Published in final edited form as:

Methods. 2012 December ; 58(4): 392–399. doi:10.1016/j.ymeth.2012.07.015.

Studying protein complexes by the yeast two-hybrid system

Seesandra V. Rajagopala^{1,*}, Patricia Sikorski¹, J. Harry Caufield², Andrey Tovchigrechko¹, and Peter Uetz^{2,*}

¹J Craig Venter Institute, Rockville, MD 20850, USA

²Center for the Study of Biological Complexity, Virginia Commonwealth University, PO Box 842030, 1015 Floyd Ave., Richmond, VA 23284, USA

Abstract

Protein complexes are typically analysed by affinity purification and subsequent mass spectrometric analysis. However, in most cases the structure and topology of the complexes remains elusive from such studies. Here we investigate how the yeast two-hybrid system can be used to analyse direct interactions among proteins in a complex. First we tested all pairwise interactions among the seven proteins of *Escherichia coli* DNA polymerase III as well as an uncharacterized complex that includes MntR and PerR. Four and seven interactions were identified in these two complexes, respectively. In addition, we review Y2H data for three other complexes of known structure which serve as “gold-standards”, namely Varicella Zoster Virus (VZV) ribonucleotide reductase (RNR), the yeast proteasome, and bacteriophage lambda. Finally, we review an Y2H analysis of the human spliceosome which may serve as an example for a dynamic mega-complex.

Keywords

Y2H; ribonucleotide reductase; DNA polymerase; MntR/PerR complex; phage lambda; proteasome

Introduction

Proteins act by either interacting with other proteins or small molecules such as sugars. In fact, many enzymes that interact with their low molecular weight substrates also form protein complexes, such as fatty acid synthase, a complex of 12 chains [1]. In the past 15 years high-throughput methods have been developed to map both protein-protein interactions on a large scale, using either the yeast two-hybrid system for binary interactions [2–4] or affinity purification and mass spectrometry for protein complexes [5–8]. In principle, a two-hybrid screen using all the proteins of a complex should yield all the binary interactions within that complex, but this is rarely the case; in most cases only a few interactions are discovered [9, 10]. On the other hand, the two-hybrid system has a certain preference for transient interactions that are lost during complex purification because of the necessary washing steps [4]. In addition, two-hybrid screening data derived from genome-scale projects is not complete or/and comprehensive. Thus, little overlap is often observed

© 2012 Elsevier Inc. All rights reserved.

*corresponding authors: rajgsv@gmail.com, peter@uetz.us.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

between the protein interaction datasets generated by protein-complex purification and two-hybrid studies [4].

In order to infer direct interactions from complex purification data, either the matrix or spoke model has been applied to lists of co-purified proteins. More recent protein complex purification studies used the *socio-affinity index* (SAI) to infer the direct interactions between complex members [5–8]. A related strategy uses data from complex purification data sets to identify interacting proteins [11]. These strategies are based on the observation that certain pairs of proteins are more frequently found in multiple purifications than others, and are thus predicted to be closer or even directly associated in the complex. Similarly, solutions have been presented to computationally identify complexes from Y2H interaction networks [12–15].

While these experimental and computational attempts to map protein-protein interactions have produced a massive amount of data, structural analysis of complexes and interacting proteins has lagged behind. One of the goals of PPI studies is thus to identify complexes that are amenable for crystallization. Often, several strategies need to be combined to reconstruct the topology of complexes that cannot be crystallized, including proteomics, cryo-electron microscopy and others.

Recently, we have developed multiple variants of the yeast two-hybrid system and shown that different two-hybrid systems detect markedly different subsets of interactions in the same interactome [16]. Ten different configurations of bait-prey fusions were required to detect up to 67% of a set of gold-standard interactions, whereas individual vector pairs detected only 25% on average [17].

Here we describe and review a similar strategy, using yeast two-hybrid assays to map the interactions within several complexes. In addition, we analyze several well-characterized protein complexes, ranging from ribonucleotide reductase (four subunits) to spliceosomes (>200 proteins), and compare structural data to published Y2H data. Our data shows that a majority of interactions in a complex can be identified by systematic Y2H screening and that Y2H assays often detect subcomplexes within a larger complex that may be amenable to crystallography while the whole complex may be not crystallizable.

Materials and Methods

Yeast two-hybrid (Y2H) assays, All full-length ORF clones of two protein complexes (Tables 1 and 2, Fig. 1) were selected from the *E. coli* ORFeome collection [18]. The entry clones were transferred into prey and bait yeast two-hybrid expression vectors using Gateway LR reactions (namely pGADT7g and pGBGT7g for N-terminal fusions and pGBKCg-atg for C-terminal fusions; the latter vector having an in-frame start codon for ORFs that do not have one) [19]. After bacterial transformation, miniprep plasmid DNA of all two-hybrid prey and bait clones were transformed into yeast two-hybrid strains *MATa* Y187 (prey clones) and *MATa* AH109 (bait clones). Yeast two-hybrid matrix screening was conducted as described in [20]. In brief, a yeast strain expressing a single protein as a bait fusion was mated to individual preys of all other protein complex members (Fig. 1A,B). After mating, the colonies were transferred to the Y2H protein-protein interaction selective medium, and the interacting protein pairs were identified by the resulting positive yeast colony. The positive interactions show a clear colony growth at a certain level of 3-Amino-1,2,4-triazole, whereas no growth was usually seen in the negative control (auto-activation), i.e. the bait mated with the empty prey vector strain (Fig. 1B).

