Mapping the Disease Protein Interactome: Toward a Molecular Medicine GPS to Accelerate Drug and Biomarker Discovery

Benoit Coulombe*

Institut de recherches cliniques de Montréal (IRCM), Montréal, Québec, Canada H2W 1R7

Abstract

Genomic approaches such as genome-wide association studies (GWAS), disease genome sequencing projects, and genome-wide expression profiling analyses, in conjunction with classical genetic approaches, can identify human genes that are altered in disease, thereby suggesting a role for the encoded protein (or RNA) in the establishment and/or progression of the disease. However, many technical difficulties challenge our ability to validate the role of these disease-associated genes and gene products. Moreover, many identified genes contain open reading frames (ORFs) that have yet to be annotated, that is, the function (or activity) of the encoded protein is unknown. As a result, translating the genomic information available in public databases into useful tools for understanding and curing disease is a very slow and inefficient process. To overcome these difficulties, we have developed a technology platform, termed the “molecular medicine GPS” (mm-GPS), which is aimed at defining high-quality maps of interaction networks involving disease proteins. These maps are used to identify network dysfunctions in disease cells or models and to develop molecular tools such as RNA interference (RNAi) and small-molecule inhibitors to further characterize the molecular basis of disease. In this article, I review our progress in producing high-quality maps of human protein interaction networks, and I describe how we used this information to identify new factors and pathways that regulate the RNA polymerase II transcription machinery. I also describe how we utilize the mm-GPS platform to guide more efficient efforts leading from disease-associated genes to protein interaction networks to small-molecule inhibitors, and consequently, to accelerate drug and biomarker discovery.

Keywords

Protein interaction networks; disease-associated genes; RNA polymerase II; transcription factors; biomarker and drug discovery; technology platform

Interaction Networks As a New, Systems-Based “Descriptor” of Proteins

Defining the entire network of interaction partners (i.e., interactors) of a protein provides various types of information on this protein. For example, some of the protein’s interactors are likely to be other proteins involved in the same pathway, process, or function. Consequently, and through implication by association, the function and cellular localization

*To whom correspondence should be addressed. Dr. Benoit Coulombe, Gene Transcription and Proteomics Laboratory, Institut de recherches cliniques de Montréal (IRCM), 110 avenue des Pins Ouest, Montréal, Québec, Canada H2W 1R7. Tel.: (514) 987-5662. Fax: (514) 987-5663. benoit.coulombe@ircm.qc.ca.

of the previously characterized interactors of a protein will inform on its function and localization within the cell. In addition, the network of interactors of a protein will often include regulatory factors that participate in modulating its activity. The value of this information, which must be validated using appropriate functional and biochemical assays, will increase with the quality of the interaction map.

A number of experimental methods (see Figure 1 for a comparative description) have been developed to identify and characterize protein–protein interactions, such as coimmunoprecipitation (Co-IP) experiments,²¹ the BRET²⁴ and FRET³⁴ techniques, affinity chromatography,³⁵ and phage display;³¹ however, the most popular methods in recent years have been the yeast two-hybrid (Y2H) system⁷ and protein affinity purification coupled with mass spectrometry (AP-MS) (see¹¹ for a review and¹⁵ for an example). Co-IP, BRET, FRET, affinity chromatography, and phage display approaches have mainly been used to confirm direct, pairwise interactions between already known partners. The Y2H technology is currently the most standardized technique in identifying and mapping protein–protein interactions.⁷ Although the Y2H is a method that is known to be prone to generate high rates of false positives,⁷ it has been recently reported that the quality of the high-throughput yeast two-hybrid data sets can be substantially improved when measured against a set of high-confidence physical binary interactions.³⁷ This set would include direct physical interactions within well-established complexes as well as conditional interactions such as those that are dependent upon posttranslational modifications.³⁷ The AP-MS technology has also significantly advanced the understanding of protein complexes and their composition.^{11,15} The AP-MS method allows for the purification of protein complexes under native, near physiological conditions.^{28,30}

