembly of the primary cilium. eLife *3*, e03083.

Wepf, A., Glatter, T., Schmidt, A., Aebersold, R., and Gstaiger, M. (2009). Quantitative interaction proteomics using mass spectrometry. Nat. Methods *6*, 203–205.

Wiśniewski, J.R., Nagaraj, N., Zougman, A., Gnad, F., and Mann, M. (2010). Brain phosphoproteome obtained by a FASP-based method reveals plasma membrane protein topology. J. Proteome Res. *9*, 3280–3289.

Wiśniewski, J.R., Hein, M.Y., Cox, J., and Mann, M. (2014). A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. Mol. Cell. Proteomics *13*, 3497–3506.

Yam, A.Y., Xia, Y., Lin, H.T., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. Nat. Struct. Mol. Biol. *15*, 1255–1262.