Three Dimensional Modeling: Homology models of the VZV R1 (ORF19) and R2 (ORF18) proteins were produced with I-TASSER server [21] and superimposed onto the 2BQ1 asymmetric holocomplex structure from PDB (<http://www.rcsb.org>). The final model was rendered in Pymol (<http://www.pymol.org>).

Results and Discussion

Bacterial complexes

The *Escherichia coli* DNA polymerase III complex—In order to validate how well multiple variants of the two-hybrid system work for mapping the topology of protein complexes, we have selected the *Escherichia coli* DNA polymerase III complex as an example of a well-characterized protein complex (protein complex 42 in [7]). Hu et al. affinity-purified this complex and identified seven subunits, namely, AsnB, DnaE, DnaX, HolC, HolD, HolE and TrmD. We did not include other proteins known to be more loosely associated with DNA polymerase III such as the epsilon subunit (DnaQ) which catalyzes the 3' to 5' proofreading exonuclease activity of the holoenzyme.

Escherichia coli DNA polymerase contains two sub-complexes: the catalytic polymerase/exonuclease sub-complex (with alpha, beta, delta and epsilon subunits), plus the DnaX complex, a heptamer that includes the tau and gamma products of the dnaX gene and confers structural asymmetry that allows the polymerase to replicate both leading and lagging strands. We have curated all the known protein-protein interactions between co-purified complex members from several studies (Fig. 1C, Table 2).

In order to validate whether we can capture all the known interactions of the *E. coli* DNA polymerase III proteins, we subjected its proteins to an Y2H matrix screening. The yeast two-hybrid assay was able to detect 70% known interactions (seven out of 10 interactions, including homo-dimers) and was able to connect four protein (out of seven proteins) in the complex compared to five proteins connected by literature data (Fig. 1C). Some of the known interactions are not reproduced here, for example the theta (HolE) subunit of DNA polymerase III binds to the epsilon subunit (DnaQ) but not to the alpha subunit (DnaE) [22]. This binding appears to enhance the interaction between alpha (DnaE) and epsilon (DnaQ) as well as slightly stimulating epsilon activity [23]. Since DnaQ is not a member of the protein complex described here, the interaction between HolE → DnaQ → DnaE has not been tested. These results demonstrate that Y2H screening can detect most direct interactions within a complex and aid the mapping of its topology.

The MntR complex is a largely uncharacterized protein complex that has been identified as “complex 34” by AP/MS [7]. This complex consists of eight proteins of which five proteins have been characterized and three uncharacterized proteins (Table 3, Fig. 1D). Even though the complex contains several well-characterized proteins the direct interaction and functional associations among the complex members are unknown.

The Y2H screening of all the members of this complex identified seven interactions (plus four homo-dimers) and based on these interactions we were able to precisely map the topology of the complex (Fig. 1B,D), confirming that these proteins actually form a complex. These protein-protein interactions should help to characterize the function of uncharacterized proteins in the complex. For example, the interaction between the transcriptional regulators MntR and PerR supports the predicted DNA-binding and transcriptional activity of the latter. While PerR has been studied in *Bacillus subtilis*, its ortholog in *E. coli* is poorly understood. PerR and related members of the LysR family have been shown to interact with other members of the family to form heterodimers, but the physiological significance of this is unknown [24]. In order to study the function of these

complexes in more detail, the topology of interactions suggests several strategies to crystallize this complex, e.g. either as dimers (YjhC-YjhC), trimers (MntR-HycG-MetN or MntR-MetN-YjhC), tetramer (MntR-HycG-MetN-YjhC), or as hetero-hexamers. Such crystallization experiments are now under way.

Previously published complexes—In addition to the new interactions in the previous section, we have also re-visited the literature on several protein complexes that have been studied by Y2H methods. They cover sizes ranging from ribonucleotide reductase, a small protein complex of four subunits, to the spliceosome, one of the largest complexes in living systems with hundreds of proteins. These complexes are discussed in the order of increasing complexity.

Ribonucleotide Reductase—Ribonucleotide reductases (RNRs) convert ribonucleotides to deoxyribonucleotides and thus provide the raw materials for DNA synthesis. All living systems that are based on DNA as their genetic material, either encode RNRs, or must obtain their dNTPs from an outside source, as several parasites and endosymbionts do [25]. Most eukaryotes and some of their viruses encode so-called type I RNRs that are characterized by a $\alpha_2\beta_2$ quaternary structure [26].

RNRs are interesting complexes for several reasons (reviewed in [25–27]). First, they all appear to be homologous, yet show a variety of quaternary structures and enzymatic mechanisms. For instance, class I RNRs form $\alpha_2\beta_2$ quaternary structures while class II enzymes are either α monomers or α_2 dimers. Second, different species use different cofactors (e.g. thioredoxin vs. formate), are aerobic or anaerobic, and have different modes of interaction. Third, they are highly dynamic complexes that undergo conformational changes during a reaction cycle, which also affects their interactions. Fourth and last, RNRs and their subunit interactions are affected by allosteric regulators such as dGTP.