Novel approaches, such as LUMinescence-based Mammalian IntERactome mapping (LUMIER),¹ protein-fragment complementation assay (PCA),²⁰ and high throughput imaging of protein localization²⁹ have also been developed to help map protein–protein interactions in space and time in mammalian cells (see Figure 1 for a comparative description). Although these approaches have not been widely used to date, they will most certainly serve to enhance the confidence of protein–protein interactions by helping to describe the local topology of protein interaction networks.⁷

In a recent workshop bringing together experts in the field of protein interaction mapping, a consensus emerged that the generation of an accurate, complete generic map of the human interactome will require the combination and standardization of various approaches developed in different laboratories and the establishment of rigorous quality control parameters (International Interactome Initiative (i3) Steering Workshop, September 25, 2009, Toronto, Canada).

High-Resolution Maps of the Protein Interactome for the RNA Polymerase II Transcription Machinery

For many years, our laboratory has been involved in studying the regulatory mechanisms of transcription by RNA polymerase II (RNAPII) in mammals. To identify novel regulatory factors of the RNAPII transcription machinery, we have adapted the classical AP-MS

procedure to systematically analyze the protein interaction network for various components of this machinery in the human cellular soluble fraction^{3,15} (see Figure 2 for a schematic representation of our proteomics pipeline). Our proteomic procedure couples (i) affinity purification of Tandem Affinity Peptide (TAP)-tagged proteins (the baits)^{9,17,26,28} from the cellular soluble fraction using gentle conditions in order to preserve weak, transient interactions; (ii) identification of copurified proteins (the preys) using sensitive high-accuracy mass spectrometry;^{22,36} (iii) validation of protein–protein interactions using a computational algorithm trained through machine learning to minimize the rate of both false positives and false negatives;^{15,19} and (iv) schematic representation of protein–protein interactions and networks thus generated to visualize protein connectivity.^{13,14} A key aspect of our AP-MS procedure consists in the reciprocal tagging of the interaction partners identified in our experiments. This step is important for confirming specific interactions and expanding the data set.

In our proteomics procedure, the accuracy of the resulting mapped network increases proportionally with the number of baits used in the AP-MS experiments. As compared to the data set we published in 2007,¹⁵ which was built using 32 baits, our most recently published data set used 77 baits.³ Coupled with other technical improvements of the procedure, especially regarding MS accuracy and sensitivity, our 77-bait data set (2009) permits significantly higher resolution mapping of the human RNAPII transcription network.

Our 32-bait data set (2007) allowed for the identification of two novel factors—MEPCE¹⁵ and LARP7¹⁸—that regulate the activity of the positive transcription elongation factor, P-TEFb. P-TEFb is recruited to the RNAPII elongation complex, where it functions by phosphorylating the C-Terminal Domain (CTD) of the RNAPII largest subunit RPB1 and negative elongation factors (such as NELF and DSIF) to stimulate transcriptional elongation.^{23,25,38} The two newly identified factors favor the sequestration of P-TEFb away from chromatin DNA. Indeed, the methylphosphate capping enzyme MEPCE and the RNA-binding protein LARP7 associate with and stabilize the 7SK snRNA, which, in association with inhibitory proteins, termed HEXIMs, binds to P-TEFb and prevents its recruitment to transcribing RNAPII complexes^{15,18} (see Figure 3, lower part, for a schematic representation). The formation of transcriptionally active P-TEFb requires its dissociation from the HEXIM-7SK inhibitory complex. In sum, the discovery of MEPCE and LARP7 provided the first example that our proteomics procedure is well suited to identify novel factors that regulate multisubunit transcription factors in human cells.

The 32-bait data set (2007) also identified a set of proteins that are tightly connected to RNAPII.¹⁵ Accordingly, these proteins were named RNAPII-Associated Proteins (RPAPs, namely, RPAP1, 2, 3, and 4).^{3,8,15,16} The 77-bait data set (2009) revealed that RPAP3 is part of an 11-subunit protein complex containing POLR2E (RPB5), a subunit shared by all three nuclear RNAPs, and a set of factors previously characterized as being involved in protein complex assembly (according to their Gene Ontology [GO] term).³ These include the chaperone cofactors RPAP3 and PIH1D1; the classical prefoldins PFDN2 and PFDN6; and the prefoldin-like proteins UXT, PDRG1, and C19orf2/URI.³ Interestingly, C19orf2/URI is a well-characterized protein known to bind POLR2E and shown to have a conserved role in the mammalian Target of Rapamycin (mTOR) signaling pathway.¹² Our functional analysis

of the RPAPs revealed that RPAP4 acts in concert with microtubule assembly to regulate nuclear import of the two largest RNAPII subunits RPB1 and RPB2 during biogenesis of this 12-subunit enzyme⁸ (see Figure 3, upper part, for a schematic representation). The mechanism of action of the RPAPs is currently being further investigated.

