Université de Montréal

A proteome-wide strategy reveals a novel mechanism of control of cell cycle progression through modulation of cyclin mRNA stability

par

Vincent Messier

Département de biochimie

Programme de Biochimie

Faculté de médecine

Université de Montréal

Thèse présentée à la faculté des études supérieures

en vue de l'obtention du grade *Philosophiae Doctor (Ph.D)*

en biochimie

8 February 2013

Université de Montréal

Faculté des études supérieures

Cette thèse s'intitule:

A proteome-wide strategy reveals a novel mechanism of control of cell cycle progression through modulation of cyclin mRNA stability

Présentée par:

Vincent Messier

a été évaluée par un juré composé des personnes suivantes:

Pascal Chartrand	Président-rapporteur
Stephen Michnick	Directeur des recherches
Damien D'Amours	Membre du juré
Jan Skotheim	Examinateur externe
Martin Albert	Représentant du doyen de la FES

Sommaire

La quantité de données générée dans le cadre d'étude à grande échelle du réseau d'interaction protéine-protéine dépasse notre capacité à les analyser et à comprendre leur sens; d'une part, par leur complexité et leur volume, et d'un autre part, par la qualité du jeu de donnée produit qui semble bondé de faux positifs et de faux négatifs. Cette dissertation décrit une nouvelle méthode de criblage des interactions physique entre protéines à haut débit chez *Saccharomyces cerevisiae*, la complémentation de fragments protéiques (PCA). Cette approche est accomplie dans des cellules intactes dans les conditions natives des protéines; sous leur promoteur endogène et dans le respect des contextes de modifications post-traductionnelles et de localisations subcellulaires. Une application biologique de cette méthode a permis de démontrer la capacité de ce système rapporteur à répondre aux questions d'adaptation cellulaire à des stress, comme la famine en nutriments et un traitement à une drogue.

Dans le premier chapitre de cette dissertation, nous avons présenté un criblage des paires d'interactions entre les protéines résultant des quelques 6000 cadres de lecture de *Saccharomyces cerevisiae*. Nous avons identifié 2770 interactions entre 1124 protéines. Nous avons estimé la qualité de notre criblage en le comparant à d'autres banques d'interaction. Nous avons réalisé que la majorité de nos interactions sont nouvelles, alors que le chevauchement avec les données des autres méthodes est large. Nous avons pris cette opportunité pour caractériser les facteurs déterminants dans la détection d'une interaction

par PCA. Nous avons remarqué que notre approche est sous une contrainte stérique provenant de la nécessité des fragments rapporteurs à pouvoir se rejoindre dans l'espace cellulaire afin de récupérer l'activité observable de la sonde d'interaction. L'intégration de nos résultats aux connaissances des dynamiques de régulations génétiques et des modifications protéiques nous dirigera vers une meilleure compréhension des processus cellulaires complexes orchestrés aux niveaux moléculaires et structuraux dans les cellules vivantes.

Nous avons appliqué notre méthode aux réarrangements dynamiques opérant durant l'adaptation de la cellule à des stress, comme la famine en nutriments et le traitement à une drogue. Cette investigation fait le détail de notre second chapitre. Nous avons déterminé de cette manière que l'équilibre entre les formes phosphorylées et déphosphorylées de l'arginine méthyltransférase de *Saccharomyces cerevisiae*, Hmt1, régulait du même coup sont assemblage en hexamère et son activité enzymatique. L'activité d'Hmt1 a directement un impact dans la progression du cycle cellulaire durant un stress, stabilisant les transcrits de *CLB2* et permettant la synthèse de Cln3p. Nous avons utilisé notre criblage afin de déterminer les régulateurs de la phosphorylation d'Hmt1 dans un contexte de traitement à la rapamycin, un inhibiteur de la kinase cible de la rapamycin (TOR). Nous avons identifié la sous-unité catalytique de la phosphatase PP2a, Pph22, activé par l'inhibition de la kinase TOR et la kinase Dbf2, activé durant l'entrée en mitose de la cellule, comme la phosphatase et la kinase responsable de la modification d'Hmt1 et de ses fonctions de régulations dans le cycle cellulaire. Cette approche peut être

généralisée afin d'identifier et de lier mécanistiquement les gènes, incluant ceux n'ayant aucune fonction connue, à tout processus cellulaire, comme les mécanismes régulant l'ARNm.

Abstract

The quantity of data generated within the framework of protein-protein interaction network large-scale studies exceeds our capacity to analyze them and to understand their meaning; on one hand, by their complexity and their number, and on the other hand, by the quality of the produced data, which are populated with spurious interactions. This dissertation describes new applications of a protein-fragments complementation assay (PCA) to screen for interactions among all proteins in the budding yeast *Saccharomyces cerevisiae*. This approach is carried out in intact cells, with proteins expressed in their native contexts and under their endogenous promoter, thus assuring correct post-translational modifications and subcellular localization. A further novel application of PCA is described for investigating proteome wide changes in response to cellular adaptation to stresses, such as nutrient starvations and drug treatments. Finally, as a result of the latter strategy applied to characterizing proteome-wide response to the immunosuppressant drug, rapamycin, I describe the discovery of an unforeseen mechanism of modulating cell cycle progression through control of cyclin mRNA stability.

In the first chapter of this dissertation, I present a pairwise screen of interactions among proteins resulting from the ~6000 open reading frames in *Saccharomyces cerevisiae*. We identified 2770 interactions among 1124 proteins. We estimated the quality of our screen by comparing our results to curated gold standard data and coverage of known interactions to all previous

studies. The majority of our interactions were novel, but overlap with data from previous studies was as high as 40%. PCA is based on refolding of the reporter protein from complementary *N*- and *C*- terminal fragments following interaction of the two proteins to which they are fused. Thus, reporter activity is sterrically limited to interactions in which the termini of the proteins to which the complementary reporter fragments are fused are sufficiently close in space. In the case of our reporter, this limit was 8 nm. Thus PCA is a molecular ruler, providing information on both direct protein-protein interactions and sterrically restricted distances between proteins in complexes. We benchmarked and demonstrated correct topological relationships for a number of known complexes, including the proteasome, RNA polymerase II and the nuclear pore complex. Thus our study provided, for the first time, a topological map of complex organization in a living cell. The integration of the results from such efforts with those of gene regulation dynamics and protein modifications will lead to a fuller understanding of how complex cellular processes are orchestrated at a molecular and structural level in the living cell.

In chapter 2, I describe the results of an application of PCA to study the dynamic rearrangement of the proteome under a specific stress; treatment of cells with rapamycin. The results of these efforts were the identification of a novel mechanism of cell cycle control at the level of cyclin mRNA. Specifically, we discovered that the balance between the phosphorylated and dephosphorylated forms of the *Saccharomyces cerevisiae* arginine methyltransferase, Hmt1, regulates both its assembly into a hexamer and its

enzymatic activity. The Hmt1 activity modulates cell cycle progression through stabilizing the B cyclin *CLB2* mRNA. We then used PCA to identify the Hmt1 regulators under rapamycin treatment. We identified the catalytic subunit of the PP2a phosphatase, Pph22, activated by the inhibition of TOR, and the kinase Dbf2, activated during entry into mitosis, as the phosphatase and the kinase responsible for the modification of Hmt1 and for its regulatory functions in the cell cycle.

I thus, in the end close the circle I began in this summary, going from large-scale discovery of protein-protein interactions, to mapping dynamics of proteome changes during an adaptation and finally to mechanistic insight into a primordial control mechanism in cellular dynamics. The strategies that we devised to discover this mechanism can be generalized to identify and mechanistically link genes together, including those of unknown function, to any cellular process.

Table of contents

Table of figures

Abbreviations

RBD	Ras binding domain on Raf
mRNA	messenger Ribonucleic acid
DNA	Deoxyribonucleic acid
NMR	Nuclear Magnetic Resonance
S phase	Synthesis phase
M phase	Mitosis phase
G1	Gap 1
G2	Gap 2
G0	Gap 0
Cdk	Cyclin dependant kinase
APC	Anaphase Promoting Complex
Y2H	Yeast two-Hybrid
PCA	Protein-fragment Complementation Assay
PIN	Protein Interaction Network
hnRNP	heterogeneous nuclear RiboNucleoProtein
GAP	GTPase Activating Protein

DHFR	DiHydroFolate Reductase
Rluc	*Renilla* LUCiferase
PKA	Protein Kinase A
cAMP	cyclic Adenosine MonoPhosphate
PPI	Protein-Protein Interaction
ORF	Open Reading Frame
PPV	Positive Predicted Value
SC	Synthetic Completed media
YFP	Yellow Fluorescent Protein
MIPS	Munich Information Center for Protein Sequences
CC	Cell Compartment
BP	Biological Process
MF	Molecular Function
TAP-MS	Tandem Affinity Purification coupled Mass Spectrometry
ROC curve	Receiver Operating Characteristic curve
GO	Gene Ontology

ESCRT	Endosomal Sorting Complex Required for Transport
MVB	MultiVesicular Bodies
RNA pol	RNA polymerase

Amino acids codes

Amino acid	Three letter code	One letter code
Alanine	ala	A
Arginine	arg	R
Asparagine	asn	N
Aspartic acid	asp	D
Cystein	cys	C
Glutamic acid	glu	E
Glutamine	gln	Q
Glycine	gly	G
Histidine	his	H
Isoleucine	ile	I
Leucine	leu	L
Lysine	lys	K
Methionine	met	M
Phenylalanine	phe	F
Proline	pro	P
Serine	ser	S
Threonine	thr	T
Tryptophan	try	W
Tyrosine	tyr	Y
Valine	val	V

Acknowledgements

During my time in graduate school, I have been in contact with many people who have transformed my views. I thought that graduation would be a simple trip, where I would acquire important knowledge and useful techniques that could help me in my future career. The truth is more complicated and elaborate. As I think of every person that I met, had discussions with, and challenges that I encountered and, more importantly, overcame were by far more important and useful than what university claims to be.

I would like to first thank Dr. Stephen Michnick for his contribution on fostering both my professional and personal maturation. Steve was the best guide for a graduate school trip. While he did the road several times already with other past students, he chooses to do it again with me. The paths were not always easy to get through and, being in Montreal, some were probably under construction. He foresaw obstacles in the way, leaving me to choose between getting through them and finding other ways. Sometimes I stopped following his advice, perhaps thinking that he was lost even more than I was. Luckily for me, he was waiting on the side of the road so that I did something more productive and efficient to get to the end of the journey. Now, being at the end of the trip, I feel that he not only cares that his students present time in his laboratory to be productive, but also he worries even more for their future, as he leaves them with the map for a successful career.

At the beginning of my doctoral studies, I became acquainted with a young scientist in Mcgill University. She is the 'queen of the scope' and had no troubles to tell me that my first paragraphs in the science draft were just worse than a pile of 's#&t'. Dr Jackie Vogel is not only a talented researcher but also an inspiration to me. I saw a strong woman starting her lab with a sea of ideas, little help, but all the ambition of the world to succeed in it. It was probably not an easy period, but she was ready to share with me and her lab members, advice, cups of coffee, chocolate and conversations that I still value. She taught me microscopy and transmitted to me some of her passion to see these small glowing dots on a black screen. Her generosity in the time she gives to others and the confidence she shares to those who have worked with her contribute significantly to the attitude and motivation of people around, for which I am thankful to have been part of.

I described Dr Daniel Zenklusen as the 'God of FISH' several times, probably because he is. He gave me time and advices to develop FISH proficiency. More importantly, we shared our enthusiasm for the latest project, which would be a dim reflection of what it is now without his input. I admire his capacity to see details, the good and the bad ones. He shapes ideas to push them to their purest form, while optimistically looking at flaws in these diamonds as a continuous chance for improvements. I believe that the future will continue to make our path overlap in many fruitful collaborative efforts.

It is inconceivable to me that I could be writing this section without several thoughts for my comrades, present and past members of the Michnick group. I wish to thank you one by one for your support, your enthusiasm and the bonds of friendship that make the Michnick group a special boat to get in to. Po Hien and Mohan for the ride that we shared to get our PhD, the laughs and the talks that came with it. To the old timers of the group, François-Xavier, Alexis and Ingrid, who did it before us and left us with a strong foundation. Louis-Philippe, Sinan and madame Didi, the new ones, for making me feel old and see that I progress through advice that I received and give back to you guys. To Jean-François and Edi for the shared beers, the ski trips and the unique friendship. To Jacqueline for the conversation on ecolo-, grano- and shaggying up life. To Valérie, “les Nathalies”, Jean-Marc, Livia, Émilie and Émily, the past post-doc friends, who shared their time and talents with me. To Luz, Durga, Emmanuelle, Bram, Abdellali and Punam, the new post-doc crowd, who were there during these past tough months. To Stevo, Mercedes, Kirill and Yelena (the yeast group) for your assurance in me while being part of a heroic accomplishment. To Michael and Beat for the shared diners, the conversation and that we feel all at home. To Carole, Ibrahima, Emmanuel and Christian to have been influential during my time in the lab. There will always be a place for you all in my thoughts.

For the time, that I invested in my studies, is a time that I could not always be with you, my family. I wish to thank my parents and my family members. Even if you did not know what I could be doing in the lab, you were

understanding of my joy, my pain and all other feelings that could come and go in these years in school. There are no words for my love and admiration that I bear for each of you and only hope that you will always be as proud of me.

To my friends that were there during this time., especially, Karine Grimard, Amélie Forget and Francis Bélanger for their constant genuine friendship, using my ear to share their passion and their own senses to support my taught times. Your company has helped me to preserve my sanity and enjoy life.

I also thank my thesis committee for their advice. Sometimes for the kick that made me think better and work faster. They were supportive in both my academic progression and in the ideas that were discussed in the frame of proposed projects.

Finally, for the evaluation of this dissertation and the interesting discussion that will issue from the upcoming defense, I wish to acknowledge my jury members.

Dedication

"For many smiles, many tears and the times that we shared laughter, they were always the best."

"There are three things that a child can teach an adult:" To be happy for no reason; to be always busy doing something;" And to know how to demand - with all one's might - what one wants"

Paulo Coelho

Overview

Proteins are operators in all cellular processes including the cell cycle, the central subject of this thesis. Further, matter, energy and information transfer through biochemical pathways and their organization, are determined by specific protein-protein interactions at every phase of the cell cycle. For example, protein interactions during cell cycle progression such as mitosis are indispensable for the structure of sub-cellular organelles [1], the transport machinery across the various biological membranes [2], the packaging of chromatin [3], the network of microtubule filaments [4], the periodic decay of mitotically regulated molecules [5], and the cyclin-dependent kinase signal transduction [6], the regulation of DNA region by transcription factors [7], to name a few. Protein interactions during cell cycle progression have been the object of intense research for many years due to their importance in development and disease. Aberrant protein bindings are implicated in a number of neurological disorders such as Creutzfeld-Jacob, Alzheimer's disease [8] and cancer, during all developmental stages and metastasis [9]. Although disregulated protein interactions are of interest clinically, the main focus of this study is in the finely tuned recruitment of proteins and molecules involved in the regulation of cell division.

Within cells, proteins interact with other proteins, metabolites and nucleic acids. Protein interactions are measured using a variety of assays, described in method sections of chapters 1 and 2. These assays are based on

new innovations in various fields of biochemistry and biophysics, increasing the sensitivity and accuracy in molecular interaction cartography. Recently, many techniques were designed with the aim of scaling up to measure interactions on a proteome-wide level [10-14]. The false-positive and false-negative rates for networks generated from such high-throughput methods are high, stressing the need for discovering new experimental techniques and new methods that will allow to study protein interactions in intact cell, and to perform studies in various growth conditions and cellular states with endogenously expressed full-length proteins in their native post-translationally modified states and cellular locations. This knowledge could be generated to understand how genetic or environmental perturbations could be linked to these networks in disease development or to uncover how perturbations, such as drugs, of individual network components could reduce or eliminate cellular dysfunctions.

I present here the application in *Saccharomyces cerevisiae* of a protein-fragments complementation assay (PCA) reporter of protein-protein interactions, that is unique because: (I) it is performed in intact cells and (II) it can be used to address dynamic behaviors of protein complexes under non-toxic conditions. The goal of this dissertation is to demonstrate the limitations, the determining factors in detecting interactions and the strength of this approach in yeast and to show how it can be used to make a substantial and unique contribution to biology. In the first chapter of this dissertation, I focus on describing the PCA in the context of a systematic screen of all pairwise

interactions among ~6,000 yeast proteins. In the second chapter, I applied the PCA strategy to understand the mechanism regulating an arginine methyltransferase assembly and activity under stressful conditions, such as starvation and rapamycin-induced TOR kinase inhibition. These enzyme regulations have a direct impact on RNA-binding protein capacities to stabilized cyclin transcripts and overall cell cycle progression. As a result, this strategy can be generalized to identify and mechanistically link genes, including those of unknown function, to any cellular process.

1.Chapter 1: An in Vivo Map of the Yeast Protein Interactome

Kirill Tarassov,[1*] **Vincent Messier,[1*]** Christian R. Landry,[1,2*] Stevo Radinovic,[1*] Mercedes M. Serna Molina,[1] Igor Shames,[1] Yelena Malitskaya,[1] Jackie Vogel,[3] Howard Bussey,[3] Stephen W. Michnick[1,2†]. ***Science***, 2008 June 13; 320(5882): 1465-70. Epub 2008 May 8.

1.1.Overview

Communication at the level of the organism requires the transduction of physical cues or chemical information, such as turgor pressure, toxin presence or available nutrient sources, into biochemical signals. Chemical and physical signals are processed by specific interactions among heterologous proteins through non-covalent or covalent modifications. Some of these molecular transducers arose from co-evolution of complementary binding interfaces that specifically interact when simultaneously localized in space and time with corresponding molecular partners, while accumulating mutations that disfavor spurious contacts [15-17]. For example, the *Haemophilus influenza* SspB protein homodimer evolved a symmetrical binding surface consisting of an α-helix and β-strand forming a hydrophobic cluster between two copies of the molecule [18]. The SspB homomeric interface is essential to bind and target for degradation truncated peptide resulting from stalled ribosome activity upon severe starvation [19]. Computationally design SspB mutations at position 12, 15, 16 and 101 that optimized the stability between the heterodimer mutants, $SspB_{FAFI}$ and $SspB_{LALI}$, would actually assemble with high stability both into SspB homomer and heteromer in solution. In contrast, design mutations at the

same position that optimize the difference in energy between the heterodimer and both homodimer mutants, $SspB_{LSLA}$ and $SspB_{YGMF}$, show a 99% specificity for the heteromeric assembly while bearing lower stability [17]. It is reasonable to think that, as in the design of SspB mutants, protein complex specificity in intact cell is achieved at the cost of stability, rendering *in vitro* purification conditions a constraint to their detection.