In the course of our studies on human herpesviral interactomes [28, 29], we have analyzed RNRs of several viruses. However, only the RNR of Varicella Zoster Virus (VZV) was analyzed with multiple Y2H vectors (Fig. 2, [16]) which allowed us to draw some more general conclusions about interactions that can be detected within protein complexes. First, under favorable conditions, majority of interactions in a complex can be detected by Y2H assays. In the case of VZV RNR, all interactions among the four RNR subunits were found: R1-R1, R1-R2, and R2-R2. Second, full-length proteins not necessarily work in such assays. Here, the full-length R2 protein neither interacted with R1 nor with itself. However, fragments clearly worked in multiple orientations. As is shown in Fig. 2B, N- and C-terminal fragments of R2 interacted in a homo-dimeric fashion (ORF18C-ORF18C), as pseudo-homo-dimers (ORF18N-ORF18C), as well as true hetero-dimers (ORF18C-ORF19). Third, and finally, the N- and C-terminal fusions of the Gal4 DNA-binding and activation domains had very distinct patterns of interaction. For instance, while an N-terminal fusion of ORF18C interacted with almost all other proteins, both N- and C-terminal fusions of ORF18N interacted very specifically only with ORF18C.

These observations are largely supported by the crystal structure of the RNR holocomplex (Fig. 2A) [30, 31], even though the biological complex appears to be a dynamic assembly of proteins in which the R2 subunits move and rotate [32]. This fact has impaired crystallization of the biological complex and it is still not entirely clear how the subunits move during a catalytic cycle *in vivo*.

Our homology model superimposed on the available asymmetric template (Fig. 2A) shows C- and N-terminals either located completely away from the interface areas or having likely

disordered/flexible ends, thus leaving enough room for accommodating the Y2H fusion constructs in all cases.

The proteasome—The proteasome is a complex assembly of proteins found within all eukaryotic cells and some bacterial species. In its fully-assembled state, this protein complex recycles peptides within the cell by degrading misfolded or otherwise unnecessary proteins. Each full 26S proteolytic complex is formed by a 20S core particle (CP) and at least one regulatory particle (RP), including six AAA-ATPase subunits and thirteen non-ATPase subunits. In eukaryotes, each of these 26S proteasome particles may be found in the cytoplasm and the nucleoplasm [33]. Several sets of yeast two-hybrid analyses have been used alone and in concert with other methods to clarify the protein interactions required for assembly of a functional proteasome. Work by Cagney et al. [34] employed a yeast two-hybrid approach to search for interactions between 31 known proteasome proteins and an array of nearly 6000 different *S. cerevisiae* proteins. This genome-wide screen revealed 55 potential protein-protein interactions, more than a third of which involved only proteasome components rather than non-proteasomal proteins. The specific pairs of interactions found in these screens demonstrated how many proteins of the 19S and 20S subcomplexes specifically interact with other proteins within those subcomplexes (Fig. 3A,B). One interaction between a 20S α ring protein, Pre8, and a 20S β ring protein, Pup1, was observed, confirming structural arrangements seen in the 20S crystal structure [35]. Some interactions did occur between subunits, as comparisons with the crystal structure of the 20S core showed. Out of 14 interacting protein pairs in the 20S subcomplex, just 3 involved proteins not predicted to be neighbors in the crystal structure. As Cagney et al. studied yeast proteasomal proteins in an Y2H system; it is quite likely that endogenous proteasomal proteins bridged the two hybrid proteins. This background may have contributed to artifactual interactions not seen in the 20S crystal structure. Use of the Y2H system may therefore be better suited to the study of heterologous protein pairs from sources other than yeast.

Work by Davy et al. [36] used the Y2H system to generate interaction data for 30 proteasome subunits from *Caenorhabditis elegans*. Each of the proteins assayed in this study are orthologous to those of *Saccharomyces cerevisiae*. Though this study found many interactions correlating with those seen in with yeast proteasome proteins, a number of new interactions were also observed (Fig. 3). Among these are four interactions between 19S subunit proteins and three 20S α subunits.

An independent study of the *human* proteasome [37] also revealed protein-protein interactions by Y2H, finding 114 potential interacting pairs. Many of these interactions were found to be structurally relevant. Interactions between the α and β rings of the 20S core were observed, confirming observations by Cagney et al. (Fig. 3B, Supplementary Table 1). Subsequent studies have confirmed many of the interactions revealed by Y2H assays using other methods. A mass spectrometry-based approach [38] generated 64 potential proteasome-associated protein-protein interactions beyond those seen between proteasome subunit proteins.

Unfortunately, the fragile nature of the 26S proteasome has rendered it difficult to crystallize for extensive X-ray crystallographic analysis. The 20S core particle crystal structure prepared by Groll et al. [35] provides an example of many potential protein interactions on the basis of the barrel-like structure of this subunit, but it provides no evidence for protein interactions beyond those of the core particle. Combined data from crystallography, cryoelectron microscopy, and protein co-purifications have been used to clarify proteasomal structural organization predicted by Y2H results [39–41]. One such model prepared by Lasker et al. displays the overall architecture of the 26S proteasome holoenzyme. While the

overall structure is based on cryo-EM mapping and molecular modeling (including a model prepared by Förster et al. [42]), the specific locations of each protein are based on crystal structures, residue-specific lysine crosslinks, and known protein-protein interactions, including those found through Y2H screens.