Mapping the Interactome of Disease-Associated Protein: A Key Step in the Development of a Molecular Medicine GPS to Accelerate Drug and Biomarker Discovery

The protein interaction networks of P-TFb and RNAPII demonstrate the efficiency of our AP-MS procedure to identify novel interactors of previously characterized (annotated) proteins and inferring function with previously uncharacterized proteins (in this case, MEPCE, LARP7, and the RPAPs). Many large-scale initiatives, including genome-wide association studies (GWAS), disease genome sequencing projects, and genome-wide expression profiling analyses, identify a rapidly increasing number of human genes that are systematically altered in certain diseases, suggesting that they are putatively involved in the establishment and/or the progression of the disease. However, significant technical difficulties must be overcome in order to validate the role of these disease-associated genes, a task that is further complicated by the fact that many identified ORFs are not annotated and that the function or activity of the encoded protein is unknown. For this reason, translating the information found in public genomic databases into useful tools for understanding and curing disease is a very slow and inefficient process.

To address this problem, we have developed a technology platform, termed the “molecular medicine GPS” (mm-GPS), which is aimed at improving the use of genomic information and accelerating the development of new treatments to cure disease (see Figure 4 for a schematic representation).

In this procedure, we first identify disease-associated ORFs by mining various public databases (e.g., OMIM, Online Mendelian Inheritance in Man; NCBI) that provide gene/protein annotations relating to disease conditions. The candidates are then classified according to their Gene Ontology (GO) terms and the association with disease is confirmed by performing a thorough examination of the available literature. This examination of the relevant literature can result in the removal (unconfirmed, ambiguous or disputed association with disease) or the addition (identification of additional disease-associated ORFs as a result of our literature survey) of candidates. ORFs on this list are prioritized by a multidisciplinary panel of experts, including basic scientists and clinicians, using criteria such as the occurrence of the disease in the population, the availability of relevant information in the literature, and the amenability of the system to biomarker and drug discovery. In this respect, priority is given to diseases for which collections of biological samples are easily available through collaborations with clinicians who have developed well characterized patient cohorts.

In a second step, the selected ORFs are used in AP-MS experiments, as described above, to map their interaction networks. For this purpose, we like combining the use of HEK293

cells, for which we and others have already generated significant, although far from complete, protein–protein interaction data, with cell lines specifically originating from the disease/tissue of interest. Mapping interaction networks generally identifies both upstream (i.e., regulators) and downstream (i.e., regulated proteins) effectors of the tagged proteins, as we have shown for components of the transcription machinery^{3–5,8,15,16,18} and as others have shown for proteins specifically involved in disease conditions.^{2,6,10,12,27,32,33}

Consequently, given that a protein associated with a specific disease condition represents a putative biomarker for this particular disease, its interactors putatively represent additional biomarkers that can regulate the disease phenotype by acting as upstream or downstream effectors. We then proceed to develop specific assays for monitoring changes in the expression (immunoblotting when antibodies are available or, alternatively, MS-based Multiple Reaction Monitoring (MRM)) and/or the intracellular localization (immunohistochemistry when antibodies are available) of the identified proteins in biological samples obtained from patients. Examples from our current work include the use of biopsies for muscular diseases and blood samples for metabolic diseases such as diabetes. These experiments characterize specific modulations of key components of the mapped network in cells and tissues of disease origin. One goal of such an approach is to identify sets of modulated effectors for which we can eventually develop specific and sensitive assays amenable to commercialization and use in the Clinic. Generally speaking, these biomarkers can be used as tools for (i) early prognostic of disease, (ii) monitoring the success of treatment, (iii) the stratification of patients in order to administrate personalized treatments according to specific conditions, and (iv) monitoring the efficacy of new drugs and treatments.