Proteins interact through their interfaces into stable macromolecular structures for which a panel of high-throughput screen strategy were design to capture. The quality of an unbiased PPI screen is composed of two distinct elements: (I) its completeness and (II) its fidelity to reproduce the real PPI network. Protein complexes have been extensively studied by many small-scale techniques and have been curated to create catalogues, such as The Munich Information Center for Protein Sequences (MIPS). These curated protein complexes, such as the nuclear pore, the ribosome, the RNA polymerases, the proteosome, to name a few, can represent *bona fides* interactions between proteins that could be considered to estimate false-negative discovery rate in a given screen. The completeness of a reported screen can be evaluated by verifying the number of observed interactions within this standard. The total coverage achieved by a screen is determined by both the limitations and sensitivity of the detection method. Therefore, the minimal amount of false-negatives can be reached at the union of the data from different approaches. Equally important, each of these screens should be carried out under conditions that maximize the sensitivity of the detection technique. The fidelity of a PPI

method, on the other hand, is more difficult to determine. Unfortunately, there is no clear dataset of protein that would not physically interact. Therefore, a compromise standard set is generated from logical inference; creating imaginary "impossible" interactions by selecting pairs of proteins with divergent expression profiles or mutually exclusive compartmentalization and assuring that none of these imaginary interactions were ever reported in all published data. This approach overestimates the number of impossible interactions, but provides a means to access the false-positive discovery rate.

The fidelity of a protein-protein interaction detection technique can be increased by systematic removal of artifacts. These can be detected by applying appropriate controls to the data, usually by determining technical limitations of a technique that recover false interactions. For PCA, there are two potential technical causes of false-positive interactions. PCA is based on protein interaction-mediated refolding of the reporter fragments to yield an active enzyme [20]. It is, however, possible that in some cases spontaneous folding of reporter from the fragments can occur. An appropriate control for such cases is to probe the proteome with fragments fused to a protein that does not interact with any protein in the cell (e.g. GFP) or the fragment itself, expressed alone. In our screen, we identified approximately 300 proteins that interact with fragment alone, notably highly abundant proteins. A second possibility is that non-specific interactions could be generated due to trapping of complexes between two proteins that normally wouldn't interact, due to irreversible folding of the reporter protein. In fact, we have shown that this is not the case: PCA reporter

proteins do reversibly unfold when protein complexes dissociate for several reporter proteins, but we had not tested this for the reporter used in this study, *murine* dihydrofolate reductase (DHFR) [21, 22]. We verified the reversibility of DHFR PCA folding using the adenosine 3′,5′-monophosphate–dependent dissociation of the yeast protein kinase A complex as a test system. This heterotetrameric complex has two catalytic and two regulatory subunits. The adenosine 3′,5′-monophosphate (cAMP) binds to and promotes conformational change of the regulatory subunits that release the catalytic subunits. An irreversible reporter folding would retain the catalytic subunit under cAMP treatments. This test insured that we would not report interactions resulting of irreversible trapping by the fragments.

DHFR PCA depends on physical contact between the reporter fragments. This could be limited by the distance between the two protein termini to which the fragments are fused. This is an opportunity to study protein complex membrane spanning protein complex topologies. The structure of the RNA polymerase II is appealing as it is determined at 2.8 Å resolution and contains the largest described variety of protein subunits [23]. We could technically use the linker peptide, separating the subunit of the RNA polymerase II and our reporter fragment, as a molecular ruler that could theoretically give a resolution of the position of complex subunits at the resolution of a single amino acid. Similarly, the reporter fragments cannot refold through a membrane. Therefore, we can use the reporter fragments fused to the C termini of membrane spanning proteins to test for their close proximity

and also their orientation into the same cellular compartment [24-27]. These particularities of the PCA would provide complementary information on the nature of proteins and their complexes.

The motivation for determining protein interaction networks is to establish potential protein functions and relationships. Gene annotation is the process of conferring biological information to sequence, consisting of (I) predicting encoded gene sequence in the genome and (II) conferring biological information upon these genes. Gene function annotation is a challenge, given that proteins, in their different states, (e.g. subcellularly localized or post-translationally modified) can be quite versatile, assuming multiple functions at any given time and under different conditions [28]. The difficulty of the gene annotation task is reflected in the number of genes that remain to be annotated. In the popular eukaryotic model, *Saccharomyces cerevisiae*, almost 30% of the genes are refractory to past annotation efforts [29]. A variety of factors are likely to contribute to the relatively large number of non-annotated yeast genes, including genetic redundancy, lack of strong phenotype, and the possibility that not all genes are real open reading frames (ORF). However, there is a clear need for new approaches and strategies for specific problems, including the characterization of individual genes and their role in nature. For example, the rationale for assigning gene function based on protein interactions relies on a "guilt by association" argument, suggesting that physically interacting proteins shared common functions known *a priori* for one of the protein. This argument is justified for defined protein complexes, but the accuracy of the argument

decreased when (I) there is no *a priori* annotation for both of the interacting proteins, (II) function shared by the two interacting proteins is novel and differ from any known *a priori* relationship between functions (e.g. between a transcription factor and a metabolic enzyme) and (III) although the proteins physically interact, these proteins shared no functions. These limitations have not prevented biologist from using these information as a starting point to characterize proteins with unknown functions. For example, using the yeast two-hybrid assay, the YDR016C gene was shown to function with the spindle-pole body as part of the Dam1 complex [11]. It was revealed to be an essential component of the DASH complex required for mitotic spindle integrity [30]. Assembling proteins of unknown function into networks linked to specific cellular processes can thus provide new mechanistic insights and potentially provide new therapeutic targets for the development of drugs.

Background

In the following manuscript that was published in an article format in the journal *Science*, we systematically screened for pairwise protein-protein interactions in *Saccharomyces cerevisiae* using a protein-fragment complementation assay (PCA) based on the dihydrofolate reductase (DHFR) reporter enzyme. We identified 2770 interactions among 1124 proteins. These interactions were mostly novel on the fact that PCA is efficient to detect protein interaction within and between protein complexes and that we perform the screen on a wider array than previous experiments reported in MIPS and by TAP-MS [31]. We demonstrate that PCA can identify protein complex structure for a crystallized protein complex, in the case of the yeast RNA polymerase II, and globally, originating from all stable crystallized complexes of yeast-homologous proteins deposited in PDB. Then by comparing likelihood that protein interaction can be detected between membrane-spanning proteins and the cell compartment location of their C-termini, we demonstrate topological relationship between these proteins. Finally, using the cyclic adenosine 3′,5′-monophosphate (cAMP) dependent dissociation of the yeast protein kinase A (PKA) complex as a test system to show full reversibility of the DHFR PCA, we insured that the PCA itself does not trap spurious complexes, altering the thermodynamics of binding. The *in vivo* extended network provides insights into fundamental cellular processes, including cell polarization and autophagy, pathways that are evolutionarily conserved and central to both development and human health.

An *in Vivo* Map of the Yeast Protein Interactome

Kirill Tarassov,[1]* **Vincent Messier,[1]*** Christian R. Landry,[1,2]* Stevo Radinovic,[1]* Mercedes M. Serna Molina,[1] Igor Shames,[1] Yelena Malitskaya,[1] Jackie Vogel,[3] Howard Bussey,[3] Stephen W. Michnick[1,2]†. ***Science***, 2008 June 13; 320(5882): 1465-70. Epub 2008 May 8.

[1]Département de Biochimie, Université de Montréal.

[2]Centre Robert-Cedergren, Bio-Informatique et Génomique, Université de Montréal.

[3]Department of Biology, McGill University.

*These authors contributed equally to this work.

†To whom correspondence should be addressed.

1.2. Abstract

Protein interactions regulate the systems-level behavior of cells; thus, deciphering the structure and dynamics of protein interaction networks in their cellular context is a central goal in biology. We have performed a genome-wide *in vivo* screen for protein-protein interactions in *Saccharomyces cerevisiae* by means of a protein-fragment complementation assay (PCA). We identified 2770 interactions among 1124 endogenously expressed proteins. Comparison with previous studies confirmed known interactions, but most were not known, revealing a previously unexplored subspace of the yeast protein interactome. The PCA detected structural and topological relationships between proteins, providing an 8-nanometer–resolution map of dynamically interacting complexes *in vivo* and extended networks that provide insights into fundamental cellular processes, including cell polarization and autophagy, pathways that are evolutionarily conserved and central to both development and human health.

1.3. Introduction and results

The elucidation of protein-protein interaction networks (PINs, or interactomes) holds the promise of answering fundamental questions about how the biochemical machinery of cells organizes matter, information, and energy transformations to perform specific functions [32]. An essential and rarely addressed question is whether protein complexes and PINs that are reconstructed or reconstituted in vitro or removed from the normal context in which they are expressed reflect their organization in living cells. For eukaryotes, the test bed for large-scale analysis of PINs is the yeast *Saccharomyces cerevisiae*, where several PIN analyses have been performed using yeast two-hybrid screens (Y2H) [10, 11] or tandem affinity purification followed by mass spectrometric analyses (TAP-MSs) [12-14]. Each approach captures specific features of protein interactions; two-hybrid methods are best at measuring direct binary interactions between pairs of proteins, whereas affinity purification techniques best capture stable protein complexes. However, neither approach measures interactions between proteins in their natural cellular context, nor are not easily amenable to studying protein complexes that are transiently associated or dynamic under different conditions, that do not survive in vitro purification, or that cannot be transported to the nucleus and form interactions in the absence of other stabilizing interactions as necessitated in Y2H screening. Protein-fragment complementation assays (PCA) provide an alternative approach to detect protein-protein interactions (PPIs) in their natural context. In the PCA strategy, two proteins of interest are fused to

complementary fragments of a reporter protein. If the proteins of interest interact physically, the reporter fragments are brought together and fold into their native structure, thus reconstituting the reporter activity of the PCA (Fig. 1A). PCA strategies provide a simple, direct means for the detection of PPIs *in vivo*, and do so with endogenously expressed full-length proteins in their native post-translationally modified states and cellular locations [33]. Further, PCAs provide spatial and topological information about PPIs. Thus, a large-scale PCA screen would provide direct insights into the global structural organization of PINs as they exist in the living cell.

1.3.1.Genome-wide in vivo screen.

We have performed a systematic binary screen for PPIs at a genome-wide scale in *S. cerevisiae* using a PCA based on the murine dihydrofolate reductase (mDHFR) assay adapted to yeast describe in section 1.5. (Fig. 1A). The DHFR PCA is a survival-selection assay based on a mutant of mDHFR that is insensitive to the DHFR inhibitor methotrexate but retains full catalytic activity and allows detection of PPIs with as few as 25 to 100 complexes per cell [24, 34]. We created unique homologous recombination cassettes for all 5756 consensus genes with both the F[1,2] and F[3] complementary N- and C-terminal DHFR fragment sequences described in section 1.5..

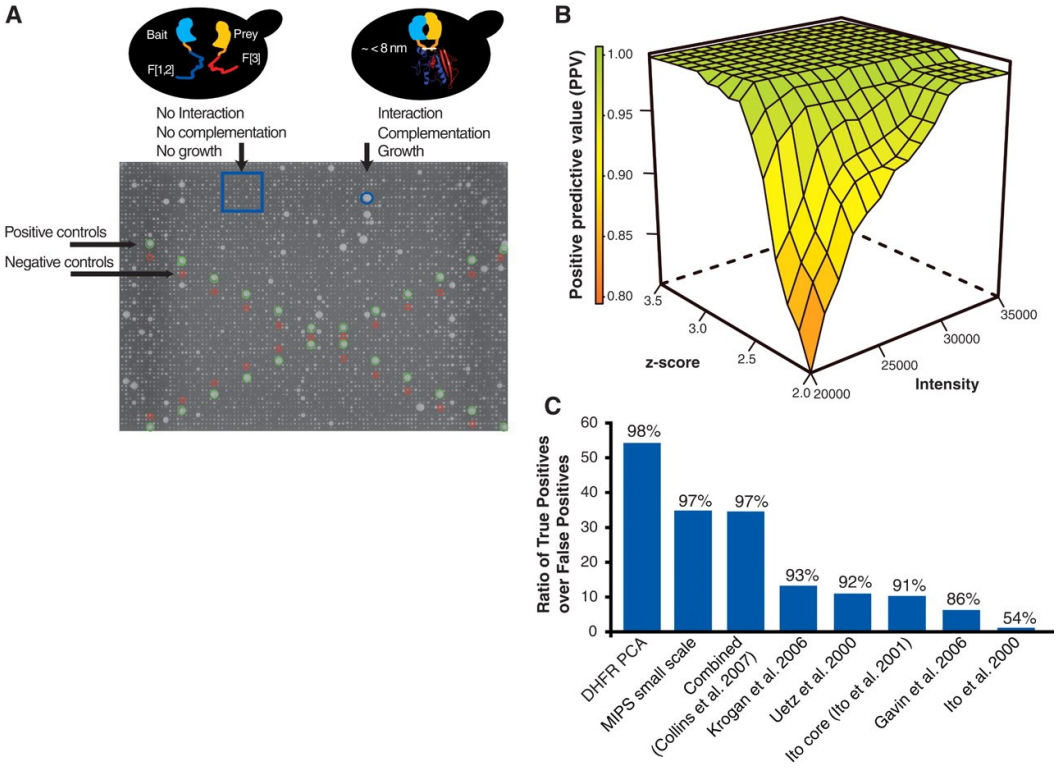

Figure 1 *In vivo* PCA screen of the yeast PIN.

(**A**) Strategy for high-density array screening of the yeast PIN by DHFR PCA. Both positive [green circles (*MAT***a**/α, CDC19 fused to DHFR fragment [1,2] (CDC19-F[1,2]), and MCK1 fused to DHFR fragment [3] (MCK1-F[3]) and negative [red circles (*MAT***a**/α, CDC19-F[1,2], and CLN3-F[3])] controls are included on each plate to ensure that each transfer and selection step has occurred correctly. (**B**) PPV score as a function of raw colony intensity and *z* score (relative colony intensity on plates). This score represents the ratio of the number of true positive interactions over the sum of the true positive and false positive interactions predicted from the reference sets. (**C**) The ratio of true positives to false positives in the DHFR PCA network compared with other large-scale data sets [10, 11, 13, 14, 35, 36]. The achieved PPV is indicated above the bars.

Successful cassette transformation of *S. cerevisiae* haploids was achieved for 4326 (75%) open reading frames (ORFs) with the DHFR F[1,2] fragment in MATa and 4804 (83%) ORFs with the DHFR F[3] fragment in MATα strains, with a final combined coverage of 5367 (93%) of all ORFs (table S1, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_tables.zip). The entire screening process was performed on solid-phase medium (Fig. 1A and Fig. 2). Briefly, MATa strains (F[1,2] fragment fusions) served as baits and were mated individually with all MATα (F[3]) strains on high-density arrays. The resulting diploids were transferred to a minimal medium [synthetic complete (SC)] plate to select for methotrexate resistance (reconstituted mDHFR activity, with native *S. cerevisiae* DHFR inhibited), and colony growth was recorded using automated analysis of digital images (Fig. 1A and Figs 3 and 4). PPIs were determined based on the growth of the diploid colonies measured by the pixel intensities on the selection plates (Figs. 2-4). In total, 3247 individual highly reproducible (Fig. 5) bait screens were performed, resulting in more than 15 million individual matings.

1.3.2.Data filtering, quality assessment, and overlap with existing PINs.

We experimentally accounted for two potential sources of false positives in a PCA screen: trapping of nonspecific complexes due to irreversible folding of the mDHFR reporter protein, and potential spontaneous complementation (folding) of the DHFR PCA fragments.

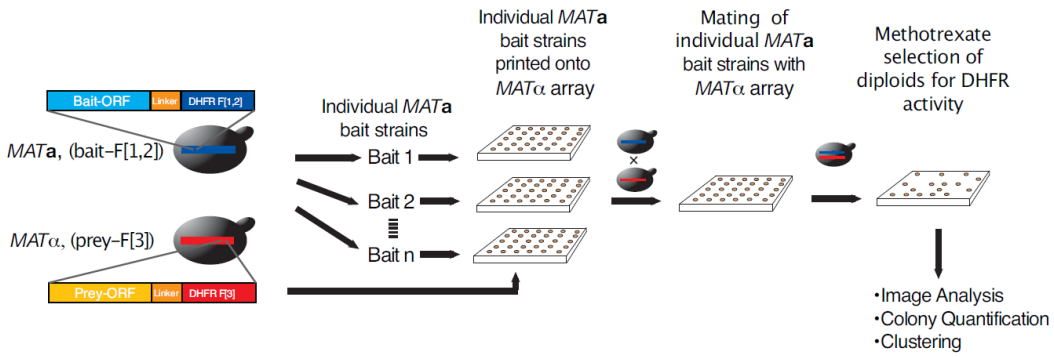

Figure 2: ***In vivo*** **PCA screen of the yeast protein interaction network.**

Strategy for single bait versus prey array screening of the yeast PIN by DHFR PCA.

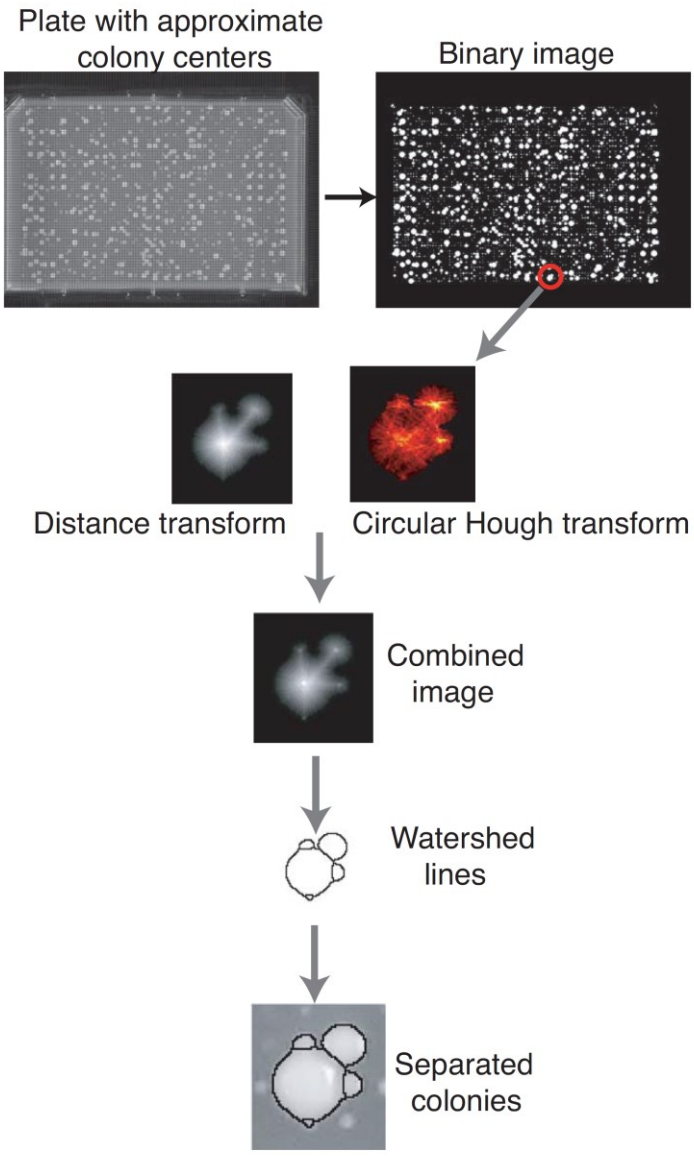

Figure 3: Automated extraction of colony intensities on plates.