In summary, the three Y2H studies found a total of 183 PPIs, while structural studies revealed a total of 38 PPIs. No Y2H study found more than 40% of the interactions assumed to take place in the proteasome, but all three studies together found 79% of all interactions because each study found a different subset (Supplementary Table 1). This is certainly due to the fact that proteins from different species as well as different Y2H systems were used.

Clearly, Y2H screens produced a substantial number of false positives in these studies, (as shown by interactions not seen in the proteasome structure), but some of them may be truly physiological interactions (i.e. taking place *in vivo*), given that the α and β subunits of the proteasome are closely related proteins that probably interact in more than the canonical combinations even within the assembled proteasome (Fig. 3).

Notably, cross-linking studies found a total of 32 interactions, of which nine (32%) were seen in the structure. However, these studies may have not been as comprehensive as the Y2H screens, but it is remarkable that they also produced a number of false positives that is similar to those found in Y2H screens.

Phage Lambda—Bacteriophage lambda virion consists of ~14 different proteins and a total of ~1075 subunits. After its discovery in 1960 hundreds of studies have revealed its structure in great detail and we know now that its subunits are connected by at least a dozen different protein-protein interactions (Fig. 4A,B). Interestingly, not even in this extremely well-studied phage all interactions are known for sure. Rajagopala et al. [43] therefore curated the literature and compiled a list of 33 published PPIs among lambda proteins (note that this list includes PPIs among proteins that are *not* in the particle). This set of 33 known interactions is considered as a “gold-standard” set here. These authors then tested all possible protein pairs encoded by the lambda genome for interactions using a matrix-based Y2H screen that employed 6 different Y2H vectors (Fig. 4). Several lessons can be learned from this analysis. First, the screen detected more than half of all previously known interactions (including 4 interactions among regulatory proteins not present in the virion). Rajagopala et al. speculate that the remaining interactions were not detected because of the lack of chaperones, assembly factors, post-translational modifications, or other effects. Second, each vector pair detected a certain fraction of the known “gold-standard” set of PPIs; which interaction was detected depended primarily on the vector pair used (Fig. 4C). For instance, pGBKT7/pGADC produced the largest absolute number of “gold-standard” interactions although the pDEST vectors produced the largest *fraction* of gold-standard interactions. 11 out of 16 interactions were detected only by one vector pair, namely pGBKT7g/pGADCg (5 PPIs), pDEST (4), and pGBKCG/pGADCg (2). The other 5 PPIs were detected with multiple vectors each. Interestingly, only 4 of the gold-standard PPIs were detected exclusively with N-terminal fusions, the system with which the vast majority of all Y2H screens are carried out. Two gold-standard PPIs were detected exclusively with C-terminal fusions. The majority was detected with either NC fusions or multiple vectors.

The spliceosome—The spliceosome is probably the largest protein complex known, at least in terms of complexity, containing more than 200 different proteins as well as multiple RNAs [44]. It is thus much bigger than the ribosome. Although spliceosomal proteins assemble into smaller complexes, such as the U1, U2, U4/6 and U5 snRNPs, they transiently associate with each other during the process of splicing, which makes them a “functional” complex. Given its complexity, the spliceosome has not been crystallized and there are only

certain subcomplexes whose structure is reasonably well-known, such as the U1 snRNP [45]. Hegele et al. [46] recently presented a systematic analysis of protein-protein interactions among 244 spliceosomal proteins, of which 141 proteins are classified as “core proteins”, based on their high abundance. Initially, Hegele et al. cloned 244 proteins known to be associated with the human spliceosome into yeast two-hybrid bait and prey vectors and tested them in a pairwise fashion for interactions. This study found a total of 632 interactions among 196 of the 244 proteins. Notably, 390 interactions were among non-core and 242 interactions were found among core proteins. Importantly, this study also curated 311 binary interactions previously published in 201 papers. Of these 311 PPIs, 72 were reproduced by Hegele’s Y2H analysis. While this number corresponds to only 23% of the published PPIs, the fraction of reproduced PPIs among the core proteins was 41% (43 PPIs). Reproducing only a quarter of all published interactions certainly have the same reasons as discussed above: first, only one vector pair was used in the Y2H analysis. Second, splicesomal interactions may depend on both RNA support as well as tertiary interactions. Third, many interactions are dynamic and thus rather weak, and may require additional factors such as assembly proteins or other catalysts [47]. The Hegele paper unfortunately did not attempt to integrate their data into structural models of the spliceosomal subunits, so it remains unclear to what extent these interactions facilitate a structural understanding of the spliceosome. However, many interactions suggested local subcomplexes that can be subjected to crystallography and other analyses.