Third, the information obtained by mapping disease-associated protein interaction networks is used to develop specific molecular tools that we utilize in functional assays to further characterize network components in normal conditions and in conditions associated with disease. Inhibitors, such as RNAi and known pharmacological inhibitors that specifically target components of networks involved in disease, represent invaluable tools for illuminating their roles and functions. This knowledge is essential for the identification of targets for the development of novel small-molecule inhibitors amenable to drug discovery.

In sum, we expect the information obtained using this mm-GPS will improve our understanding of the molecular bases of several diseases, will help identify biomarkers to be used as diagnostic tools, and will provide knowledge and tools for accelerating the drug discovery process for some diseases. Overall, the mm-GPS is intended to guide biomedical scientists in their effort to map the disease protein interactome and translate this knowledge to cure disease.

Acknowledgments

I thank members of my laboratory for their contribution to the experimental work described in this article and for helpful discussions. My research is supported by grants from the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Cancer Research Society (CRS), the Fonds de la recherche en santé Québec (FRSQ) and the Canada Foundation for Innovation (CFI).

References

- Barrios-Rodiles M, Brown KR, Ozdamar B, Bose R, Liu Z, Donovan RS, Shinjo F, Liu Y, Dembowy J, Taylor IW, Luga V, Przulj N, Robinson M, Suzuki H, Hayashizaki Y, Jurisica I, Wrana JL. High-throughput mapping of a dynamic signaling network in mammalian cells. *Science*. 2005; 307:1621–1625. [PubMed: 15761153]
- Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S, Hopf C, Huhse B, Mangano R, Michon AM, Schirle M, Schlegl J, Schwab M, Stein MA, Bauer A, Casari G, Drewes G, Gavin AC, Jackson DB, Joberty G, Neubauer G, Rick J, Kuster B, Superti-Furga G. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol*. 2004; 6:97–105. [PubMed: 14743216]
- Cloutier P, Al Khoury R, Lavalley-Adam M, Faubert D, Jiang H, Poitras C, Bouchard A, Forget D, Blanchette M, Coulombe B. High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods*. 2009; 48:381–386. [PubMed: 19450687]
- Cloutier P, Coulombe B. New insights into the biogenesis of nuclear RNA polymerases. *Biochem Cell Biol*. 2010; 88:211–221. [PubMed: 20453924]
- Cote-Martin A, Moody C, Fradet-Turcotte A, D'Abramo CM, Lehoux M, Joubert S, Poirier GG, Coulombe B, Laimins LA, Archambault J. Human papillomavirus E1 helicase interacts with the WD repeat protein p80 to promote maintenance of the viral genome in keratinocytes. *J Virol*. 2008; 82:1271–1283. [PubMed: 18032488]
- Djouder N, Metzler SC, Schmidt A, Wirbelauer C, Gstaiger M, Aebersold R, Hess D, Krek W. S6K1-mediated disassembly of mitochondrial URI/PP1gamma complexes activates a negative feedback program that counters S6K1 survival signaling. *Mol Cell*. 2007; 28:28–40. [PubMed: 17936702]
- Figeys D. Mapping the human protein interactome. *Cell Res*. 2008; 18:716–724. [PubMed: 18574500]
- Forget D, Lacombe A-A, Cloutier P, Al-Khoury R, Bouchard A, Lavalley-Adam M, Faubert D, Jeronimo C, Blanchette M, Coulombe B. The protein interaction network of the human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Mol Cell Proteomics*. 2010 epub ahead of print.
- Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. 2002; 415:141–147. [PubMed: 11805826]
- Giglia-Mari G, Coin F, Ranish JA, Hoogstraten D, Theil A, Wijgers N, Jaspers NG, Raams A, Argentini M, van der Spek PJ, Botta E, Stefanini M, Egly JM, Aebersold R, Hoeijmakers JH, Vermeulen W. A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet*. 2004; 36:714–719. [PubMed: 15220921]
- Gingras AC, Gstaiger M, Raught B, Aebersold R. Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol*. 2007; 8:645–654. [PubMed: 17593931]
- Gstaiger M, Luke B, Hess D, Oakeley EJ, Wirbelauer C, Blondel M, Vigneron M, Peter M, Krek W. Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science*. 2003; 302:1208–1212. [PubMed: 14615539]
- Hu Z, Mellor J, Wu J, Delisi C. VisANT: an online visualization and analysis tool for biological interaction data. *BMC Bioinform*. 2004; 5:17.
- Hu Z, Mellor J, Wu J, Yamada T, Holloway D, Delisi C. VisANT: data-integrating visual framework for biological networks and modules. *Nucleic Acids Res*. 2005; 33:W352–W357. [PubMed: 15980487]
- Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B.

Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell*. 2007; 27:262–274. [PubMed: 17643375]

16. Jeronimo C, Langelier MF, Zeghouf M, Cojocaru M, Bergeron D, Baali D, Forget D, Mnaimneh S, Davierwala AP, Pootoolal J, Chandy M, Canadien V, Beattie BK, Richards DP, Workman JL, Hughes TR, Greenblatt J, Coulombe B. RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Mol Cell Biol*. 2004; 24:7043–7058. [PubMed: 15282305]
17. 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. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*. 2006; 440:637–643. [PubMed: 16554755]
18. Krueger BJ, Jeronimo C, Roy BB, Bouchard A, Barrandon C, Byers SA, Searcey CE, Cooper JJ, Bensaude O, Cohen EA, Coulombe B, Price DH. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res*. 2008; 36:2219–2229. [PubMed: 18281698]
19. Lavalley-Adam M, Coulombe B, Blanchette M. Detection of locally over-represented GO terms in protein-protein interaction networks. *J Comput Biol*. 2010; 17:443–457. [PubMed: 20377456]
20. Michnick SW, Ear PH, Manderson EN, Remy I, Stefan E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat Rev Drug Discovery*. 2007; 6:569–582. [PubMed: 17599086]
21. Miernyk JA, Thelen JJ. Biochemical approaches for discovering protein-protein interactions. *Plant J*. 2008; 53:597–609. [PubMed: 18269571]
22. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics*. 2005; 4:2010–2021. [PubMed: 16249172]
23. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell*. 2006; 23:297–305. [PubMed: 16885020]
24. Pfleger KD, Eidne KA. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat Methods*. 2006; 3:165–174. [PubMed: 16489332]
25. Price DH. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol*. 2000; 20:2629–2634. [PubMed: 10733565]
26. Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*. 2001; 24:218–229. [PubMed: 11403571]
27. Ranish JA, Hahn S, Lu Y, Yi EC, Li XJ, Eng J, Aebersold R. Identification of TFB5, a new component of general transcription and DNA repair factor IIIH. *Nat Genet*. 2004; 36:707–713. [PubMed: 15220919]
28. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*. 1999; 17:1030–1032. [PubMed: 10504710]
29. Rosochacki SJ, Matejczyk M. Green fluorescent protein as a molecular marker in microbiology. *Acta Microbiol Pol*. 2002; 51:205–216. [PubMed: 12588095]
30. Ruffner H, Bauer A, Bouwmeester T. Human protein-protein interaction networks and the value for drug discovery. *Drug Discovery Today*. 2007; 12:709–716. [PubMed: 17826683]
31. Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. 1985; 228:1315–1317. [PubMed: 4001944]