The DHFR PCA results were inferred from the growth of diploid colonies on plates containing methotrexate. Images of the plates were taken after a 90-hour growth period with a 4.0 Mega pixel camera (Powershot A520, Canon). Plate images were saved in JPG format at a resolution of 180 dpi and a size of 2,272 × 1,704 pixels. In order to extract the intensity of the colonies, we used available image recognition routines available in the Matlab image analysis toolbox and we modified parameters for it to be able to differentiate colonies that are in proximity to each other. The quality of the position of the grid and the recognition algorithm was examined through visual inspection of all plates.

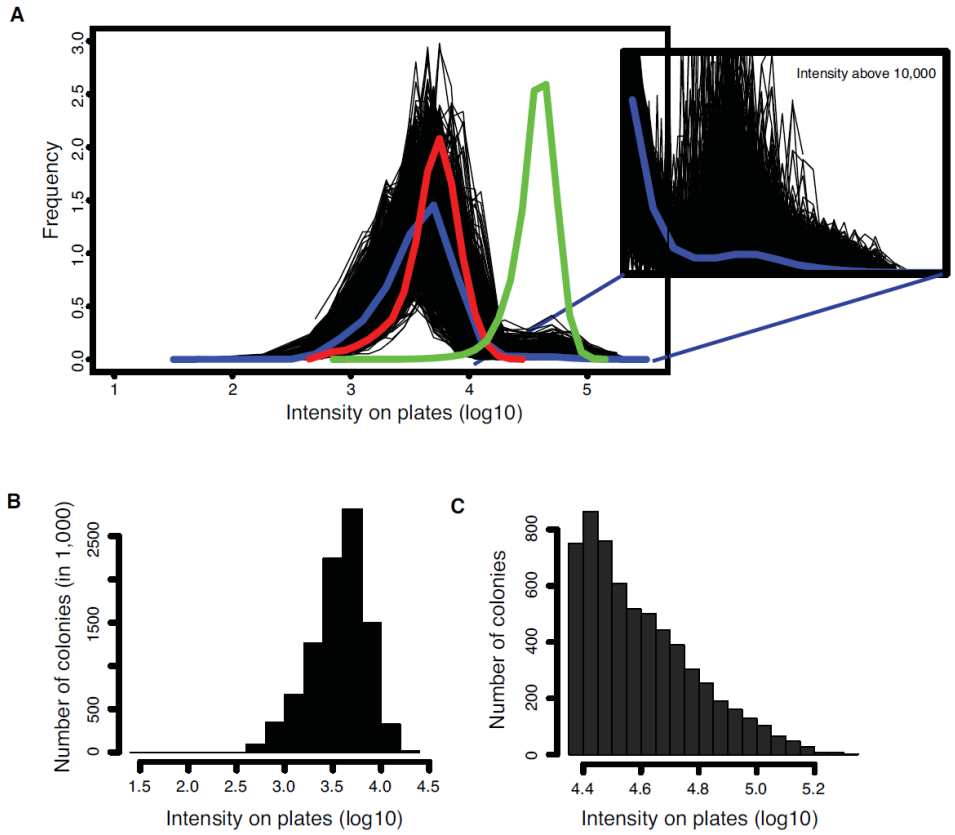

Figure 4: Distribution of colony intensities on plates.

(A) Raw intensities prior to filtering. Black lines represent the intensity distributions on individual plates. The blue line represents the distribution of colony intensities across the entire experiment. The red and green lines represent the intensity distribution of the negative and positive controls respectively. The second panel shows the distribution of colony intensities above 10,000 to illustrate the growth of the methotrexate resistant diploid strains. (B) Intensities of colonies below the threshold. (C) Intensities of colonies above the threshold after filtering positions corresponding to baits and preys that interact with the control fragments.

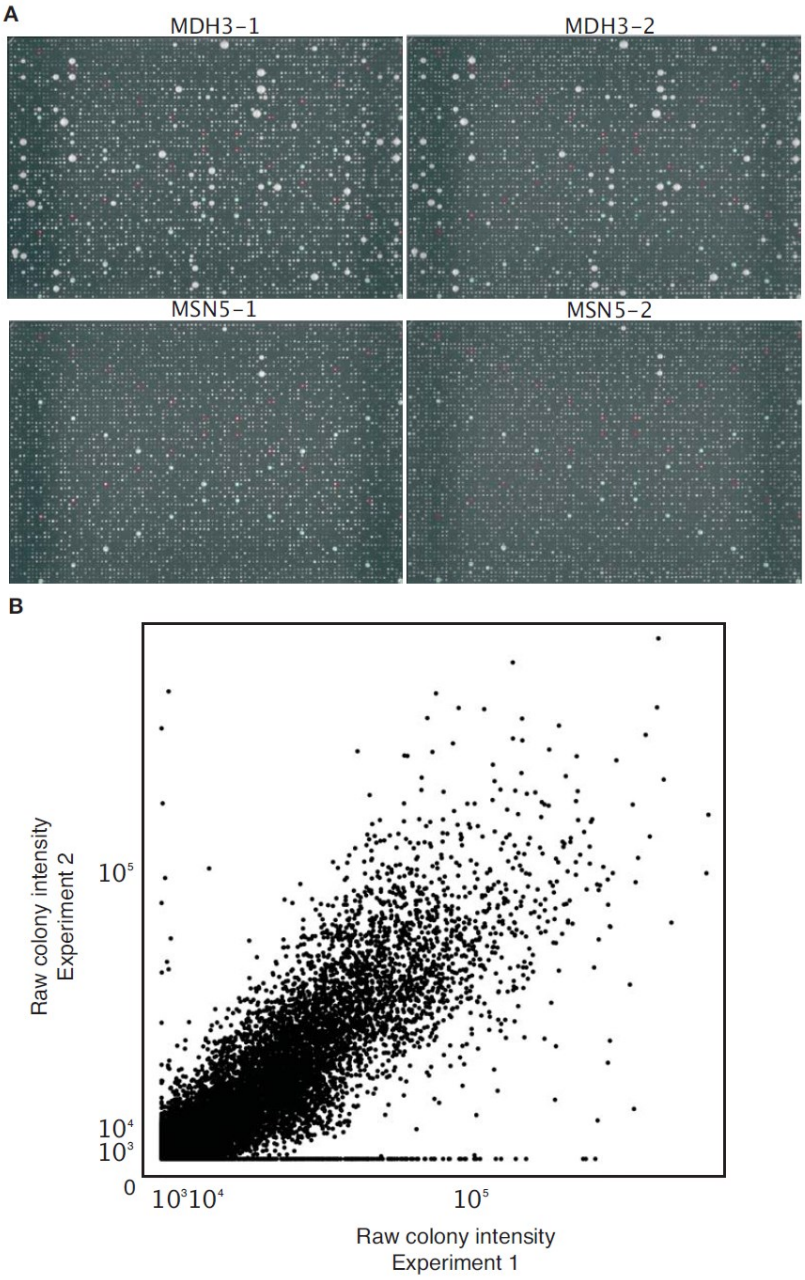

Figure 5: Reproducibility of the screening process.

In order to evaluate the reproducibility of the screening process, we repeated the screen for 48 baits selected for representing the distribution of number of colonies growing above background in the first screen. (A) Example of plates that were repeated. Top row, MDH3 and bottom row, MSN5 used as baits. Green and red circles represent respectively positive and negative controls. (B) Raw colony intensity of experiment two plotted against the raw colony intensity of experiment one (Pearson correlation: 0.86, $P < 2.2e\text{-}16$).

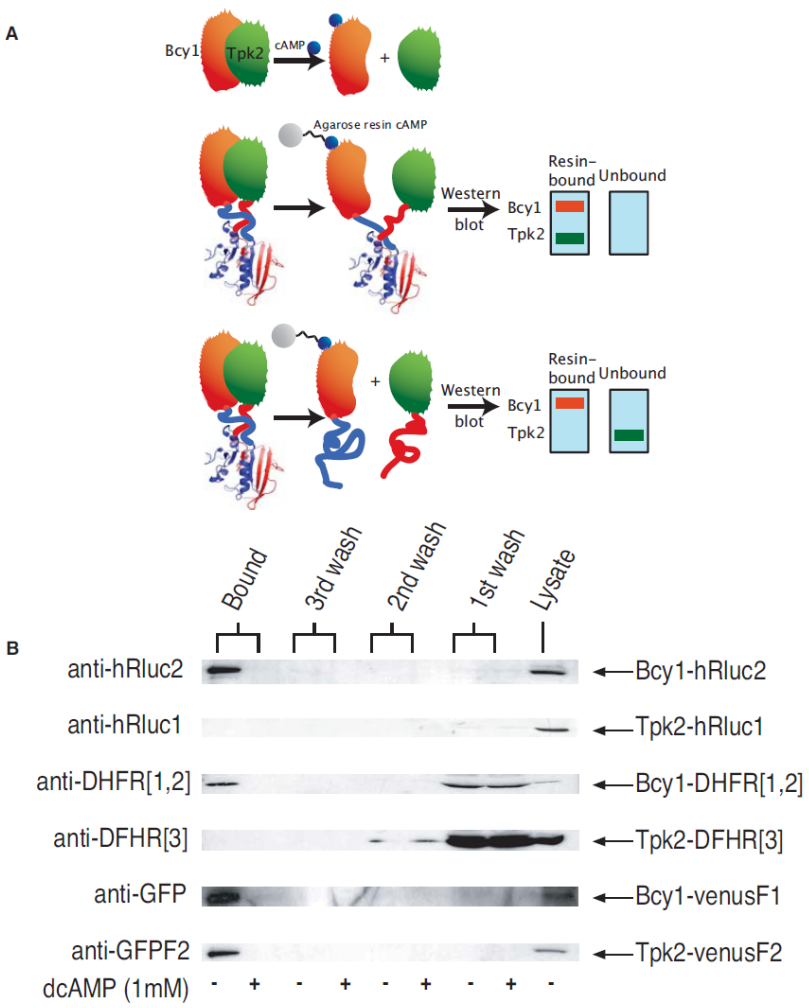

Figure 6: The DHFR PCA is reversible.

(A) Upper panel. cAMP-mediated dissociation of the yeast PKA regulatory (Bcy1p) and catalytic (Tpk2p) subunits. Middle panel. Schematic representation of an irreversible PCA for PKA and predicted results of the Bcy1 subunit binding to cAMP-conjugated agarose beads. For the irreversible PCA, Bcy1 and Tpk2 dissociate but remain trapped by the folded PCA reporter protein. Lower panel. For a reversible PCA, reporter protein fragments unfold and dissociate when Bcy1 binds to cAMP-conjugated agarose beads and thus Bcy1 remains bound to the resin while Tpk2 is found in the unbound supernatant fraction. (B) The DHFR PCA is fully reversible. As reported previously [21], the Rluc PCA is reversible; Bcy1 is found in the cAMP-conjugated agarose fraction while Tpk2 is found in the wash. Precisely the same result is found for the DHFR PCA, suggesting that it is reversible, while Venus YFP PCA is irreversible.

First, we used the adenosine 3′,5′-monophosphate–dependent dissociation of the yeast protein kinase A complex as a test system [21] to show that the DHFR PCA is fully reversible, and thus the trapping of complexes is unlikely (Fig.6). Second, we screened all the strains against the individual F[1,2] and F[3] complementary fragments or fragment-peptide linker sequences. This allowed us to eliminate 344 promiscuous, highly expressed proteins (Fig. 7 and table S2, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_tables.zip), several of which are also often observed as false-positives in affinity purifications. We next identified a threshold of colony intensity above which we could infer PPI. The Munich Information Center for Protein Sequences (MIPS) complexes were used as a standard set of true positives, along with 266,858 true negative interactions between proteins expressed in different cellular compartments or having negatively correlated expression [31, 36]. After several filtering steps describe in section 1.5. and benchmarking on the reference PPIs, we obtained a high-quality data set containing 2770 interactions among 1124 proteins that reach a positive predictive value (PPV) of 98.2% (Fig. 1B and tables S3 and S4, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_tables.zip). This resulted in data having precision (number of true positives relative to false positives) comparable to the MIPS small-scale experiments and all previous large-scale data sets (Fig. 1C and Fig. 8).

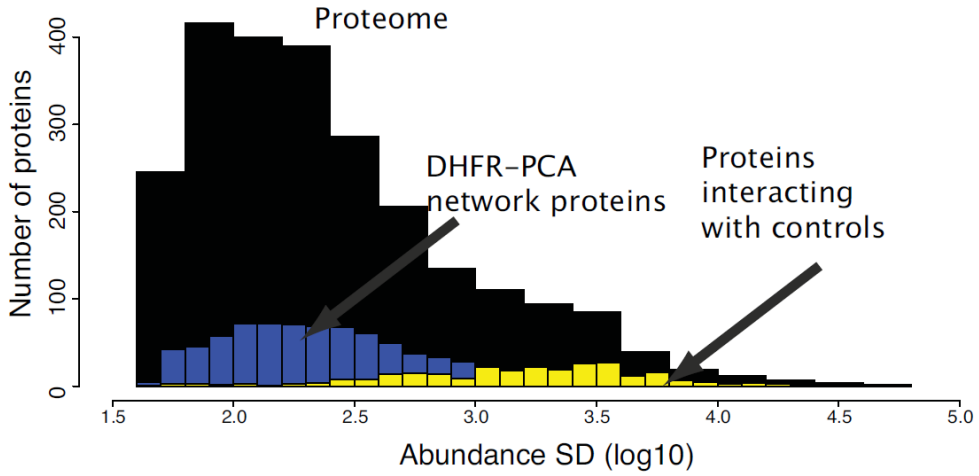

Figure 7. Distribution of protein abundance.

The distribution of protein abundance for cells grown on the same (SC, SD + glucose (from [37])) medium used in the DHFR PCA screen of the entire proteome (black), proteins of the DHFR PCA network (blue) and proteins interaction with the control fragments (yellow).

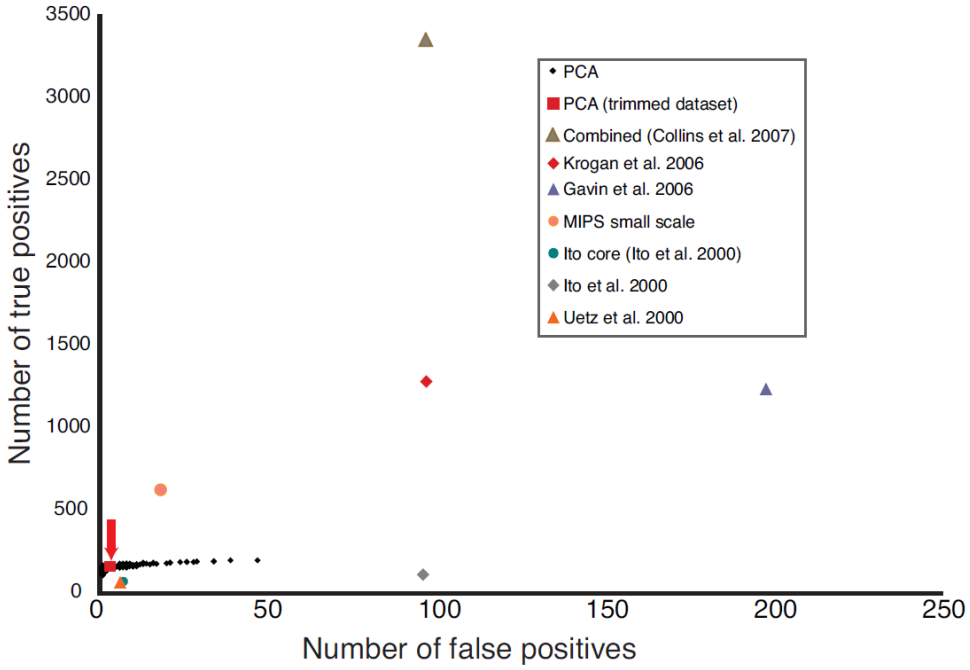

Figure 8: Quality assessment for the PCA networks and other PPI networks.

The curve represents the total number of true positive interactions and the total number of false positive interactions as a function of the score thresholds for defining PPIs in the DHFR PCA screen (ROC curve). Values for published datasets are shown as well as values of the final DHFR PCA networks. Sources for the other networks are described in the section 1.5.13.

The proteins in the DHFR PCA network are highly enriched in cellular compartments [for example, organelle membranes ($P < 10^{-12}$), proteasome regulatory particles ($P < 10^{-8}$), the nucleolus ($P < 10^{-7}$), and the cell cortex ($P < 10^{-7}$)] that were less represented in comprehensive TAP-MS results [36] (tables S5 and S6, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_tables.zip). The high sensitivity of the DHFR PCA assay is reflected in the abundance of the proteins that populated our network, which are on average only slightly more expressed than the proteome [the median log10(protein abundance) = 2.32 versus 2.28; Wilcoxon rank sum test, P= 0.19] and spanning the whole distribution of protein abundance (Fig. 7). Because this study was performed *in vivo*, with a technique never used at this scale and in a different medium than previous experiments, we expected that many interactions would be previously undiscovered. An examination of major databases of PPIs reveals that most of the interactions (~80%) we report are among protein pairs for which no data had been previously reported (Fig. 9A). However, when considering only PPIs that could be detected by both DHFR PCA and the other experiments, we confirmed between 16 and 41% of PPIs reported in previous large-scale screens, suggesting excellent concordance between the results of our and very different methods (Figs. 9B and 10). Further, PPIs derived from PCA represent pairwise interactions, which contrast with TAP-MS PINs, which identify clusters and thus complexes of interacting proteins. PPIs detected by PCA are therefore either within, between, or outside these complexes and thus complement these previous studies.