A good example how Y2H screens can dissect the interactions with a large complex is an analysis of the U1 snRNP. This subcomplex of the spliceosome consists of 10 U1 proteins and a ring of seven Sm proteins in yeast. Several groups have solved the overall structure of the complex by cryo-EM analysis and crystallography [48, 49]. However, their model of the *human* U1 snRNP contains only three U1 proteins and the Sm ring (7 proteins). It remains unclear where the remaining proteins are that are clearly parts of the *yeast* U1 snRNP. Ester & Uetz [50] showed that a subcomplex of 3 proteins (Snu71, Prp40, and Luc7) can be detected within the yeast U1 snRNP and these authors mapped their interaction domains. It is hoped that such subcomplexes are stable enough for crystallization, so that their atomic structure can be resolved eventually.

Conclusions

From the data discussed here, it becomes clear that the Y2H system can contribute significantly to the understanding of protein complexes. However, there is also much room for improvements. First, multiple vectors need to be used routinely to achieve maximum coverage. In addition, the existing vectors can be certainly further improved. Notably, vector variations as mentioned here have not even been used for alternative Y2H systems such as the split-ubiquitin system or protein fragment complementation assays. Second, and similar to crystallography, it is likely that comparative studies will shed much light on protein complexes. For instance, while crystallization of *E. coli* and eukaryotic ribosomes was unsuccessful for many decades, the ribosomes (or ribosomal subunits) of other species could be crystallized more easily. Similarly, it is quite likely that many interactions may be easier to detect with proteins from non-model organisms although this issue has not been systematically studied on a larger (experimental) scale [51]. That said, it will be interesting to see how complexes and their interactions look in different, distantly related species. There is certainly much to be discovered by comparative analysis using improved methodology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

SVR and PU were supported by NIH grant R01GM079710; AT was supported by NSF grant 1048199.

Abbreviations

AD	Activation domain
AP/MS	affinity purification/mass spectrometry
DBD	DNA-binding domain
PPI	protein-protein interaction
RNR	ribonucleotide reductase
SAI	<i>socio-affinity index</i>
VZV	Varicella Zoster Virus
Y2H	yeast two-hybrid

References

- Jenni S, Leibundgut M, Boehringer D, Frick C, Mikolasek B, Ban N. Science. 2007; 316:254–261. [PubMed: 17431175]
- Titz B, Rajagopala SV, Goll J, Hauser R, McKevitt MT, Palzkill T, Uetz P. PLoS ONE. 2008; 3:e2292. [PubMed: 18509523]
- Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamar G, Yang M, Johnston M, Fields S, Rothberg JM. Nature. 2000; 403:623–627. [PubMed: 10688190]
- Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, Sahalie J, Hirozane-Kishikawa T, Gebreab F, Li N, Simonis N, Hao T, Rual JF, Dricot A, Vazquez A, Murray RR, Simon C, Tardivo L, Tam S, Svrikapa N, Fan C, de Smet AS, Motyl A, Hudson ME, Park J, Xin X, Cusick ME, Moore T, Boone C, Snyder M, Roth FP, Barabasi AL, Tavernier J, Hill DE, Vidal M. Science. 2008; 322:104–110. [PubMed: 18719252]
- Butland G, Peregrin-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A. Nature. 2005; 433:531–537. [PubMed: 15690043]
- Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G. Nature. 2006; 440:631–636. [PubMed: 16429126]
- Hu P, Janga SC, Babu M, Diaz-Mejia JJ, Butland G, Yang W, Pogoutse O, Guo X, Phanse S, Wong P, Chandran S, Christopoulos C, Nazarians-Armavil A, Nasser NK, Musso G, Ali M, Nazemof N, Eroukova V, Golshani A, Paccanaro A, Greenblatt JF, Moreno-Hagelsieb G, Emili A. PLoS Biol. 2009; 7:e96. [PubMed: 19402753]
- Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandhi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF. Nature. 2006; 440:637–643. [PubMed: 16554755]
- Rajagopala SV, Titz B, Goll J, Parrish JR, Wohlbold K, McKevitt MT, Palzkill T, Mori H, Finley RL Jr, Uetz P. Mol Syst Biol. 2007; 3:128. [PubMed: 17667950]