32. Taylor IW, Linding R, Warde-Farley D, Liu Y, Pesquita C, Faria D, Bull S, Pawson T, Morris Q, Wrana JL. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol.* 2009; 27(2):199–204. [PubMed: 19182785]
33. Tian Q, Feetham MC, Tao WA, He XC, Li L, Aebersold R, Hood L. Proteomic analysis identifies that 14-3-3{zeta} interacts with {beta}-catenin and facilitates its activation by Akt. *Proc Natl Acad Sci US A.* 2004; 101:15370–15375.
34. Tramier M, Gautier I, Piolot T, Ravalet S, Kemnitz K, Coppey J, Durieux C, Mignotte V, Coppey-Moisan M. Picosecond-hetero-FRET microscopy to probe protein-protein interactions in live cells. *Biophys J.* 2002; 83:3570–3577. [PubMed: 12496124]
35. Wilchek M, Chaiken I. An overview of affinity chromatography. *Methods Mol Biol.* 2000; 147:1–6. [PubMed: 10857080]
36. Yates JR, Cociorva D, Liao L, Zabrouskov V. Performance of a linear ion trap-Orbitrap hybrid for peptide analysis. *Anal Chem.* 2006; 78:493–500. [PubMed: 16408932]
37. 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. High-quality binary protein interaction map of the yeast interactome network. *Science.* 2008; 322:104–110. [PubMed: 18719252]
38. Zhou Q, Yik JH. The Yin and Yang of P-TEFb Regulation: Implications for Human Immunodeficiency Virus Gene Expression and Global Control of Cell Growth and Differentiation. *Microbiol Mol Biol Rev.* 2006; 70:646–659. [PubMed: 16959964]

METHOD	DESCRIPTION	REF
Bioluminescence Resonance Energy Transfer (BRET)	Takes advantage of the resonance energy transfer between a luminescent donor and a fluorescent acceptor	24
Fluorescence Resonance Energy Transfer (FRET)	Takes advantage of resonance energy transfer between a fluorescent donor and a fluorescent acceptor, provided that the two fluorophores are different	34
Co-ImmunoPrecipitation (Co-IP)	Uses the specificity of antibodies to purify a target protein with its interaction partners	21
Yeast-Two Hybrid (Y2H)	Relies on a transcriptional readout for the detection of protein-protein interactions through the reconstitution of a functional transcriptional activator	7
Affinity Purification coupled with Mass Spectrometry (AP-MS)	Allows for the purification of protein complexes under native conditions and the identification of direct or indirect interacting partners without prior knowledge of complex composition	11, 15
LUMinescence-based Mammalian IntERactome mapping (LUMIER)	An automated high-throughput technology to map protein-protein interaction networks systematically in mammalian cells. Use of <i>Renilla</i> luciferase enzyme fused to proteins of interest, which are then coexpressed with individual Flag-tagged partners in mammalian cells	1
Protein-fragment Complementation Assay (PCA)	Comprises a re-assembly of separate fragments of a protein reporter molecule fused to the two proteins of interest. The reporter will reconstitute its activity only when the two fused protein are in proximity of each other	20
Affinity Chromatography	Based on the highly specific interaction of a protein of interest immobilized on a solid support with a protein that specifically binds to it while passing through the column	35

Figure 1.
Table comparing the main methods used for the analysis of protein-protein interactions, protein complexes, and networks.

Figure 2.

Schematic representation of our affinity purification coupled to mass spectrometry (AP-MS) procedure. Cells expressing a protein of interest (bait) having a Tandem Affinity Purification (TAP) tag are used to identify protein interaction partners (i.e., interactors or preys) for this protein. The proteins that copurify with the TAP-tagged protein are identified by mass spectrometry and validated using a computational procedure aimed at minimizing false positive and false negative rates. Reciprocal tagging of affinity purified partners serves to confirm some interactions and to expand the data set.

Figure 3.

Identification of novel interaction partners of the RNA polymerase II machinery. The upper part illustrates the function of the RNA Polymerase II-Associated Proteins (RPAPs) in regulating the nuclear import and biogenesis of the enzyme RNA polymerase II (RNAPII). The detailed role of the RPAPs in this process is currently under investigation. The lower part illustrates the function of the methylphosphate capping enzyme MEPCE and the 7SK-binding protein LARP7 in regulating the activity of the Positive Transcription Elongation Factor b (P-TEFb). Both MEPCE and LARP7 stabilize the 7SK RNA, thereby maintaining P-TEFb in an inactive state. Activation of P-TEFb requires release of the inhibitory factors HEXIM, 7SK, MEPCE, and LARP7, and association with BRD4.

Figure 4.

Chart summarizing our Molecular Medicine GPS (mm-GPS) platform. This technology platform is aimed at guiding research from disease-associated Genes to Protein interaction networks to Small-molecule inhibitors. The various components of the mm-GPS are illustrated.