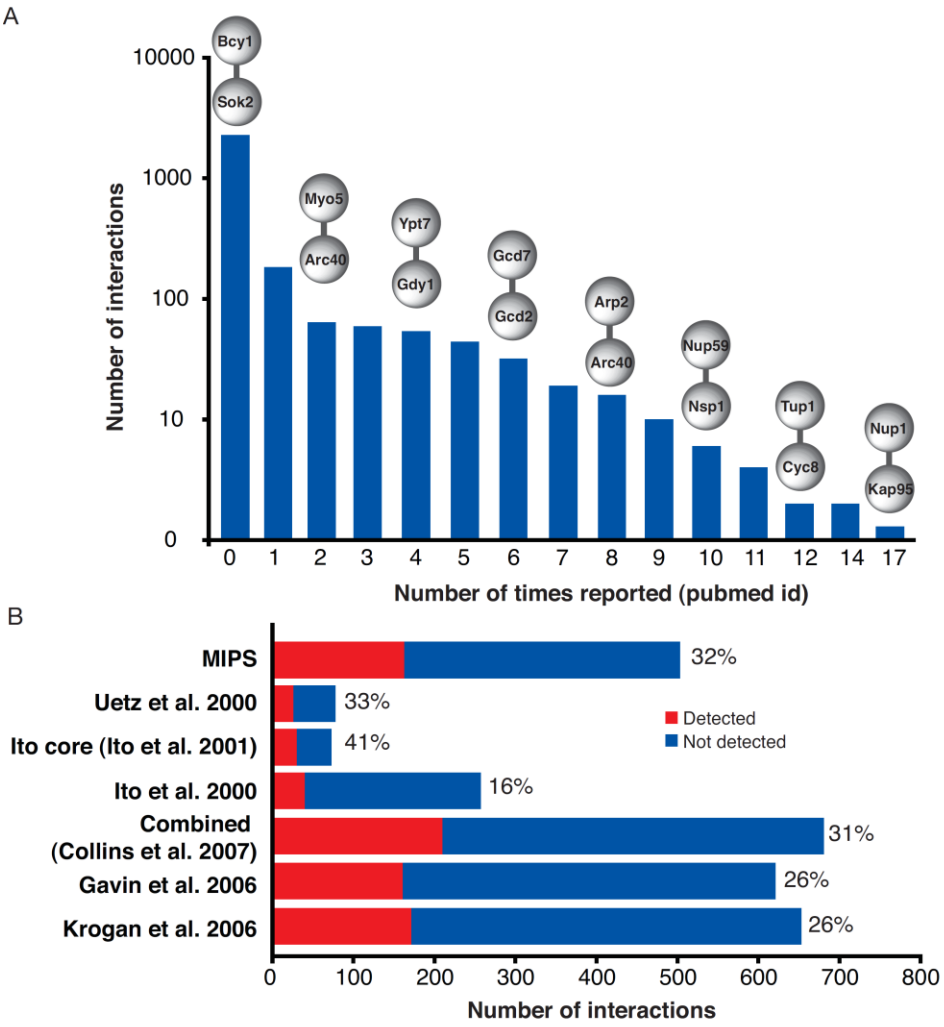

Figure 9: Overlap of the DHFR PCA network with other large-scale experiments.

(A) Most DHFR PCA PPIs are new, since they score 0 within the distribution of the number of times a known interaction has been independently deposited in major PPI repositories. Examples of interactions are shown above the bars. (B) The overlap of the DHFR PCA network is substantially increased when only the interactions that could be discovered are considered, i.e. only identified successful baits and preys are considered. Bars indicate the number of PPIs that could have been discovered by PCA. In red is the number of interactions that were discovered. Percentages indicate the percentage of interactions that were discovered by PCA out of the total possible.

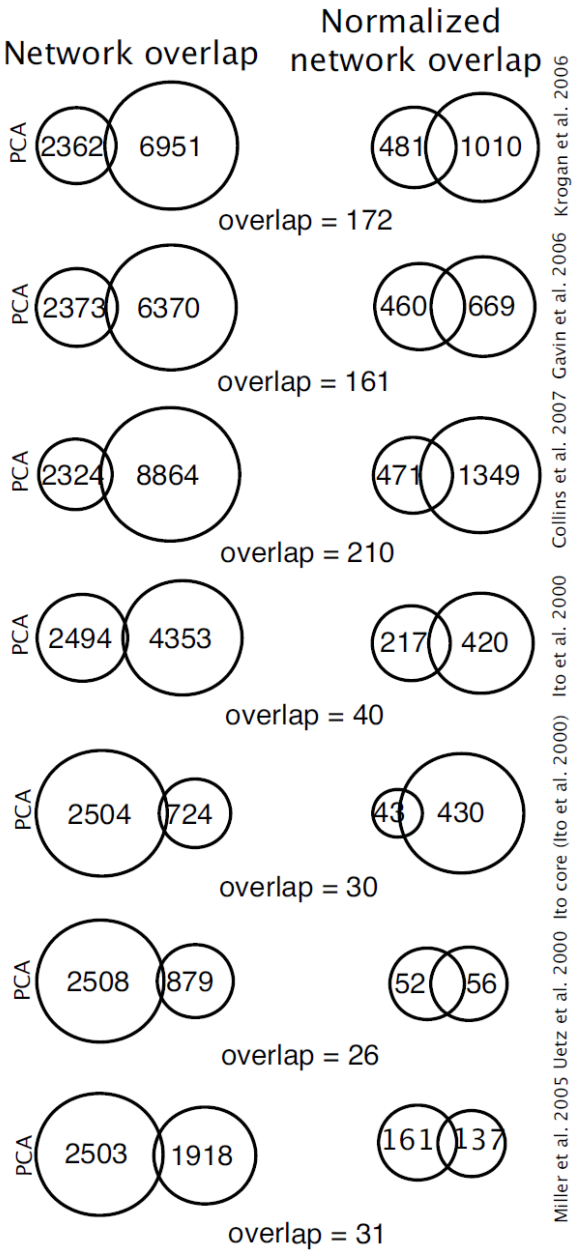

Figure 10: Overlap of the DHFR PCA network with other large-scale experiments.

On the left is the overlap between the different networks. On the right are the same overlaps, but only for those interactions that could have been detected in both experiments; i.e. cases in which the interactions were tested for in both experiments.

For instance, 10% of the DHFR PCA PPIs map within specific complexes in the combined analyses of the two TAP-MS data sets [13, 31], and 36 and 38% of the DHFR PCA PPIs are between one protein found in a complex and one protein not in the published data set, or two proteins not in the data set, respectively. We identified several interactions among complexes (15%), which probably mediate the integration of biological processes among PIN modules. For instance, PPIs occur among complexes that are more related in their functional annotations than would be expected to occur by chance [interacting protein pairs had a semantic similarity score of cell compartments (CCs) of 3.44 versus 1.64, $P < 10^{-100}$; of biological processes (BPs) of 3.48 versus 1.51, $P < 10^{-80}$; and of molecular functions (MFs) of 3.53 versus 2.3, $P < 10^{-10}$]. For example, we see interactions between Dhh1p and Lsm4p, both involved in the RNA metabolic process but part of the CCR4 and the RNA-splicing complex, respectively. Another example is the interaction between Reg1p and Snf1p: subunits of the serine/threonine phosphoprotein phosphatase and SNF1 complex, respectively, but both involved in the regulation of carbohydrate metabolic processes (table S7, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_tables.zip). Finally, we report 286 interactions involving one uncharacterized protein with proteins of known function (n = 278 interactions) or between two uncharacterized proteins (n = 8) [38], which will aid in their functional annotation.

1.3.3.General organization of the yeast DHFR PCA PIN.

Because we detected PPIs as they occurred in intact cells, with the faithful representation of gene expression timing and protein localization, we predicted and observed stronger coregulation of interacting protein pairs (Pearson r = 0.2 versus r = 0.1, P << 0.001) than was expected for random networks of the same size with the same protein connectivity. This is also mirrored in the enrichment of interactions among proteins that share the same BPs, MFs, and CCs and a depletion of interactions among genes of different categories (Fig. 11 and Fig. 12). PPIs among categories are somewhat more enriched in the PCA-determined network as compared with TAP-MS studies. For instance, 64, 56, and 63% of DHFR PCA interactions map to different BPs, CCs, and MFs, whereas these numbers are smaller in the TAP-MS PINs [58, 46, and 57% [14] and 51, 49, and 58% [13]]. Much of this increased enrichment of the cross-cellular components reflects interactions among proteins that the DHFR PCA method covers more of than TAP-MS; these are interactions that appear to represent the natural exchange of proteins between, for instance, the endoplasmic reticulum, Golgi, mitochondrial envelope, and vacuolar proteins, whereas others reflect the organization of complex cellular processes. For example, high enrichments in interactions between proteins localized to the bud and bud neck with those of the cell cortex, cytoskeleton, plasma membrane, and sites of polarized growth reflect the roles of these proteins in several compartments during cell division.

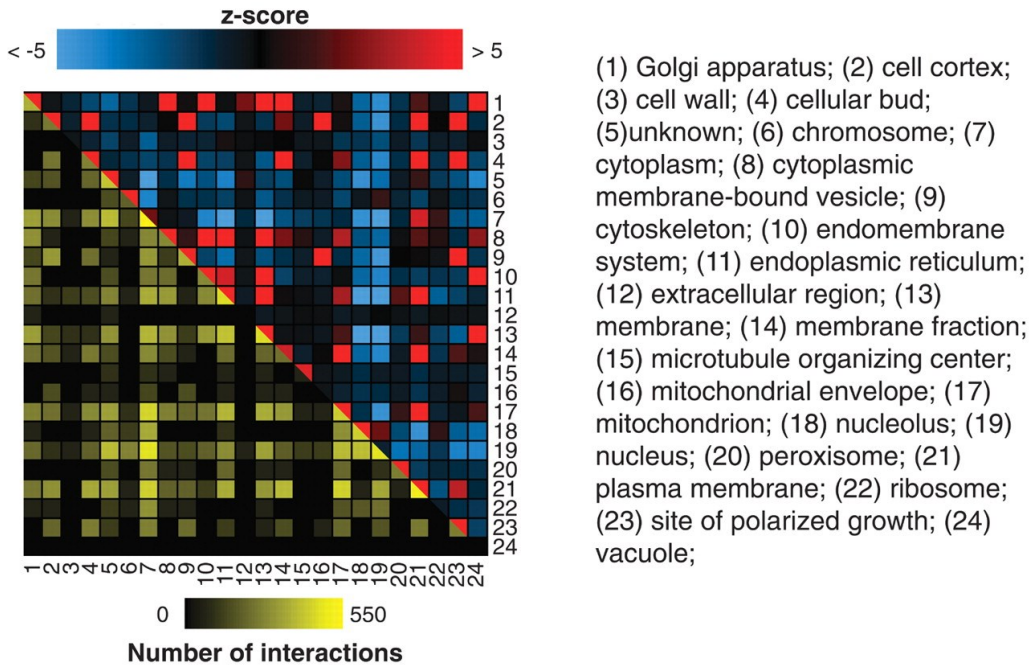

Figure 11. Interactions are enriched within GO categories.

The DHFR PCA network covers several classes of protein function, location, and biological process. The colors above the diagonal represent positive and negative deviations from the expected number of interactions between two cell compartments. A positive *z* score indicates a larger number of interactions within or between two categories as compared with a random network. A negative *z* score indicates a smaller number of interactions than expected. A *z* score of 2 or –2 corresponds to a *P* value of 0.05, and a *z* score of 5 or –5 to a *P* value of 5×10^{-7}. Values below –5 and above 5 were given these minimal and maximal values. Entries below the diagonal indicate the observed numbers of interactions on a $\log_{10}$ scale.

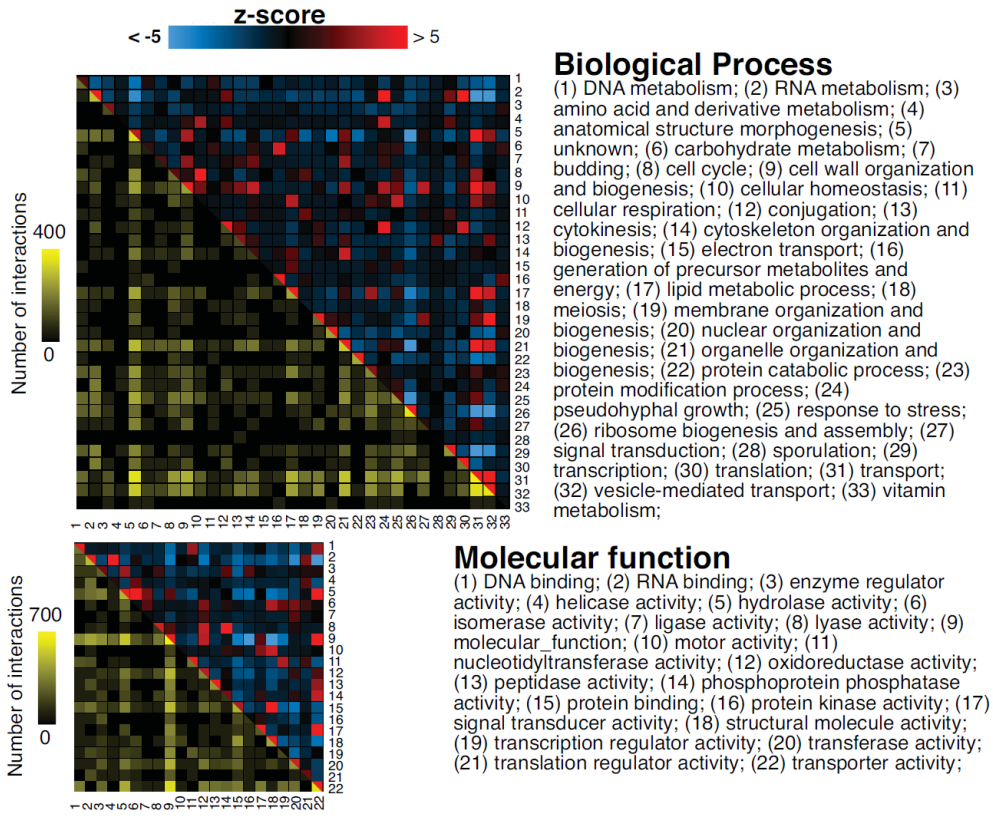

Figure 12: Interactions are enriched within Gene Ontology categories.

The DHFR PCA network covers several classes of protein function, location and biological process. The colors above the diagonal represent positive and negative deviations from the expected number of interaction between two categories, Biological Process or Molecular function. A positive z-score indicates a larger number of interactions within or between two categories compared to a random network. A negative z-score indicates a smaller number of interactions than expected. A z-score of 2 or -2 corresponds to a P-value of 0.05 and a z-score of 5 or -5 to a P-value of 5×10^{-7}. Values below -5 and above 5 were given these minimal and maximal values. z-scores were calculated by generating 10,000 random networks. Entries below the diagonal indicate the observed number of interactions on a log10 scale.

We also saw strong compartmentalization of interactions; for example, for nuclear and nucleolar proteins, which show enrichment in interactions between proteins in these two compartments but strong depletions in interactions with those of any other compartment. Equally, patterns of cross-process and molecular function categories reflect differences in complexity and organization (Fig. 12). For example, among molecular functions, RNA binding is specifically enriched in interactions between helicase and translation regulatory functions, whereas the more general transporter activity category shows links to diverse functions. The observation that PCA interactions detect links among functionally related categories is supported by a semantic analysis of the full Gene Ontology (GO) hierarchies. Proteins that show interactions with different GO Slim annotations have higher semantic similarities in their GO terms than expected by chance (CCs, 1.52 versus 0.94, $P < 10^{-231}$; BPs, 2.04 versus 1.35, $P < 10^{-122}$; and MFs, 1.89 versus 1.64, $P < 10^{-8}$), and may thus represent interactions relating information among these processes and CCs that allows their integration into higher-order networks. As we describe below, these interactions reveal specific spatial and topological relationships between known and previously unknown complexes underlying both known and previously unknown cellular processes.

1.3.4. Global structure and topology of the *in vivo* interactome.

PCA-detected interactions are interpreted differently than purified protein complexes or binary (one-to-one) interactions determined in Y2H

screens, allowing us to address how protein complexes and PINs are spatially and topologically organized in living cells. Whether an interaction can be observed depends on the distance between the C termini of two proteins and the length of the polypeptide linker separating bait and prey proteins to the PCA fragments [24, 34] (Fig. 1A). Given that the linkers used in this study were of 10–amino acid residues, for a given protein complex we expected to detect only binary (direct) or near-binary (indirect, C termini within 82Å, but mediated by one or more other proteins) interactions for protein pairs separated by less than this distance. We first tested this prediction by exhaustively screening all pair-wise interactions (n = 45 possible pairs) in the well-known RNA polymerase II complex (Fig. 13). We found that we were 5.7 times more likely to detect an interaction if the C-termini were within 82 Å (Fisher's exact test, P = 0.01) (Fig. 14B). Interactions that were detected but not predicted could be due to alternative assemblies of this complex in intact cells, to changes in their configuration under different conditions, or to protein dynamics that cannot be interpreted from crystal structures. We then asked whether spatial restraint on observable interactions is reflected in the complete DHFR PCA network.

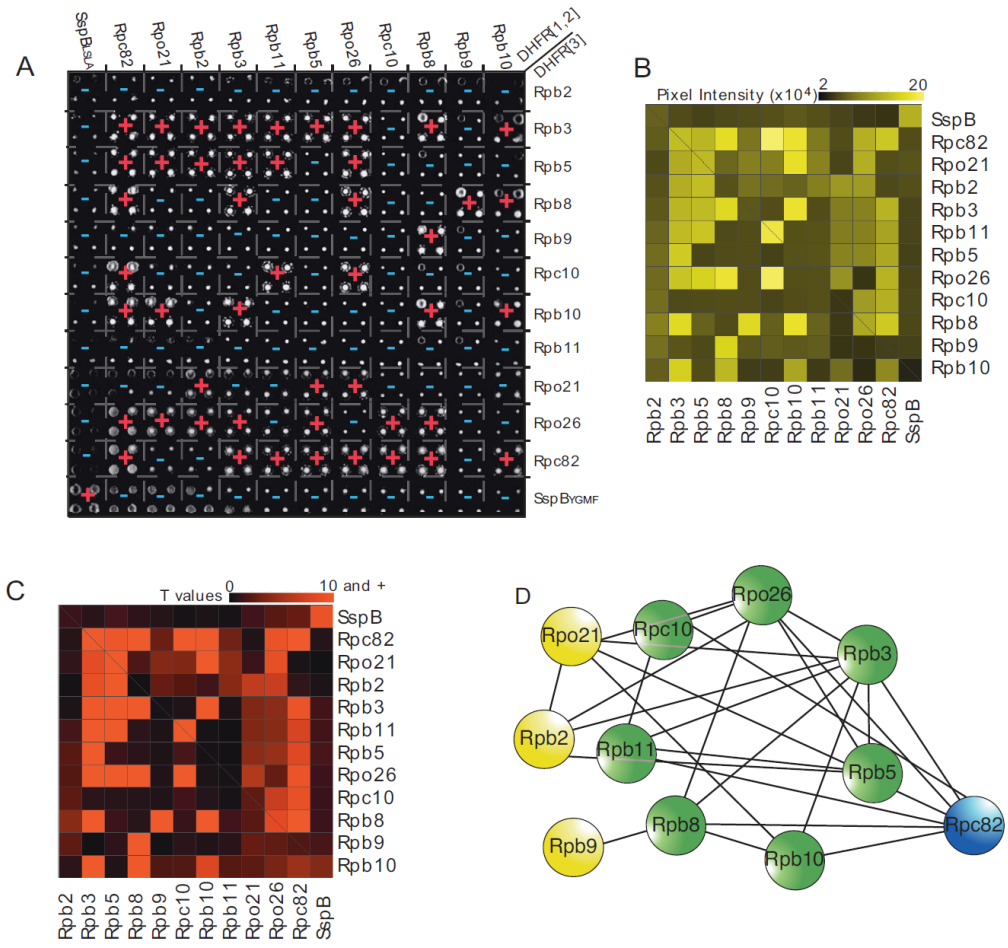

Figure 13. RNA polymerase II complex reconstitution through DHFR PCA.