10. Rajagopala SV, Hughes KT, Uetz P. *Proteomics*. 2009; 9:5296–5302. [PubMed: 19834901]
11. Schelhorn SE, Mestre J, Albrecht M, Zotenko E. *Mol Cell Proteomics*. 2011; 10:M110 004929. [PubMed: 21451165]
12. Hirsh E, Sharan R. *Bioinformatics*. 2007; 23:e170–176. [PubMed: 17237088]
13. Li WY, Liu Y, Huang HC, Peng YX, Lin YJ, Ng WK, Ong KL. *Ieee-Acm Transactions on Computational Biology and Bioinformatics*. 2007; 4:233–250.
14. Spirin V, Mirny LA. *Proc Natl Acad Sci U S A*. 2003; 100:12123–12128. [PubMed: 14517352]
15. Zheng H, Wang H, Glass DH. *Ieee Transactions on Systems Man and Cybernetics Part B- Cybernetics*. 2008; 38:5–16.
16. Stellberger T, Hauser R, Baiker A, Pothineni VR, Haas J, Uetz P. *Proteome Sci*. 8:8. [PubMed: 20205919]
17. Chen YC, Rajagopala SV, Stellberger T, Uetz P. *Nat Methods*. 2010; 7:667–668. [PubMed: 20805792]
18. Rajagopala SV, Yamamoto N, Zweifel AE, Nakamichi T, Huang HK, Mendez-Rios JD, Franca-Koh J, Boorgula MP, Fujita K, Suzuki K, Hu JC, Wanner BL, Mori H, Uetz P. *Bmc Genomics*. 2010; 11:470. [PubMed: 20701780]
19. Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S, Vidal M. *Methods Enzymol*. 2000; 328:575–592. [PubMed: 11075367]
20. Rajagopala SV, Uetz P. *Methods Mol Biol*. 2011; 781:1–29. [PubMed: 21877274]
21. Zhang Y. *BMC Bioinformatics*. 2008; 9:40. [PubMed: 18215316]
22. Studwell-Vaughan PS, O'Donnell M. *The Journal of biological chemistry*. 1993; 268:11785–11791. [PubMed: 8505306]
23. Taft-Benz SA, Schaaper RM. *J Bacteriol*. 2004; 186:2774–2780. [PubMed: 15090519]
24. Knapp GS, Hu JC. *PLoS ONE*. 2010; 5:e15189. [PubMed: 21187915]
25. Nordlund P, Reichard P. *Annu Rev Biochem*. 2006; 75:681–706. [PubMed: 16756507]
26. Lundin D, Torrents E, Poole AM, Sjoberg BM. *Bmc Genomics*. 2009; 10:589. [PubMed: 19995434]
27. Eklund H, Uhlin U, Farnegardh M, Logan DT, Nordlund P. *Prog Biophys Mol Biol*. 2001; 77:177–268. [PubMed: 11796141]
28. Fossum E, Friedel CC, Rajagopala SV, Titz B, Baiker A, Schmidt T, Kraus T, Stellberger T, Rutenberg C, Suthram S, Bandyopadhyay S, Rose D, von Brunn A, Uhlmann M, Zeretzke C, Dong YA, Boulet H, Koegl M, Bailer SM, Koszinowski U, Ideker T, Uetz P, Zimmer R, Haas J. *PLoS Pathog*. 2009; 5:e1000570. [PubMed: 19730696]
29. Uetz P, Dong YA, Zeretzke C, Atzler C, Baiker A, Berger B, Rajagopala SV, Roupelieva M, Rose D, Fossum E, Haas J. *Science*. 2006; 311:239–242. [PubMed: 16339411]
30. Sommerhalter M, Voegtli WC, Perlstein DL, Ge J, Stubbe J, Rosenzweig AC. *Biochemistry*. 2004; 43:7736–7742. [PubMed: 15196016]
31. Uhlin U, Eklund H. *Nature*. 1994; 370:533–539. [PubMed: 8052308]
32. Uppsten M, Farnegardh M, Domkin V, Uhlin U. *J Mol Biol*. 2006; 359:365–377. [PubMed: 16631785]
33. Peters JM, Franke WW, Kleinschmidt JA. *The Journal of biological chemistry*. 1994; 269:7709–7718. [PubMed: 8125997]
34. Cagney G, Uetz P, Fields S. *Physiol Genomics*. 2001; 7:27–34. [PubMed: 11595789]
35. Groll M, Ditzel L, Lowe J, Stock D, Bochtler M, Bartunik HD, Huber R. *Nature*. 1997; 386:463–471. [PubMed: 9087403]
36. Davy A, Bello P, Thierry-Mieg N, Vaglio P, Hitti J, Doucette-Stamm L, Thierry-Mieg D, Reboul J, Boulton S, Walhout AJ, Coux O, Vidal M. *EMBO Rep*. 2001; 2:821–828. [PubMed: 11559592]
37. Chen C, Huang C, Chen S, Liang J, Lin W, Ke G, Zhang H, Wang B, Huang J, Han Z, Ma L, Huo K, Yang X, Yang P, He F, Tao T. *Proteomics*. 2008; 8:508–520. [PubMed: 18186020]
38. Guerrero C, Tagwerker C, Kaiser P, Huang L. *Mol Cell Proteomics*. 2006; 5:366–378. [PubMed: 16284124]