(A) Results of the RNA polymerase II complex PCA network through an exhaustive screen for interactions among the ten subunits. Colonies for diploid strains that show resistance to methotrexate are indicated with red "+" and for those showing no resistance, with blue "-". (B) Mean colony pixel intensity values extracted from the high-resolution image in (A) by quantification of total colony pixel intensities. (C) t-scores for colony pixel intensities ranging from 0 to 10 and higher ($P < 0.0001$) resulting from the comparisons with control colonies. (D) Summary of the results of the RNA Pol II DHFR PCA screen where edges represent a physical interactions corresponding to a t-score of 4 and higher ($P < 0.05$) and nodes are the individual RNA Pol II subunits. Nodes colored in yellow, blue and green are respectively RNA Pol II exclusive proteins, RNA Pol III exclusive protein and RNA Pol II and RNA Pol III shared proteins.

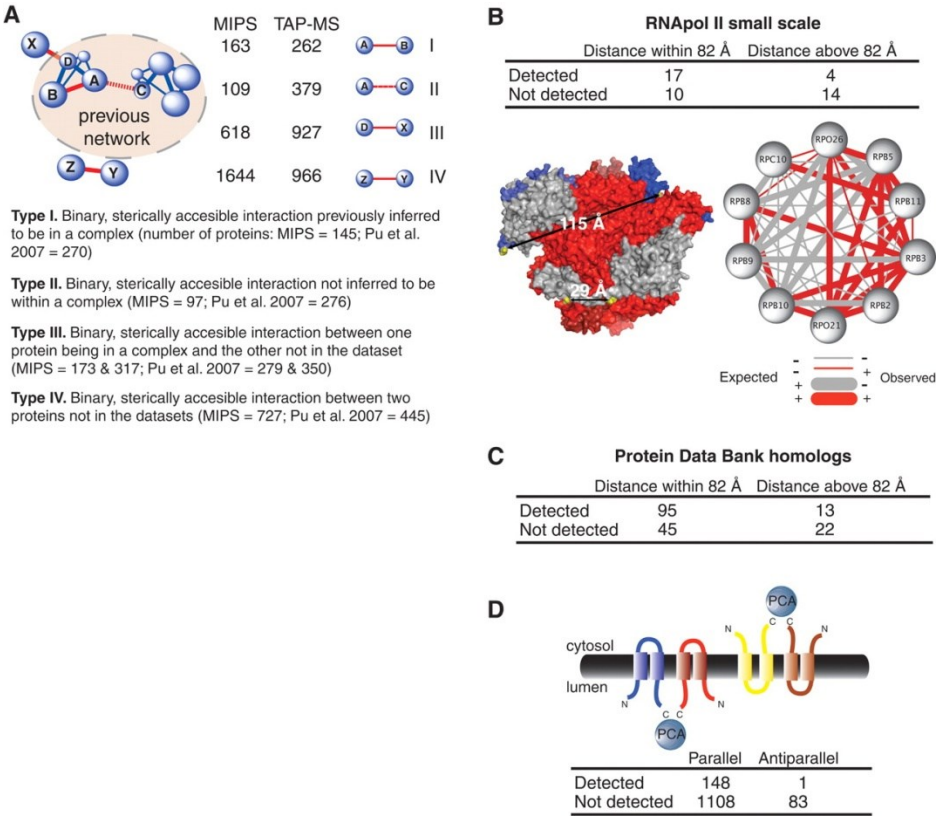

Figure 14. The DHFR PCA results provide structural and topological insights.

PCA fragments have to be in proximity to each other in order to fold into the active structure of the reporter protein. **(A)** PCA PPIs versus protein complexes. Comparison of the PCA network with databases of curated protein complexes (MIPS) and inferred from computational analysis of TAP-MS [31] allows classification of four types of PCA interactions: in which both proteins are found within a complex (type 1), are inferred to be in two separate complexes (type 2), one protein is in a complex and the other is not in the network (type 3), or both are absent from the network (type 4) [31]. Columns of numbers indicate the number of PCA PPIs observed for each data set and each category. **(B)** A thorough DHFR PCA screen of the RNA polymerase II complex [Protein Data Bank (PDB) number 1I3Q] detects predicted interactions among the 10 subunits. **(C)** An interaction is 3.5 times more likely to be detected for a pair of proteins known to interact if the C termini of these proteins are within 82 Å of each other in the case of stable crystallized complexes of yeast-homologous proteins deposited in the PDB. (D) Membrane protein topology and PPI detection by PCA. A protein interaction is 12 times more likely to be detected if the C termini are in the same cell compartment.

An examination of homologous protein complexes with solved structures showed that we were 3.5 times more likely to detect an interaction between a pair of proteins that have C termini closer than 82 Å than for those with longer distances between C termini (Fisher's exact test, $P < 0.002$) (Fig. 14C). Further, we found that the interacting protein pairs possess domains known to mediate PPI more often than they would be expected to have possessed by chance (7.3% of protein pairs have domains known to mediate PPI versus 0.6%, $P << 0.001$). Thus, the data will be useful for predicting spatial relationships and the bases of molecular recognition among proteins, protein domains, and peptide recognition motifs. Finally, because the C termini of proteins have to be in close proximity and also oriented into the same cellular compartment, PCA provides information about membrane protein topology (Fig. 14D) [24-27]. Our results demonstrate that the topology of interacting membrane proteins is also reflected in the PIN; specifically, that membrane proteins that colocalize to the same cellular compartment are 12 times more likely to show an interaction if they have a parallel rather than an antiparallel orientation (Fisher's exact test, $P < 0.0005$) [39]. These PPIs between membrane-associated and soluble proteins will serve to predict cross-compartment functional relationships, such as interactions of endoplasmic reticulum–associated membrane receptors and cytosolic or nuclear effector proteins.

1.3.5.Bird's-eye view of the yeast *in vivo* PIN.

The general predictions described above led us to pose specific hypotheses for how protein complexes and networks are organized in living cells. Unsupervised hierarchical clustering of the 2770 DHFR PCA interactions provides an overview of the *in vivo* PIN (Fig. 15 and file S1, http:// www.sciencemag.org /content /suppl /2008 /05 /08 /1153878.DC1 /1153878s_data.zip). A number of crystallographically or biochemically well characterized complexes are organized as clusters along the diagonal, confirming that their organization in cells reflects their predicted structures in vitro. Also, substructures of these clusters are consistent with those of previously affinity purified subcomplexes. For instance, the nuclear pore contains a number of distinct subclusters, three of which clearly correspond to known subfractionated complexes (the Nup84 subcomplex includes Nup85, Nup120, and Nup145, which are in the network, and Nup84 and Seh1p, which are not in the network; a second subcomplex that includes Nup57, Nup49, and Nsp1, which are in the network, and Nic96, which is not in the network; the Nup82 subcomplex includes Nup159, Nup82, and Nsp1, which interacts with Nup166 for its proper localization) [40]. These subcomplexes also represent groups of proteins that have been hypothesized to form direct contacts in a detailed architectural map of the assembly of the nuclear pore complex [41]. Our results now suggest that such substructures exist in intact cells.

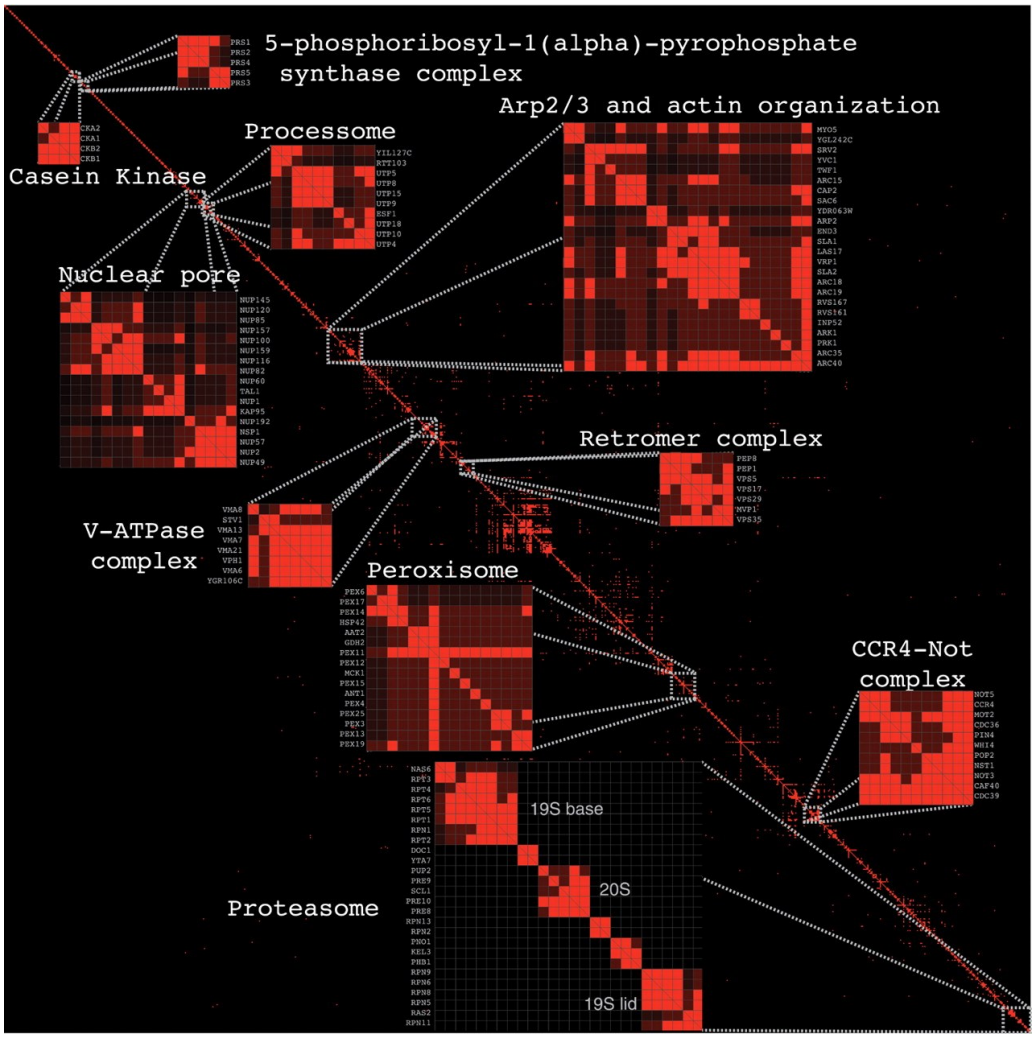

Figure 15. The DHFR PCA network is modular and interconnected.

Clustering of the DHFR PCA network reveals numerous known complexes, within which the substructure represents known subunits. Proteins that have interaction patterns similar to those of other proteins and that interact together are grouped along the diagonal.

Similarly, the proteasome partitions into subcomplexes that correspond to the composition of characterized fractions and of structures that can be visualized in intact cells [42]. Complexes described in vitro can therefore accurately reflect those seen *in vivo* by PCA and as reported by whole-cell electron tomographic studies of protein complexes [43]. PPIs between complexes that reflect the cross-compartmental and cross-functional interactions described above (Fig. 11 and Fig. 12) can be visualized as off-diagonal interactions on this map (Fig. 15). These represent links among several network modules that have been well described and shown to be central to eukaryotic cell biology. Our map therefore allows us to identify previously unknown multifunctional PINs and to associate and integrate other proteins to these processes (Fig. 16). First, we showed that, starting from the Arp2-actin organization network, we are able to describe new connectivity among the complex network of interactions that integrates actin filament assembly and patch formation with secretion and cell-wall synthesis and ultimately with membrane abscission and cell separation during cell division. Second, we showed that starting from the retromer complex, we can physically integrate the protein-sorting machinery and trafficking with the synthesis of autophagosomes, links that were previously suggested from genetic and cell-biological observations.

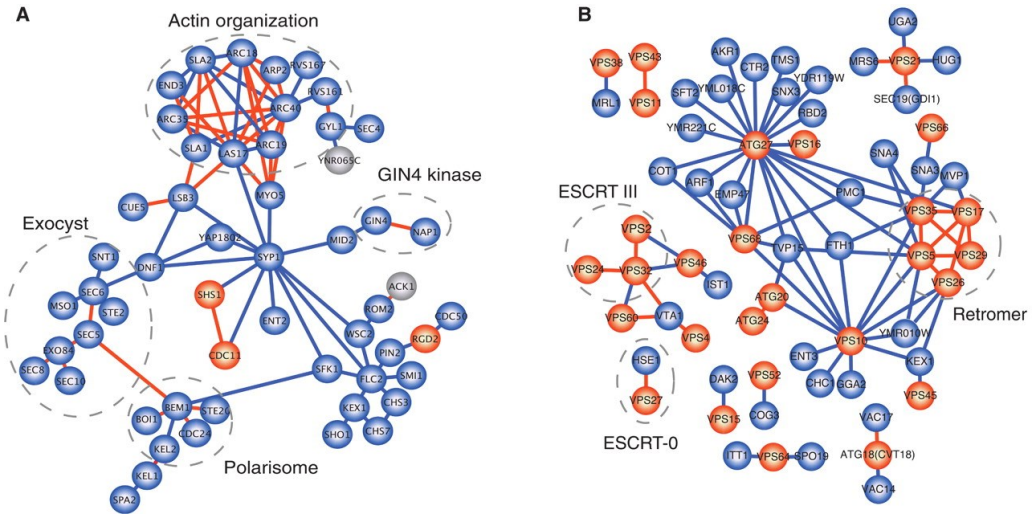

Figure 16. The yeast DHFR PCA network provides insights into both cell polarity and autophagy.

Blue edges denote previously unknown interactions (*10*) and orange edges denote interactions reported at least once in major databases. (**A**) Network at the bud neck. Physical association, detected by PCA between proteins that localize to the bud and bud neck (blue) and proteins that localize predominantly to the bud neck (orange), and proteins lacking localization data (gray), can be used to assemble the structure of a polarity PIN. This PIN shows both known interactions and previously unknown coupling between proteins involved in actin-filament organization and patch and assembly with proteins acting in secretion and cell-wall synthesis. The interactions between protein complexes in the PIN reflect the complex transition of proteins between the bud tip and the bud neck, which function in cell polarity from bud emergence to cytokinesis. (**B**) Autophagy network. Interactions directly connecting proteins involved in autophagy (ATG), vacuolar protein sorting (VPS), and cytoplasm-to-vacuole targeting (CVT) (orange) and other proteins (blue) are shown.

1.3.6.*In vivo* network at the bud neck.

Successful completion of a mitotic cell cycle, and the creation of a viable daughter cell, involves the tight coordination of nuclear events with mechanisms that control cell morphology. The formation of a bud is an example of this temporal and spatial coordination. The bud is the recipient of the segregated material from a mother cell, and ultimately will form the daughter cell. DNA, organelles, proteins, and mRNAs are selectively transported into the bud from the mother and after cytokinesis are enclosed within the daughter cell [44]. After bud-site selection, recruitment and assembly of proteins that act in polarized growth and bud emergence occur at the incipient bud site. We can construct a network (Fig. 16A) that captures the dynamic assembly and localization of "polarisome" proteins and both known and previously unknown interactions to proteins that provide input signals from the cell-cycle machinery via specific cyclin–cyclin-dependent kinase 1 (Cdk1) complexes and the Rho signaling proteins that control polarity and cell integrity [45]. These mechanisms organize and polarize the cytoskeleton and the secretory apparatus at the bud tip and bud neck during cell-cycle progression. Many polarisome proteins that localize to the bud and bud neck (Fig. 16A, blue), or predominantly to the neck (Fig. 16A, orange) are found in the PCA network. Bem1 plays a central role in bud polarization through its ability to build scaffolds, at sites of polarized growth, of an activator (Cdc24) and an effector (Ste20) of the Cdc42 Rho-like guanosine triphosphatase (GTPase) [46]. Kel1 and Kel2 also act as scaffolds for polarity components at the bud tip and

bud neck and were shown to couple to Bem1 and Spa2. Further, the exocyst complex functions in the vectoral transport of vesicles from the Golgi to the bud and promotes plasma membrane expansion, and the Arp2/3 complex, by mediating the assembly of actin patches, promotes membrane recycling through endocytosis [47]. An extensive network of proteins containing the Arp2/3 actin-assembly complex, its activator Las17, and effectors of actin organization is represented in our network and recapitulates many known protein interactions (Fig. 16A, orange edges), but extends the level of connectivity among components (Fig. 16A, blue edges), especially for Sla2, Las17, and Arc40. However, Arc40 of the Arp2/3 complex is linked via Rvs161 to the GTPase-activating protein (GAP) Gyl1, known to function in actin-patch formation and polarized exocytosis; Gyl1 is also connected to the exocyst through the Sec4 GTPase and to Ynr065C (Fig. 16A), a large protein of previously unknown function and for which localization data are unavailable. In a further extension of this actin patch–assembly complex, we found that Las17 and Myo5, a type I myosin that associates with actin patches, interact with Syp1, a protein implicated in actin cytoskeletal organization. Further, we showed that Syp1 physically associates with multiple proteins, including the bud-neck septins Cdc11 and Shs1 and the cell-surface sensors Mid2 and Wsc2, which activate the cell-integrity pathway through Rom2.Collectively, the interactions among distinct complexes seen by PCA represent a potential regulatory network involved in bud polarization, bud-neck organization, and cytokinesis, a network

that captures the dynamic transitions of polarity and exocyst components between the bud tip and the bud neck during the cell cycle.

1.3.7. *In vivo* network of autophagy.

Autophagy is the process whereby organelles and the cytosol are engulfed within membrane vesicles for delivery to the lysosome/vacuole for degradation and macromolecule recycling and is involved in development, response to stress, and pathogen resistance. Dysfunction of this conserved eukaryotic process is associated with neurodegenerative conditions, namely Huntington's, Alzheimer's, and Parkinson's diseases, and with cancer [48]. Proteins involved in autophagy are rich in interactions in the yeast DHFR PCA network (Fig. 16B), including the endosomal sorting complexes required for transport (ESCRTs) ESCRT-0 and ESCRT-III, the retromer complex, and other known interactions (Fig. 16B, dashed gray circles). Vps32, Vps24, and Vps2 are three of the four subunits of the ESCRT-III complex that are responsible for the sorting of transmembrane proteins into the multivesicular body (MVB) pathway. Dysfunction of this complex leads to autophagosome accumulation and neurodegeneration in mammals [49]. ESCRT-0 is required for sorting ubiquitinated membrane proteins before vacuolar degradation. Vta1 is a member of the MVB pathway and is known to bind to Vps60 and Vps4, regulating the activity of the latter [49]. Proteins destined for secretion or for delivery to intracellular compartments follow the same route and are sorted in the trans-Golgi network. In yeast, the lysosomal/vacuolar proteins are sorted from other proteins by the carboxypeptidase Y receptor Vps10. These receptors

are then returned from the prevacuolar compartments to the trans-Golgi network by the retromer complex. The role of the retromer complex in protein transport is also crucial in metazoa because, for example, it is essential to the formation of important morphogen gradients along body axes [50]. Some of the interactions we uncovered shed light on this functional relationship between protein sorting and trafficking by the retromer complex and Vps10. For instance, we find that Vps10 shows physical interactions with the retromer complex, which was previously hypothesized on the basis of genetic interactions with Vps35 and Vps26 [51, 52]. These observations also suggest a topological relationship between these two proteins that add to our understanding of the structural organization of the retromer complex recently resolved by crystallography [53]. The interaction between Chc1 (clathrin heavy chain 1 human homolog), the clathrin heavy chain involved in protein transport and endocytosis, and Vps10 was also hypothesized based on genetic data that shows Vps10 is rerouted to the plasma membrane in a chc1 vps1 mutant instead of its normal travel to the endosome [54]. We confirm this functional relationship and show that it is mediated through a physical interaction. Most of the interactions we see are not previously described (75%) (Fig. 16, blue edges) and represent a substantial advance in describing the autophagy and cytoplasm-to vacuole targeting pathways. For instance, Atg27 shows a particularly large number of interactions. This protein plays a critical role in the formation of sequestering vesicles, including autophagosomes. It localizes to the Golgi apparatus, the mitochondrion, and the phagophore assembly site. Despite its

importance, it showed no interactions in recent large-scale TAP-MS experiments and only one in previous Y2H screens [35]. Recent work affirms its involvement in both bulk and specific autophagy, and it is hypothesized that Atg27 (along with Atg9, not represented here) labels the membrane source for its transport to and the formation of autophagosomes [55]. Our results suggest that Atg27 occupies a central role in autophagy because it physically interacts with the retromer complex and with many other vacuolar proteins involved in the sorting of vacuolar hydrolases; further, these results implicate uncharacterized ORFs, such as YML018C, YMR221C and YDR119W, in this process.

1.4.Conclusions.

There remain many insights to be drawn beyond the general details, overview, and examples of extended structural and functional networks reported here for the *in vivo* protein interactome, and other dimensions of the interactome remain to be explored: How dynamic are these interactions? What are the effects of growth conditions on PPI network architecture? The functional and integrative genomic tools developed for this study will enable analysis of protein-interaction dynamics on any scale to uncover mechanisms of biochemical network regulation. A wide variety of PCA reporter enzymes can be used to study temporal and spatial dynamics of protein interactions over a broad range of time scales (from seconds to many hours) and under the influence of natural or artificial perturbations [33]. Further, the topological requirements of PCA generate a protein-complex topology map at 8-nm

resolution that will provide reference data for studying the spatial dynamics of functional protein complexes by immunofluorescence or by monitoring the localization of proteins genetically tagged with fluorescent proteins. Finally, they will also provide reference constraints for determining the architecture of macromolecular assemblies [41]. The integration of the results from such efforts with those of gene regulation dynamics and protein modifications will lead to a fuller understanding of how complex cellular processes are orchestrated at a molecular and structural level in the living cell.

Supplementary Table and File is linked to the online version of the paper at www.sciencemag.org/cgi/content/full/1153878/DC1.

Acknowledgements We thank N. Aubin-Horth and D. Klionsky for comments on the manuscript, J.-F. Paradis for technical and management support, J. Vlasblom for providing scripts and data files, and N. Barkai for providing the gene expression profile compendium. Supported by grants from Genome Canada, Genome Quebec, and the Canadian Institutes of Health Research (CIHR) (MOP-152556) to S.W.M. S.W.M. holds the Canada Research Chair in Integrative Genomics. C.R.L. is a Natural Sciences and Engineering Research Council of Canada and CIHR Strategic Training Program in Bioinformatics postdoctoral fellow.

Author contributions V.M., S.R., M.S., I.S., and Y.M. performed the experiments; S.R. designed the automated robotic procedures; K.T., C.R.L., and S.W.M. analyzed the results; and H.B. helped plan the research and edited the manuscript, which C.R.L., S.W.M., V.M., and J.V. wrote.

1.5. Methods

1.5.1. Adaptation of the mDHFR Protein-fragment Complementation Assay for studies in yeast.

The mDHFR PCA (henceforth DHFR PCA) was previously developed for *E. coli,* plant protoplasts and mammalian cell lines [20, 24, 56]. To adapt the DHFR PCA for high-throughput screening in *Saccharomyces cerevisiae* we created a double mutant (L22F and F31S) that is 10,000 times less sensitive to methotrexate than wildtype scDHFR, while retaining full catalytic activity [22]. The L22F-F31S double mutant was created by introducing by site-directed mutagenesis the L22F mutation into the existing F31S mutant DHFR PCA N-terminal fragment (F[1,2]) previously developed for studies in mammalian cells [24]. The mutant was designed so that methotrexate inhibition of the wildtype scDHFR activity would be complemented by the activity of the reconstituted DHFR PCA reporter. In order for a PCA to minimally perturb the natural kinetics of protein-protein interactions and to prevent trapping of non-specific complexes, it has to be reversible; the reporter protein has to unfold and the fragments dissociate upon the disruption of a protein complex. PCAs based on the *Gaussia princeps* and *Renilla reniformis* luciferase have been shown to have this property whereas those based on GFP variants are irreversible [21, 22]. To directly assay the reversibility of the DHFR PCA, we used an *in vitro* assay that uses the cAMP dependent dissociation of regulatory (Bcy1) and catalytic (Tpk2) subunits of the yeast homologue of serine/threonine kinase PKA. In this assay, binding of the Bcy1 regulatory subunit to resin-immobilized

cAMP is probed and whether the Tpk2 catalytic subunit is dissociated or remains associated with Bcy1 can be assessed [21, 57]. These experiments allow us to directly distinguish two possible outcomes of Bcy1:Tpk2 dissociation (Fig. 6A): i), the PCA is a trap; that is, while a PKA complex may dissociate on binding to the cAMP resin conjugate, the PCA fragments do not unfold and thus both catalytic and regulatory proteins remain bound to the resin; ii), PKA subunits dissociate and the PCA fragments unfold and separate, resulting in the release of the catalytic domain from the cAMP resin. In this case, the regulatory subunit remains bound to the resin while the catalytic subunit would be found in supernatant. We tested the reversibility of the DHFR PCA in direct comparison to the Rluc and "Venus" YFP PCAs, which we previously showed to be respectively reversible and non-reversible [21]. We constructed two sets of yeast expression vectors harboring fusions of Bcy1 and Tpk2 to complementary fragments of the Rluc, YFP and DHFR PCAs under control of a constitutive TEF promoter. Bcy1 and Tpk2 genes were PCR amplified from the *S. cerevisiae* genome and subcloned into yeast expression vectors harboring mDHFR, Renilla luciferase (Rluc) or "Venus" YFP PCA fragments fused 3' to Zip-linker, replacing the Zip sequences (p413-TEF-Zip-linker-F[3], p415-TEF-Zip-linkers-F[1,2], p413-TEF-Zip-linkerhRluc- F[1] and p415-TEF-Zip-linker-hRlucF[2], p413-TEF-Zip-linker-Venus YFP F[1] and p415-TEF-Zip linker-Venus YFP F[2]). Approximately 100 ng of each of the complementary PCA fragments expression vectors were cotransformed into *S. cerevisiae* BY4743 strain (*MAT*a/α *his3Δ/his3Δ leu2Δ/leu2Δ lys2Δ/LYS2*

*MET15/met15*Δ *ura3*Δ*/ura3*Δ) and positive clones were selected in synthetic complete medium (SC) –methionine, -lysine, -uracil and –histidine for p413/p416 transformed strains. BY4743 diploid strains harboring plasmids encoding the three DHFR, Rluc or "Venus" YFP Bcy1-Tpk2 PCAs were grown in 5 ml of synthetic complete medium (SC) (-met, -lys, -his, -leu) to an OD600 of 1.0, harvested and treated with 200 units of lyticase (Sigma-Aldrich, St-Louis, MO) for 2 hrs at 30°C to digest the cell wall. We extracted soluble lysate and incubated with the cAMP resin followed by a series of washes and then probed supernatant and bound fractions with antibodies that can bind to each PCA fragment (Fig. 6B). Specifically, the protoplasm was harvested at 500xg for 30 minutes at 4°C and resuspended in 0.1% volume of the original culture volume of phosphate-buffered saline (PBS) (pH 7.4) + 1% (v/v) Triton-X-100 + protease inhibitors (100 μg/ml) PMSF and MiniComplet EDTA free protease inhibitor tablets (Roche, Indianapolis, IN) and stored on ice for 30 minutes. The lysate was clarified by centrifuging at 12,000xg for 30 minutes at 4°C. The soluble fraction of each lysate was separated into two equal aliquots (volume of 250 μL). The first aliquot was incubated for 15 minutes with 1 mM dibutyryl cyclic AMP (DBcAMP) dissolved in water (Biolog, Bremen, Germany) and the second aliquot with water only.. 30 μL of 8- (2-Aminoethylamino)adenosine-3', 5'-cyclic monophosphate (8-AEA-cAMP) crosslinked to agarose beads (Cedarlane Laboratories ltd, Canada, ON) were added to each of the two aliquots and incubated at 4°C for two hours. The agarose beads were then washed three times successively with equal volumes of phosphate-buffered

saline (PBS) (pH 7,4) + 1% (v/v) Triton- X-100 + protease inhibitors and pelleted by centrifugation (2500xg) at 4°C. The agarose beads and 25 µL aliquots of each discarded wash and of the lysates were boiled for 5 minutes with Laemmli sample buffer and loaded onto 10% polyacrylamide gel and proteins were resolved by SDS-PAGE. Proteins were transferred from gels onto PVDF membranes (BioRAD, CA) by semi-dry electroporation (Hoefer SemiPhor, Pharmacia Biotech, San Francisco, CA, USA). The membranes were blocked in Tris-buffered saline with 0.2% (v/v) Tween-20 (TBST) + 5% (wt/volume) milk for 16 hrs. The blocked membranes were incubated for 2 hrs at room temperature with the primary antibodies: anti-Rluc antibodies (Mab4410 *versus* Rluc-F[1], Mab4400 *versus* Rluc-F[2], (Chemicon, Temecula, CA)), anti-dihydrofolate reductase antibodies (D1067 *versus* DHFR-F[1,2], D0942 *versus* DHFR-F[3] (Sigma-Aldrich, St-Louis, MO)) and anti-GFP antibody (A6455 *versus* both "Venus" YFP-F[1] and "Venus" YFP-F[2]); (Molecular probe, Eugene, OR, USA), 11814460001 *versus* "Venus" YFP-F[2]; (Roche, Basel, Switzerland)). The membranes were washed 3 times with TBST and incubated for 90 minutes at room temperature with the appropriate secondary antibody coupled to horseradish peroxidase: anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (#7074; Cell Signaling technology, Danvers, MA, USA) for primary antibodies raised in rabbit (D1067, D0942, A6455) and antimouse horseradish peroxidase (HRP) conjugated antibody (#7076, Cell Signaling technology, Danvers, MA, USA) for primary antibodies raised in mouse (Mab4400, Mab4410 and 11814460001), and washed with

TBST. The membranes were revealed with the western lightning chemiluminescence reagent plus (PerkinElmer, USA) substrate of HRP with Kodak Biomax XAR film over 30 seconds and 20 minutes exposure. Comparison of the three PCAs (Fig. 6B) showed that Bcy1 remained bound to cAMP resin while Tpk2 is found in the supernatant for both Rluc and DHFR PCAs whereas both the Bcy1 and Tpk2 subunits remained associated with cAMP resin in the case of the "Venus" YFP PCA. Thus, in contrast to the irreversible Venus YFP PCA and as already demonstrated for the Rluc PCA [21], the DHFR PCA is reversible.

1.5.2. Creation of universal DHFR PCA fragment templates and creation of homologous recombination cassettes.

We set out to create two universal oligonucleotide cassettes encoding each complementary DHFR PCA fragment and two unique antibiotic resistance enzymes to allow for selection of haploid strains that have been successfully transformed and recombined with one or the other homologous recombination cassettes. The universal cassettes were constructed in three steps as follows: First, the ADH terminator (ADHterm) was PCR amplified from *S. cerevisiae* genomic DNA with a forward primer containing a BamHI restriction site prior to a 5'XbaI restriction site and a reverse primer containing a BglII 3' restriction site using the high fidelity Accuprime Pfx DNA polymerase (Invitrogen, Carlsbad, California). The BamHI/ BglII digested PCR product was subcloned into the multicloning site of the pAG25 and pAG32 plasmids (henceforth called pAG25ADHterm and pAG32 ADHterm) [58]. Second, the DHFR PCA N-

terminal (F[1,2]) and C-terminal (F[3]) fragments, each proceeded 5' by a sequence coding for a 10 amino acid (Gly.Gly.Gly.Gly.Ser)2 flexible polypeptide linker (henceforth referred to as "linker") were PCR amplified from pMT3 mammalian expression vectors harboring these constructs [24]. These PCRs were performed with a forward primer containing a 5' BamHI restriction site and reverse primer containing a 3' XbaI restriction site. The linker- F[1,2] was subcloned into pAG25ADHterm between BamHI and XbaI restriction sites 3' to the ADHterm sequence, creating the pAG25-linker-F[1,2]-ADHterm cassette. The linker-F[3] was subcloned into the multicloning site of pAG32ADHterm between BamHI and XbaI restriction sites creating the pAG32 linker-F[3]-ADHterm. Each of the plasmids used for subcloning already contained unique antibiotic resistance cassettes that in the resulting constructs are 3' to the ADHterm. Thus, the final DHFR PCA F[1,2] universal template consists of pAG25-linker- F[1,2]-ADHterm followed by TEF promoter, nourseothricin N-acetyl-transferase (NAT1) that confers resistance to nourseothricin and finally a TEF terminator. The final DHFR PCA F[3] universal template is the pAG32 linker-F[3]-ADHterm followed by TEF promoter, hygromycin B phosphotransferase that confers resistance to hygromycin B and TEF terminator. The resulting universal templates were used to create homologous recombination cassettes for each of 5,756 budding yeast genes by PCR using 5' and 3' oligonucleotides consisting of 40-nucleotide sequences homologous to the 3' end of each ORF (prior to the Stop codon) and a region approximately 20 nucleotides from the stop codon. Design of

recombination cassette and diagnostic PCR primers are described below in section 1.5.3.

1.5.3. Oligonucleotide design and synthesis.

The oligonucleotides used for the 3'-tagging of each ORF with recombinant DHFR PCA cassettes and the oligonucleotides used to perform diagnostic confirmation of successful transformations were designed as follows and are available in tables S8 and S9. Coding and downstream sequences of yeast ORFs were downloaded from the Saccharomyces Genome Database (SGD) (http://www.yeastgenome.org/) on March, 2004. Homologous regions for the forward oligonucelotides were 40 nucleotides long and their sequences corresponded to a sequence 5' to the ORF stop codon. The following sequence was added to the end of this homologous sequence: GGCGGTGGCGGATCAGGAGGC, which anneals to the 3' end of the TEF terminator region of antibiotic resistance cassettes that are 3' to linker-F[1,2]-ADHterm in pAG25 and linker-F[3]-ADHterm in pAG32 as described in section 1.5.2, creating a specific 61 bp PCR oligonucleotide for each of 5,756 ORFs. Homologous regions for the reverse oligonucleotides were 40 nucleotides long and contained the sequence of genomic DNA immediately 3' to the stop codon of each ORF. The following sequence was added 5': TTCGACACTGGATGGCGGCGTTAG, which anneals to the sequence of the linker in both linker-F[1,2]-ADHterm in pAG25 and linker-F[3]-ADHterm in pAG32, creating a 64 bp oligonucleotide for each of 5,756 ORFs. Diagnostic PCR oligonucleotides were designed to correspond to sequences of the non-

coding strand of DNA in a region from 100 to 1,000 bp downstream of each ORF. Using custom made Perl scripts, these regions were searched for short sequences (18-25 nucleotides) with the following properties: have a melting temperatures from 58 to 62 degrees, have high GC content (> 50%), have a G or C at the 3' end for better annealing of the 3' end and optimized primer extension, have no complementary ends and do not contain palindromes. The forward diagnostic oligonucleotide consisted of a region common to both homologous recombination cassettes within the terminator of the antibiotic resistance marker. The oligonucleotides were custom synthesized (IDT, Coralville, IA, USA). Primer mass was determined by mass spectrometry for quality control of the synthesis. Primers were delivered lyophilized on 15, 384-well format microtitre plates. The oligonucleotides were resuspended in sterile distilled water at a concentration of 133 μM in their indexed well positions and stored at -20°C. An aliquot of the forward and the reverse primers for generating each ORF-specific homologous recombination cassette were mixed in a 60 μl volume to final concentration of 5 μM of each primer. The primer mixtures were used for PCR amplification of the cassette and remaining volume was kept at -20°C.

1.5.4. Creation of homologous recombination cassettes

The linker-F[1,2]-ADHterm-TEFpromoter-NAT1-TEFterm in pAG25 and linker-F[3]-ADHterm- TEFpromoter-HPH-TEFterm in pAG32 were PCR amplified with the ORF-specific oligonucleotides (describe in section 1.5.3) to create the specific homologous recombination cassettes in assigned positions of

384-well PCR plates. The 30 PCR cycles were carried out with the high fidelity Accuprime Pfx DNA polymerase, with an annealing temperature of 59°C for 30 seconds, a 3-minute elongation at 68°C, in a 25 µl total volume reaction. 3 µl aliquots of 40 randomly chosen PCR products on each 384-well plate were mixed with brilliant blue tainted glycerol, resolved by gel electrophoresis of the mixture on a 1% agarose gel stained with ethidium bromide and visualized under an UV light for quality control of the PCR amplification. The PCR products were directly used for creation of recombinant strains and the remaining reaction volume was stored at -20°C. Final homologous recombination cassettes are referred to henceforth as F[1,2]- NAT1 and F[3]- HPH.