39. Forster F, Lasker K, Nickell S, Sali A, Baumeister W. *Mol Cell Proteomics*. 2010; 9:1666–1677. [PubMed: 20467039]
40. Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, Martin A. *Nature*. 2012; 482:186–191. [PubMed: 22237024]
41. Lasker K, Forster F, Bohn S, Walzthoeni T, Villa E, Unverdorben P, Beck F, Aebersold R, Sali A, Baumeister W. *Proc Natl Acad Sci U S A*. 2012; 109:1380–1387. [PubMed: 22307589]
42. Forster F, Lasker K, Beck F, Nickell S, Sali A, Baumeister W. *Biochem Biophys Res Commun*. 2009; 388:228–233. [PubMed: 19653995]
43. Rajagopala SV, Casjens S, Uetz P. *BMC Microbiol*. 2011; 11:213. [PubMed: 21943085]
44. Markus CW, Cindy LW, Reinhard Lh. *Cell*. 2009; 136
45. Stark H, Dube P, Luhrmann R, Kastner B. *Nature*. 2001; 409:539–542. [PubMed: 11206553]
46. Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Luhrmann R, Stelzl U. *Mol Cell*. 2012; 45:567–580. [PubMed: 22365833]
47. Staley JP, Woolford JL Jr. *Curr Opin Cell Biol*. 2009; 21:109–118. [PubMed: 19167202]
48. Stark H, Dube P, Luhrmann R, Kastner B. *Nature*. 2001; 409:539–542. [PubMed: 11206553]
49. Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. *Nature*. 2009; 458:475–480. [PubMed: 19325628]
50. Ester C, Uetz P. *BMC Biochem*. 2008; 9:29. [PubMed: 19014439]
51. Yu H, Luscombe NM, Lu HX, Zhu X, Xia Y, Han JD, Bertin N, Chung S, Vidal M, Gerstein M. *Genome Res*. 2004; 14:1107–1118. [PubMed: 15173116]
52. Goll J, Rajagopala SV, Shiau SC, Wu H, Lamb BT, Uetz P. *Bioinformatics*. 2008; 24:1743–1744. [PubMed: 18556668]
53. Elkins PA, Watts JM, Zalacain M, van Thiel A, Vitazka PR, Redlak M, Andraos-Selim C, Rastinejad F, Holmes WM. *J Mol Biol*. 2003; 333:931–949. [PubMed: 14583191]
54. Gulbis JM, Kazmirski SL, Finkelstein J, Kelman Z, O'Donnell M, Kuriyan J. *European journal of biochemistry/FEBS*. 2004; 271:439–449. [PubMed: 14717711]
55. Kazmirski SL, Podobnik M, Weitze TF, O'Donnell M, Kuriyan J. *Proc Natl Acad Sci U S A*. 2004; 101:16750–16755. [PubMed: 15556993]
56. Keniry MA, Berthon HA, Yang JY, Miles CS, Dixon NE. *Protein Sci*. 2000; 9:721–733. [PubMed: 10794414]
57. Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J. *Cell*. 2006; 126:881–892. [PubMed: 16959568]
58. Larsen TM, Boehlein SK, Schuster SM, Richards NG, Thoden JB, Holden HM, Rayment I. *Biochemistry*. 1999; 38:16146–16157. [PubMed: 10587437]
59. Rajagopala SV, Goll J, Gowda ND, Sunil KC, Titz B, Mukherjee A, Mary SS, Raviswaran N, Poojari CS, Ramachandra S, Shtivelband S, Blazie SM, Hofmann J, Uetz P. *Bioinformatics*. 2008; 24:2622–2627. [PubMed: 18786976]
60. Wolf DH, Hilt W. *Biochim Biophys Acta*. 2004; 1695:19–31. [PubMed: 15571806]
61. McHenry C, Kornberg A. *The Journal of biological chemistry*. 1977; 252:6478–6484. [PubMed: 330531]
62. Naktinis V, Onrust R, Fang L, O'Donnell M. *The Journal of biological chemistry*. 1995; 270:13358–13365. [PubMed: 7768937]
63. Maki S, Kornberg A. *The Journal of biological chemistry*. 1988; 263:6555–6560. [PubMed: 3283126]
64. Xiao H, Dong Z, O'Donnell M. *The Journal of biological chemistry*. 1993; 268:11779–11784. [PubMed: 8505305]

Fig. 1. Mapping interactions within *E. coli* protein complexes by Y2H

(A) The *Escherichia coli* DNA polymerase III complex as determined by AP/MS (complex 42 in [7]). (B) Yeast two-hybrid matrix screening of the DNA polymerase III complex with N and C-terminal fusion baits with N-terminal fusion prey clones. Growth of yeast indicates positive PPI; auto-activators (false positive PPI) are marked as red. (C) The topology of the protein complex based on protein-protein interactions obtained in (B) as well as known direct interactions and structural information collected from the literature [52–58] and PDB. (D) Protein complex 34 from [7]. No published information is available on the topology of this complex. As in panels (A–C) Y2H protein-protein interaction data was used to map the topology of this complex.

Fig. 2. The ribonucleotide reductase (RNR) complex of Varicella Zoster virus (VZV)
(A) Model of the VZV RNR complex based on crystal structures of *Salmonella typhimurium* RNR (2BQ1 [32]). VZV ORF18 = R2 (Uniprot P09247), ORF19 = R1 (P09248) with chains in each homo-dimer numbered as I and II. R1 chain II and R2 chain II are colored in rainbow from N-terminal (blue) to C-terminal (red), with terminals labeled as NT and CT, respectively. The C-terminal of the R2 (II) chain was replaced with the corresponding peptide taken from 2BQ1 (thick red tube on the picture) where it is in contact with the groove in the R1 (I). The remaining part of R2 (II) CT is hidden because it is disordered in 2BQ1. The C-terminal of R1 (I) and the N-terminal of R2 (II) are shown as thick tubes in regions disordered in 2BQ1. The remaining terminals are behind the structures of the symmetric homo-dimers in this view and thus not labeled. **(B)** Interactions among RNR subunits detected in Y2H assays. The Gal4 DNA-binding (DBD) and activation domains (AD) are shown as green boxes that are either fused to the N or C-terminus of the RNR proteins. ORF18 was tested as full-length protein and as N- or C-terminal fragments (and each fragment was tested as N- and C-terminal fusion to DBD and AD). Y2H data in (B) from [16].