1.5.5. Creation of recombinant strains

The strains BY4741 (*MAT***a** *his3Δ leu2Δ met15Δ ura3Δ*) and BY4742 (*MAT*α *his3Δ leu2Δ lys2Δ ura3Δ*) were transformed as described in [59] with PCR products amplified from the templates described above to create specific homologous recombination DNA fragments (in section 1.5.4). The protocol was adapted to large-scale transformation in 96 well plates as follows: 8 µl of PCR product was mixed with 10 µl of chemically-competent yeast and mixed with 72 µl polyethyleneglycol (PEG) buffer and incubated for 30 minutes at room temperature. Cells were then heat shocked for 15 minutes at 42°C in a water bath. The transformation buffer was replaced with 250 µl YPD and cells were left to recover for 4 hours at 30°C and were then plated on antibiotic containing YPD agar plates and allowed to grow for 4 days at 30°C. In all cases

the BY4741 (*MAT***a**) strain was transformed with the F[1,2]-NAT1 cassettes and BY4742 (*MAT*α) with the F[3]-HPH cassettes. The transformed strains were grown on YPD agar plates plus appropriate antibiotic agar plates (100 µg/mL nourseothricin for *MAT***a** transformed strains (WERNER BioAgents, Jena, Germany) or 250 µg/mL hygromycin B for *MAT*α transformed strains (Wisent Corporation, Quebec, Canada). Identification of successfully recombinant clones was performed as follows: Putative recombinant clones were picked by hand, cell lysis was performed by heat treatment and confirmation of the correct location of a genome insertion was determined by PCR using the diagnostic primers described above (in section 1.5.3). The yeast lysates were placed at indexed positions of the 96-well PCR plates and mixed with the ORF specific and cassette specific diagnostic primers. Annealing of the diagnostic primers was performed at 56°C and 35 PCR cycles were carried out with the regular Taq DNA polymerase, with elongation cycles of 1.5 minutes at 72°C, in a 50 µl total reaction volume. Aliquots of 30 µl from each PCR product position within each 96-well plate were mixed with brilliant blue tainted glycerol and gel electrophoresed on ethidium bromide stained 1% agarose gels. Correct recombination was confirmed based on sizes of PCR products. The success rate for obtaining positive recombinants ranged between 30 and 90% of diagnosed colonies per 96-well plate. This process was repeated for up to six rounds of recombinant colony selection if recombinants were not found in the first round. Our efforts resulted in successful creation of 4,326 *MAT***a** strains harboring unique ORF F[1,2]-NAT1 fusions and 4,804 *MAT*α

harboring unique ORF-F[3]-HPH fusions. The confirmed strains were glycerol-stocked in pre-assigned positions of 96-well plates and stored at -80°C with a total of 120 plates.

1.5.6. Optimization of DHFR PCA screening conditions

We optimized the DHFR PCA by selecting a subset of 380 *MAT***a** strains with ORFs tagged with F[1,2] mated to 380 *MAT*α strains with ORFs tagged with F[3], for a total of 145,000 crosses (in section 1.5.5). These were selected based on the knowledge that the protein products of each ORF expressed as a fusion to F[1,2] should interact with the protein product of at least one ORF expressed as fusion to F[3] when complementary *MAT***a** and *MAT*α strains are mated and resulting diploids selected for growth in the presence of methotrexate (positive reconstitution of the methotrexate-insensitive DHFR PCA reporter). The known protein-protein interactions were obtained from the hand-curated MIPS database of protein complexes (http://mips.gsf.de/genre/proj/yeast/) as of 21 July 2005. We used this array to optimize the protocols for large scale, solid phase mating; selection of a methotrexate concentration that assures minimal background growth of negative control strains (in section 1.5.8) and maximizes growth of DHFR PCA-rescued diploid strains at different temperatures while minimizing the incubation time. Briefly, methotrexate concentrations were tested between 25 µg/mL and 225 µg/mL (in 25 µg increments) with an optimal concentration found at ~ 200 µg/mL. While higher concentrations provided more stringent selection, we found that methotrexate had limited solubility above this concentration. For all concentrations of methotrexate, we found that

the optimal incubation time was 96 hours of growth at 30°C, which were then the conditions we used for the large-scale screen.

1.5.7. Test for detection of structural and topological organization of a protein complex

To test for the capacity of the DHFR PCA to provide structural and topological information, we performed a screen for interactions among the subunits of the well-characterized RNA polymerase II (RNA Pol II) complex structure, which has been determined at high resolution (PDB file 1I3Q) [21]. We constructed homologous recombination cassettes for the 10 subunits (RPB2, RPB3, RPB5, RPB8, RPB9, RPB10, RPB11, RPC10, RPO21, RPO26) and created *MAT***a** and *MAT*α strains in which each of these genes are fused to the DHFR PCA fragments according to the methods described in Section 1.5.5. We used engineered heterodimerizing mutants of the SspB dimer from *Haemophilus influenzae* to serve as a positive control for the DHFR PCA and a negative control for interactions with RNA Pol II [17]. Expression plasmids harboring these control proteins fused to the DHFR PCA fragments were generated as follows: The heterodimerizing SspB mutants, SspBLSLA and SspBYGMF were amplified by PCR and subcloned into p413-TEF-Zip-linker-DHFR F[3] and p413-TEF-Zip-linkers-DHFR F[1,2], replacing the leucine Zipper (Zip) sequence. We replaced the existing His3 metabolic enzyme selection marker cassette in p413TEF with nourseothricin N-acetyl-transferase (NAT1) or hygromycin B phosphotransferase (HPH) antibiotic selection genes. Specifically the cassette encoding (NAT1) from pAG25 was PCR amplified

and subcloned between the sequence corresponding to the His3 complementing cassette of p413-TEF-SspBLSLA-linkers-DHFR F[1,2]. These plasmids were renamed p41NAT. The cassette encoding HPH gene from pAG32 was PCR amplified and subcloned between the sequence corresponding to the His3 gene cassette of p413-TEFSspBYGMF- linker-DHFR F[3]. These plasmids were renamed p41HPH. Final products were transformed into *MAT***a** or *MAT*α (p41NAT-TEF-SSpBLSLA-linkers-DHFR F[1,2] and p41HPHTEF-SspBYGMF-linker-DHFR F[3]) and respectively selected on YPD (+100 μg/mL nourseothricin) or YPD (+250 μg/mL hygromycin B). The ten *MAT***a** strains harboring genomic fusions of DHFR F[1,2] fused to the ten RNA Pol II subunit genes or the control *MAT***a** expressing the SspBLSLA mutant fused to F[1,2] were mated in a one-to-one matrix with the ten complementary *MAT*α strains harboring genomic fusions of DHFR F[3] fused to the ten RNA Pol II subunit genes or the *MAT*α control strain expressing SspBYGMF mutant fused to F[3]. The diploid strains were grown on synthetic complete medium (SC) (-met, -lys, +methotrexate 200 μg/mL) to select for growth of methotrexate resistant diploid strains and incubated at 30°C for 120 hours (Fig. 13). The plates were then photographed with a high-resolution (4 Mega pixel, Canon) digital camera and resulting images were converted to 8-bit grayscale images and analyzed with ImageJ 1.36b software (National Institutes of Health, USA), measuring pixel intensities of the colonies over a constant area. Results were corrected by subtraction of the pixel intensity of a region without colonies to remove any background intensity of the plate (Fig. 13A, B). Based on a t-test

for differences in growth, we detected 30 of the 50 possible interactions among the RNA Pol II subunits (60%), including homodimers while the control heterodimerizing mutants of SspB fused to mDHFR fragments showed no significant interactions with any of the ten RNA Pol II subunits (Fig. 13C). The detected interactions are consistent with observations of physical proximity inside the RNA polymerase II complex (Fig. 13D and Fig. 14).

1.5.8. Large scale PPI screen

To test pairwise protein-protein interactions between the two collections of *MAT***a** and *MAT*α strains, each of 3,247 *MAT***a** strains harboring ORF-F[1,2]-NAT1 fusions (henceforth the "baits") were individually mated with each of 4,804 *MAT*α strains harboring ORF-F[3]-HPH fusions (henceforth "preys") arrayed on agar plates as described below (Fig. 2). Diploids were subsequently selected and plated on medium containing methotrexate to select for positive DHFR PCA reconstitution. Although our screen is based on a selection assay and thus amenable to a split-pool screening strategy, we performed individual crosses to avoid the overrepresentation of highly abundant interacting proteins that would result from growth competition within pools of clones. The bait-by-query-array crosses were performed as follows: First, to generate the prey array, the *MAT*α query strains were printed from 96-well glycerol stock plates (in Section 1.5.5) to yeast peptone dextrose medium (YPD) agar plates with hygromycin B (+250 μg/mL hygromycin B) to a density of 384 colonies per plates using a 96 pin robotically manipulated pin tool (0.787 mm flat round-shaped pins, custom AFIX96FP3 BMP Multimek FP3N, V&P Scientific Inc.,

San Diego, CA) and allowed to grow for 24 hours at 30°C. The colonies were transferred to four YPD agar plates with hygromycin B (+250 μg/mL hygromycin B) using a 384 pin robotically manipulated pin tool (0.356 mm flat round-shaped pins, custom AFIX384FP8 BMP Multimek FP8N, V&P Scientific Inc., San Diego, CA) to a density of 1,536 spots/plate and allowed to grow for 24 hours at 30°C. Subsequently, colonies from the four *MAT*α plates were transferred onto individual YPD agar plates using a robotically manipulated 1,536 pin tool (0.229 mm flat roundshaped pins, custom AFIX1536FP9 BMP Multimek FP9N, V&P Scientific Inc., San Diego, CA), necessary for the mating procedure with individual *MAT***a** bait strains, resulting in *MAT*α query strain arrays at a density of 6,244 pin transfers per plate in which the 4,804 *MAT*α query strains form colonies at defined positions. In addition, positive and negative controls for correct transfer of cells and selection conditions were introduced at 48 positions on the array in cross-shaped patterns that traverse the entire array (See Fig. 5B) The positive control strains consisted of 24 *MAT***a**/*MAT*α diploids harboring DHFR PCA fusions of the MCK1 and CDC19 proteins that are known to interact and for which the interactions were confirmed by the DHFR PCA (MCK1-F[3] and CDC19-F[1,2]) and 24 negative control *MAT***a**/*MAT*α diploids for the pair of proteins CLN3 and CDC19, which have not been shown to interact by any method and show no growth in the DHFR PCA (CLN3-F[3] and CDC19-F[1,2]). Remaining positions on the 6,244-position array were empty; no cells having been transferred to these positions. For each of the bait-by-prey array crosses

(Fig. 1), individual *MAT***a** bait strains were grown separately in liquid YPD with nourseothricin (+100 µg/mL nourseothricin) to saturation. The *MAT***a** bait strains saturated culture were concentrated by centrifugation and removal of approximately 60% of the supernatant. The *MAT***a** bait strain was printed on top of previously printed *MAT*α query strains on a YPD containing plates, using a robotically manipulated 1,536 pin tool (0.229 mm flat round-shaped pins, custom AFIX1536FP9 BMP Multimek FP9N, V&P Scientific Inc., San Diego, CA), and incubated at 30°C for 24 hours to allow for mating. The mated colonies were selected the next day by transferring all colonies onto YPD agar (+ 100 µg/mL nourseothricin, +250 µg/mL hygromycin B). The selected *MAT***a**/*MAT*α diploids were transferred the next day onto synthetic complete medium (SC) (-met, -lys, +methotrexate 200 µg/mL) to select for growth of methotrexate resistant diploid strains and incubated at 30°C for 90 hours. Individual plates were then photographed with a high-resolution (4 Mega pixel, Canon) digital camera and resulting images were analyzed and growing colonies quantitated using a shape recognition algorithm (in Section 1.5.9). The diploid strains created were heterozygous for the two tagged genes, thus minimizing growth defects that may result from reduced activity or abundance of the modified gene products and reducing the potential for non-specific interactions among highly expressed proteins. Bait screens for which no significant growth above background was observed could be due to low expression of the proteins under the growth conditions, to interference of the mDHFR fragments with the folding or stability of bait and prey proteins, or

simply to the absence of expression or interaction of the bait proteins with any of the query proteins under the conditions tested

1.5.9. Control experiments for spontaneous, interaction-independent DHFR protein fragment complementation.

Large-scale PPI screens often contain a large fraction of false-positive interactions, either because of experimental noise or from artifacts specific to the detection method. The latter is the most problematic because false-positive interactions are often reproducible and have strong signal to background ratios. This problem therefore must be dealt with experimentally. Previous large-scale PPI studies (e.g., [13, 14, 60]) indeed reported that a number of proteins show promiscuous patterns of interactions. These interactions are typically arbitrarily eliminated from filtered datasets. The source of spurious results in PCA could be due to spontaneous complementation (folding) of the DHFR PCA fragments into active enzyme in the absence of any physical interaction between bait and prey proteins. Specifically, some proteins fused to a PCA fragment, (e.g. F[1,2]) could complement the other PCA fragment, F[3], without necessarily interacting with the protein to which this second fragment is fused, i.e. interacting with the F[3] fragment alone. In a screen such as ours, these proteins could promiscouously interact with many proteins and thus produce a large fraction of false positive interactions. To test for this possibility, we constructed a set of control baits that consist of the 10-amino acid linker (Gly.Gly.Gly.Gly.Ser)2 fused to each of the complementary F[1,2] or F[3] DHFR fragments or F[1,2] or F[3] fragments alone. These control baits were

used to challenge the prey array and a newly constructed *MAT***a** F[1,2] array created with the same procedure described in Section 1.5.7. The fragments have to be expressed from plasmids that harbor both linker-fragment fusions or fragments alone and contain resistance cassettes compatible with our screening strategy, transformed into complementary strains and mated against the individual arrays as described in section 1.5.8. The control constructs were generated by PCR amplifying linker-fragment fusions or fragments alone from the universal PCA template constructs (in Section 1.5.2) constructed in pAG25 (linker-F[1,2]-ADHterm) and pAG32 (linker-F[3]-ADHterm). We used forward primers containing a 5' XbaI restriction site upstream of sequence coding for the linker sequence or individual DHFR PCA fragments and reverse primers coding for the 3' region of the individual fragments upstream of a 3' XhoI restriction site-coding sequence. These were used to subclone the XbaI- and XhoI-digested PCR products into the expression vector p413TEF between the XbaI and XhoI of the multicloning site 3' of the constitutive TEF promoter. The existing His2 metabolic enzyme section marker cassette in p413TEF needed to be replaced with NAT1 and HPH antibiotic selection genes to be consistent with our screening strategy as follows: The cassette encoding nourseothricin N-acetyl-transferase from pAG25 [58] flanked by the TEF promoter in 5' and terminator 3' was amplified with TEF promoter forward primer and TEF terminator reverse primer. These primers insert an NdeI restriction site into the 5' region of the PCR product and a NsiI restriction site in the 3' region. The resistance cassette was subcloned into the control

expression plasmids between the NdeI restriction and NsiI restriction sites that flank the His3 complementing cassette. These plasmids are henceforth referred to as p41NAT for the NAT1 gene, conferring resistance to the antibiotic nourseothricin of transformed yeast. The control fragment plasmids expressing a F[3] fusion were further digested with BstXI and BsmI to subclone the HPH gene from pAG32 [58], located between BstXI and BsmI, encoding hygromycin B phosphotransferase, which confers resistance to the antibiotic hygromycin B of transformed yeast. The expression of the control fragments was confirmed by Western blot using a rabbit anti- DHFR polyclonal antibody that specifically recognizes an epitope in the N-terminal F[1,2] fragment (Sigma D1067, 1:6000 (Sigma-Aldrich, St-Louis, MO)) F[1,2] or an anti-DHFR polyclonal antibody that specifically recognizes an epitope in the C-terminal F[3] fragment (Sigma D0942, 1:5000 (Sigma-Aldrich, St-Louis, MO)) (Data not shown).