Figure 3. The proteasome

(A) Structure of the 26S proteasome from *Schizosaccharomyces pombe* as determined by cryo-EM density map (adapted from [41]). The core particle is shown in red, the AAA-ATPase hexamer in blue and the Rpn subunits in gold. (B) Interactions among proteasome subunits as determined by Y2H and cross-linking assays (“X-link”). Four interactions between 19S proteins and beta subunit proteins are omitted for clarity ($\alpha 3$ -Rpt2, $\alpha 4$ -Rpt2, $\alpha 6$ -Rpt5, $\alpha 7$ -Rpt5; see Supplementary Table 1). Y2H results are derived from 3 independent studies on the proteasomes of three different species (Sc = yeast [34], Ce = *C. elegans* [36], Hs = human [37]), crosslinking has been carried out in two species (Sc = *Saccharomyces cerevisiae*, Sp = *Schizosaccharomyces pombe*) [38, 41]. Structural and modeling results are derived from cryo-EM mapping, X-ray crystallography, and molecular modeling and used as “gold-standard” interactions, shown as grey bars [41, 60]. (C) Number of interactions in Y2H and cross-linking studies and the fraction that is seen in a structural model (red) that serves as “gold standard”. Numbers in red areas are the fraction of PPIs found in the “gold-standard” structure. (A) reproduced by permission from PNAS.

A

Assembly factors, head: A, FI, Nu1, Nu3, groELS
 Assembly factors, tail: I, L?, M?, K, G, GT, Z

B

C

Figure 4. Bacteriophage lambda

(A) Phage particle and its protein components. (B) Interaction map of the proteins shown in (A). Proteins inside box are virion proteins, proteins outside the box are assembly factors that are not part of the assembled virion. It remains unclear if gpM (“M”) and gpL (“L”) are part of the virion. (C) Contribution of each vector pair to the known lambda interactome [43]. The left panel shows the number of interactions generated by each vector pair, including previously known interactions (shown in blue). The right panel illustrates the fraction of previously known interactions relative to the total number of PPIs generated by each vector. The pDEST vectors produce a high fraction of such “gold-standard” PPIs although their total number of known PPIs is still smaller than those generated by the

pGBKT7/pGADC pair. Note that these numbers include 4 PPIs among regulatory proteins not in the virion. (A) and (B) after [43].

\$watermark-text

\$watermark-text

\$watermark-text

Table 1**DNA Polymerase III complex**

[7]. See Fig. 1 for interactions.

Name	Locus	Description
AsnB	JW0660	asparagine synthetase B
DnaE	JW0179	DNA polymerase III alpha subunit
DnaX	JW0459	DNA polymerase III/DNA elongation factor III, tau and gamma subunits
HolC	JW4216	DNA polymerase III, chi subunit
HolD	JW4334	DNA polymerase III, psi subunit
HolE	JW1831	DNA polymerase III, theta subunit
TrmD	JW2588	tRNA (guanine-1-)-methyltransferase

Table 2
Protein-protein interactions of the *Escherichia coli* DNA polymerase III protein complex

Interactions between protein_A and protein_B have been curated from the literature (“References”) or are the results of our own Y2H studies (vectors used are given after “Y2H:”).

protein_A	protein_B	Interaction detection method	Reference
DnaE	AsnB	tandem affinity purification	[5]
DnaX	DnaE	biochemical	[61]
HolC	DnaE	tandem affinity purification	[5]
HolD	DnaE	tandem affinity purification	[5]
HolC	DnaX	molecular sieving and biochemical	[62, 63]
HolD	DnaX	molecular sieving	[62], [64]
HolD	HolC	x-ray crystallography, molecular sieving and biochemical	[54, 62–64]
DnaX	DnaX	Y2H: GBG-GAD	This study
HolD	DnaX	Y2H: GBG-GAD	This study
HolD	HolC	Y2H: GBG-GAD	This study
HolC	HolD	Y2H: GBG-GAD and GBKC-GAD	This study
TrmD	TrmD	Y2H: GBG-GAD and GBKC-GAD	This study
AsnB	AsnB	Y2H: GBKC-GAD	This study
DnaE	DnaX	Y2H: GBKC-GAD	This study
HolC	DnaX	Y2H: GBKC-GAD	This study
HolC	HolC	Y2H: GBKC-GAD	This study

Table 3**Protein complex 32**

[7]. See Figure 1 for interactions.

Name	Locus	Description
HycG	JW2689	hydrogenase 3 and formate hydrogenase complex, HycG subunit
MetN	JW0195	DL-methionine transporter subunit, (ATP)-binding cassette (ABC) family
MntR	JW0801	Regulator of manganese levels, transcriptional regulator of mntH [23]
PerR	JW0244	predicted DNA-binding transcriptional regulator
WcaC	JW2042	predicted glycosyl transferase, colanic acid synthesis [24]
YhjB	JW3488	two-component response regulator, predicted DNA-binding
YjhC	JW5769	predicted oxidoreductase, maybe involved in sialic acid metabolism [25]
YraM	JW3116	conserved hypothetical protein