1.5.10. Data acquisition, colony quantification and documentation, and statistical analyses.

Positive protein-protein interactions were interpreted from growth of the diploid colonies on methotrexate-containing medium as described above (in Section 1.5.7). Complete acquisition and analysis of each plate proceeded as follows: First, images of the diploid methotrexate selection (*MAT***a** bait/*MAT*α prey) array plates were taken after 96 hours of growth (as described above, Section 1.5.7) with a 4.0 Mega pixel camera (Powershot A520, Canon). Plate images were saved in JPG format at a resolution of 180 dpi and a size of 2,272 x1,704

pixels. Second, all of the 3,301 plates (3,247 plates for the 3,247 different baits, 48 repeated plates and 6 plates for the control experiments) were manually inspected during the image analysis step and positions too close to the plastic edges of the plate and therefore uninterpretable were eliminated by setting the intensities of the first and the last colonies on the first row of the array to 0. At this stage we eliminated 44 plates from the final analysis because they displayed growth of colonies at empty positions, likely resulting from grid misalignment or contamination. In order to accurately identify bait ORF/prey ORF coordinates on the array and to extract the intensity of the colonies, we developed image recognition routines from the Image Processing Toolbox of Matlab (The MathWorks, Natick, Massachusetts). Each plate contained a set of positive and negative control colonies as described in Section 1.5.7, arranged in parallel X-patterns across the plates (Fig. 1A), which allowed us to identify misalignment or mispositioning of the colony arrays on each plate due to variation in the robotic pinning process. The first step of the image analysis was to determine expected centers of the colonies arrayed in 96 columns and 64 rows. In order to adjust for variation in plating and possible rotation of an image during image acquisition we manually defined the coordinate center of a first and a last colony in a first row of the array and of the last colony of the first column using Matlab build-in tools for accepting user input. Using the coordinates of these centers the image was rotated so that each column and row of colonies lies on a straight line. Next, images were adjusted using the same parameters and centers of all colonies were calculated. For each subsequent

plate, manual reselection of colony centers was performed for colonies or positions at the extremities of a plate if array coordinate position centers deviated significantly from those of an earlier plate. Image analysis was performed in the same order as the images were acquired and thus deviations in the positioning of plate images occurred among groups of plates and were easily identified. However, we usually found that series of images taken on the same day (~100 plates per day were processed) did not have to be readjusted. Results of positional array adjustment and detection of colony centers were manually checked for all images taken during the large-scale screening (3,307 plates in 6,144 position format plus spontaneous complementation control experiments (in Section 1.5.8) and repeats of 48 bait/prey repeat plates. While the procedure described above defines approximate colony centers, we used a modified version of a previously described algorithm to accurately locate them [61]. The assumption of the algorithm is that pixels have higher intensity values at colony centers. Thus, the algorithm processes columns and rows of pixels one at a time and finds areas of maximal pixel intensity. Intersection of columns and rows where corresponding sets of pixels have the highest 75th percentile values are used to find approximate colony centers. We modified this algorithm in order to handle an array of 6,144 colonies per plate (96 well format was used in the original work). Next, we used an automated Matlab procedure to calculate colony areas and intensities on each plate as follows (Fig. 2) (please refer to Matlab documentation for detailed description of functions used for the analysis): 1), images were corrected for possible non-uniform illumination as

described in (http:// www.mathworks.com / products / image / demos.html?file= / products / demos / shipping / images / ipexrice.html) and 2), small objects that correspond to gel background, bubbles, plate edges or other anomalies were removed using the imopen function with the disk morphological structuring excluding elements of a radius smaller than 2 pixels (radius of 4 pixels was used on plate edges), 3) images were converted into binary format using the im2bw function with a threshold calculated by the graythresh function. In this format, pixels that correspond to colonies were set to 1 and background pixels to 0. Thus, connected components of binary images with pixel values equal to 1 (calculated with bwlabel and regionprops functions) corresponded to colonies. Identity (row and column number on the plate array) of each colony was calculated by comparing approximate colony centers (calculated as described above) and centers of these connected regions. Intensity of each colony was then extracted from the original image. If there was no overlap between colonies (i.e. there were no colonies that touched each other) these 3 steps were sufficient for the analysis. However, in a small number of cases, large colonies overlapped between adjacent positions and had to be deconvoluted. One of the simplest ways of discriminating intersecting circular objects is to calculate the distance transform for a binary image followed by a watershed transform. The watershed transform finds "catchment basins" (which represent circular colonies) and "watershed ridge lines" (which correspond to an edge where colonies touch each other) in an image by treating it as a surface (see Image Analysis watershed segmentation demo for details

http://www.mathworks.com/company/newsletters/news_notes/win02/watershed.html). As a result of the watershed transform, pixels that lie on the border between objects can be identified. These border pixels are set to zero on the original binary image, thus separating overlapping colonies and step 3 is repeated in order to analyze separated colonies. If the size of a connected region identified in step 3 was larger than expected for a single colony (on average, 400 pixels) or a connected region could not be matched to a position on the array (fusion of several colonies will result in centering of a connected region to deviate from an approximate colony center) the following steps were performed: 4), extraction of a rectangular subpart of an original image that fully contains a connected region that is suspected to contain fused colonies, 5), apply a distance transform. The distance transform calculates for each pixel its distance from the nearest nonzero pixel in a binary image. In the case of two intersecting circular objects it will produce a set of values with a maximum value at the centers of the objects. We found that calculating an accumulation array by the Circular Hough transform and superimposing its local maximum on a distance transform array further improved the detection of circular colonies. Hough transform is another method for detecting circular objects. It's a voting procedure that assigns a value to a pixel if it can be a center of a circle. As a result, central pixels in circular objects receive higher values. So step 6), superimpose local maxima of the accumulation array of Circular Hough transform onto a matrix produced by step 5. 7), apply the watershed transform. 8), in cases where a number of objects that were separated using steps 4 to 5

was different from the number of possible centers detected by the Circular Hough transform, we performed a watershed transform on the original grey scale image that corresponds to the connected regions requiring deconvolution. Raw colony intensities are available on the Michnick Lab website (http://michnick.bcm.umontreal.ca/).

1.5.11. Analysis of colony intensity distributions and benchmarking.

Raw colony intensity (sum of pixel values from a grayscale) distribution was approximately lognormal and centered around 4,300 (Fig. 4). This distribution showed a steep decrease in frequency at around 10,000, after which it showed an increase in frequency. This second distribution represents the population of diploid colonies able to grow on methotrexate. This is also where we saw a clear distinction between the positive (median: 38,361) and negative control intensity (median = 5,192) distributions. We therefore reasoned that the threshold above which we could infer a protein-protein interaction (PPI) should be located at around an intensity of 20,000 (see below). It is important to note that controls were diploid strains printed directly onto the plates and that these control strains did not go through the solid phase mating procedures. This explains why several colonies were smaller than the negative controls. These likely represented variability in diploid cell transfers. In order to determine the accuracy of our data, we first had to eliminate data for ORFs that showed interactions with control fragments (spontaneous fragment complementation; in Section 1.5.8). To do this, we calculated z-scores for control plates and used cutoffs for accepting an interaction based on a visual inspection of distributions

of plate scores. Cutoffs were conservatively assigned based on visual inspection of the control plate distribution by identifying all colonies with larger intensities than the background distribution. We identified 344 such proteins (table S1), which are enriched for those associated with ribosomes ($P < 2x10^{-65}$) (table S2) and are highly expressed compared to the proteome (median log10(Abundsd) = 3.27 vs 2.28 for the proteome, Wilcoxonrank test: $P < 2.2x10^{-16}$; Fig. 11). Similar highly abundant proteins are also often observed as false-positives in spurious interactions in affinity purifications in particular several ribosomal proteins and others like Cdc19p, Eno2p, Tef2p and Tef3p [14]. Global analysis of the network created by these proteins revealed that they show highly structured and similar patterns of PPIs, suggesting that they could also be identified by computational analysis. Proteins that show correlated patterns of interaction with that of the controls but that were not identified in this control experiment could then be eliminated as well (see below). After eliminating positions from the plates corresponding to these proteins, we determined the threshold above which we could confidently infer PPIs as follows: First, we used the MIPS catalog of protein complexes (ftp://ftpmips.gsf.de/yeast/catalogues/complexcat/complexcat_data_18052006) as of 18 May, 2006 as a source of True-positive (TP) interactions (n = 11,005 among 1,236 proteins), 503 of which could be potentially detected in our screen (bait and prey strains exist in our collection for each protein and crossing of each strain resulted in growth of at least a colony, suggesting that the DHFR fragment-tagged ORFs are expressed and can interact with other proteins). This

set of manually annotated MIPS complexes serves as a benchmarking standard for networks enriched for soluble proteins such as TAP-MS [36] and also in studies of helical membrane protein PINs (n = 79) [61]. A true-negative (TN) set of PPI was obtained from [36] and consists of pairs of proteins that are part of distinct complexes and are expressed in different cellular compartments or have anticorrelated patterns of gene expression. This set contains 266,858 interactions, 6,377 of which could potentially be detected in our screen as described for the TPs. These two sets of PPIs allowed us to determine at what colony intensity we could confidently infer a PPI. In order to control for plate-to-plate variation in overall growth intensity and the non-random distribution of the number of interactions among plates, we combined two criteria [62] to determine a threshold above which to call the growth of a colony: The first criterion was the absolute intensity (a sum of pixel intensities on a grey-scale) of the colonies and the second, a score derived from the distribution of intensities on the plates:

$$\text{z-score}(x) = (x-\mu_x)/\sigma_x$$

where x is the intensity of a given colony, μ_x is the average intensity of the plate and σ_x is the standard-deviation of the mean. This allowed us to eliminate colonies that had high intensity values due to the high background growth of colonies on some plates. This combination allowed us to maximize coverage of true-positive PPI at a high Positive Predictive Value (PPV) comparable to small-scale experiments, where PPV is defined as the ratio of inferred True Positive interactions over the sum of the inferred True and False Positives. This

was achieved with a pixel intensity of 23,000 and a z-score = 2.4. After removing false-positive interactions based on the control experiments (n = 344) and benchmarking on the MIPS gold standard, our final, high quality dataset includes data that reached a PPV score of 97.7%, implying that 97.7% of these interactions are predicted true-positives based on this high quality data. These cutoff results in data having precision comparable to all previous large-scale data sets including those that reaches the same precision as small-scale protein interaction studies (Fig. 1C, Fig. 8). At this cutoff and PPV score of 97.7% we observed 5,672 interactions**.** A further analysis revealed 83 highly connected proteins that mediate 2,902 interactions. Eight of these proteins constantly demonstrated a higher growth pattern than other *MAT*α strains (type 1). The remaining 75 proteins showed an unusually large number of common interactions. 23 proteins had a similar pattern of interactions to proteins that showed growth in the negative control experiments (type 2). Some of them, for example ARO8 and VAS1, were not among the 344 proteins identified as interacting with controls because their growth was just slightly below one of the two threshold cutoff values used to analyze control plates. The remaining 52 proteins showed a distinct pattern of common interactions (1,830 interactions) also typical of the 344 proteins that showed positive growing colonies in the negative control experiments. Given the ambiguity or control-like behavior of these results, all 2,902 interactions were removed from the final dataset. Our final, filtered dataset contained only 3 FP and 163 TP interactions. To determine the statistical significance of these results, we generated 10,000

random networks by randomly assigning the same number of interactions between the same set of proteins as in our final dataset to estimate the distribution of random scores, and a z-score for the observed score from this distribution. After applying a multi-step filtering procedure, the DHFR PCA network was derived from 24% of these bait screens. On average, a random network had 2.7 TP and 33.8 FP interactions. Our observed results are thus unlikely to have been found by chance alone (TP : $P = 0$, FP = $P < 10^{-7}$) and do not reflect a bias in the composition of our network with respect to the TP and TN data sets used here. A total of 3,113 colonies were found above the PPV threshold of 98.2%, which represents 2,770 unique interactions after subtracting interactions observed in both *MAT***a** bait (ORF1-F[1,2])/*MATα* prey (ORF2-F[3]) and *MAT***a** (ORF2-F[1,2])/*MATα* prey (ORF1-F[3]) (tables S3 & S4). 40% of the interactions that could have been detected in reversed crosses were detected, consistent with previous observations showing that interactions may be detected by the DHFR PCA when the individual proteins are fused to either F[1,2] or F[3] fragments, but not necessarily if the fragments are swapped between the two proteins [63]. This proportion was 50% in the RNA pol II small-scale experiment (Section 1.5.7), which suggests that we are near the upper bound with the large-scale screen. However, interactions that were observed in both directions (n = 343) can be considered as being of higher confidence given that their intensity scores on the plates is slightly higher on average (51, 216 versus 38,995), despite the fact that these two sets of proteins have only marginally different levels of protein abundances (median

log10(Abund): 2.38 vs 2.32, $P = 0.01$). This is also reflected in the similarity of their Gene Ontologies. Pairs of proteins that were shown twice to interact during the screen have higher semantic similarity on average in terms of Biological Process, Molecular Function and Cellular Compartment than the same number of interactions taken randomly from the interactions detected in one direction only (BP : 3.45, BP random : 2.65, $P < 10^{-12}$; MF : 3.88, MF random : 3.12, $P < 10^{-5}$; CC: 2.76, CC random: 2.21 ; $P < 10^{-7}$). PCA is efficiently able to detect interactions among membrane proteins with high specificity [64, 65], which contrasts with methods that have comprehensively studied protein interactions [10, 11, 13, 14, 35]. About one quarter of all genes in most genomes contain putative transmembrane (TM) helices [61]. They therefore likely represent an important fraction of the interactome and yet we have little knowledge of their patterns of interactions. The DHFR PCA network therefore represents important steps in that direction. As a consequence of the difficulty to identify protein-protein interactions among membrane proteins, these are not well represented in public databases. Two papers established genome-wide protein interaction networks for membrane proteins: 1) Miller *et al. [26]* using the split-ubiquitin assay, and 2) Xia *et al. [61]*, using computational prediction. These studies both acknowledged the challenge that the identification of protein-protein interactions among membrane proteins represents. Miller et al. identified 1,985 putative interactions among 536 proteins. Xia *et al.* predicted 4,145 interactions among 1,048 putative helical membrane proteins that they identified using computational methods. As

membrane proteins are less represented in high-quality datasets such as the MIPS (6% of proteins membrane proteins and 1% of interactions are among membrane proteins) which may limit our ability to identify false-positive interactions, we took special care in trying to identify potential spurious interactions involving membrane proteins. First, the control experiment with the fragments and the linkers alone revealed 19 (out of 344) putative helical membrane proteins that show spontaneous fragment complementation. Further, we eliminate 83 proteins that showed similar patterns of interactions, which is a characteristic of proteins showing spontaneous complementation with other proteins. This analysis further identified 45 membrane proteins that are likely to mediate spontaneous complementation of the DFHR fragments and these were removed from the dataset. In the final network, 232 of the putative membrane proteins identified by [61] show interactions in our PCA screen out of 1,124 proteins, which represents a slight enrichment compared to the entire genome. We identified 2,770 high quality interactions among 1,124 proteins in our screen. The average degree of a protein (number of interactions) is therefore ~ 1.6. Our network contains 232 helical membrane proteins that make 662 protein interactions among them, for an average degree of ~2.8. This is not much higher than for the network as a whole. This may represent an increased power of the PCA to detect pairwise interactions among 15 membrane proteins, which are less spatially constrained than proteins that are member of large complexes. In order to assess the quality of the fraction of the DHFR PCA network that represents the among helical membrane protein interactions, we examined the

overlap between our final dataset and that of the two previous membrane protein PINs. First, we found that we detected 51 of the 662 interactions predicted by [61]. Randomizations (10,000) of our network reveals that we expect an overlap of only 5 interactions between these two datasets by chance alone, which represents an enrichment of 10 fold over the random expectation ($P < 10^{-94}$). Similarly, 27 of the our 662 interactions were also identified using the split-ubiquitin method [26]. Only 1.9 interactions are expected to be in common between these two datasets, which represents an enrichment of 15 fold over the random expectation ($P < 10^{-75}$). Finally, we found that while controlling for the sharing of cellular compartment of interacting protein pairs, interacting protein pairs remain significantly enriched for semantic scores of Molecular Function and Biological processes (i.e. similarity of gene ontology categories), and this for the whole set of protein interactions (MF: 3.76, MF random: 2.48, $P < 10^{-74}$; BP: 3.34, BP random: 2.14, $P < 10^{-144}$) and for the set of helical membrane proteins interactions (MF: 3.89, MF random: 2.84, $P <10^{-18}$; BP: 2.86 , BP random: 2.15, $P < 10^{-15}$).

1.5.12. Analysis of high-quality PPI in comparison to protein abundance, gene ontology enrichment and three-dimensional structures.

1.5.12.1. Analysis of PPI versus protein abundance

Protein abundance for yeast grown under the same conditions (SC medium) as used in the DHFR PCA screen or YEPD and based on FACS analysis of GFP-tagged proteins was obtained from [37].

1.5.12.4. Analysis of Gene Ontology enrichment.

We examined whether the final set of proteins constituting the DHFR PCA network contained an overrepresentation of proteins involved in specific biological processes, with certain molecular functions or that localize to specific cell compartments**.** Gene Ontology enrichments were calculated using the method implemented in GOstat (R library) [66, 67]. This approach utilizes a conditional hypergeometric algorithm that considers the hierarchical relationships of gene ontology definitions to decorrelate the results. More precisely, due to their hierarchical organization—a GO term inherits all annotations from its more specific descendants. Gene Ontology categories are not exclusive and are thus locally correlated. This method considers this organization to limit the redundancies in the results. The algorithm considers more specific to more general terms. When testing the significant enrichment of a GO term, it removes genes that are annotated to a significantly enriched node from all its ancestors (more general term). The universe of genes (reference list to which the network proteins are compared) used for the comparisons was the entire genome, as we aimed at identifying what categories of proteins were overrepresented relative to the entire proteome. A conservative *P-value* cutoff of 0.00001 was used for these analyses.

1.5.12.5. Correlation of gene expression profiles.

Correlation of gene expression between pairs of ORFs (Pearson correlation) was calculated from gene expression profiles from more than 1,000 expression profiles compiled in [68]. The distribution of correlation coefficients for

random networks of the same size and degree distribution were estimated from 1,000 random networks.

1.5.12.4. Structural analysis of PPI

Structural domain annotations for all *S. cerevisiae* proteins were obtained from SGD in September 2007. The identity of domains known to mediate PPI were obtained from [69]. These data are derived from the mapping of protein domains onto 3D structures of resolved protein complexes deposited in the Protein Data Bank by [70, 71]. Domains mediating PPIs are those that are found at the interface of interacting proteins. The fraction of protein pairs in the DHFR PCA network that each have a domain that is known to mediate PPIs was estimated by combining these data with the domain identity obtained from SGD. The random expectation of the fraction of interacting proteins pairs that have one domain each that mediate PPI was estimated by generating random networks, as described above. Distances among C-termini of known protein complexes were determined as follows: The identity of yeast proteins that have homologs in the Protein Data Bank (PDB) was obtained from SGD. On June 2007, 46,320 PDB files corresponding to biological units of cellular organisms were obtained from PDB, among which 13,966 contained yeast homologs. The first biological unit was used (.pdb1) if more were available. We extracted the distances between the C-termini of all pairs of chains within each complex. We then went through each complex and determined whether: 1), they contain more than one chain homologous to any of the proteins that are part of the DHFR PCA network, 2), we recorded all of the possible pairs of interactions within

those complexes and 3), we merged all the interactions together and if a pair of chains was recorded more than once, we kept the one with the shortest distance between the C-temini of these two proteins. Finally, we examined whether this interaction was seen or not in the DHFR PCA screen. We considered interactions as not being seen only if it could have been detected, i.e. the pair of proteins showed at least one interaction as bait and prey or as prey and baits. This left us with 175 interactions from 129 distinct PDB entries.

1.5.11.6. Membrane topology and PPI determined by DHFR PCA.

Protein membrane topology was obtained from [39]. We reasoned that the C-termini of interacting proteins have to be oriented into the same cellular compartment in order for DHFR PCA to occur; for instance both in the cytosol or both in the lumen of the endoplasmic reticulum. Kim *et al. [39]* established the location of the C-termini of 546 proteins of which 448 have their C-termini oriented towards the cytosol. We examined all possible pairs of proteins among these 546 proteins and determined which of these are localized in at least one common cellular compartment using microscopic evidence for yeast proteins fused to green fluorescent protein [72]. It is important to note that we don't know if any of these pairs should actually show an interaction. For these possible interactions, we counted how many of each type we detected and did not detect, considering only those for which interactions could have been observed as described above (i.e. strains exist in which both proteins are fused to one or the other DHFR PCA fragments and show interactions).

1.5.13. Comparison of DHFR PCA to previously determined PPI.

In order to examine whether the DHFR PCA interactions had been reported previously, we obtained the following databases: Biogrid (www.thebiogrid.org/, version 2.0.29), mips-MPact (http://mips.gsf.de/genre/proj/mpact, version 18052006) and DIP (http://dip.doembi. ucla.edu/, June 2007). Entries in Biogrid with experimental systems defined as: "Synthetic Lethality" , "Dosage Rescue", "Synthetic Growth Defect", "Synthetic Rescue", "Epistatic MiniArray Profile", "Dosage Lethality", "Phenotypic Enhancement", "Phenotypic Suppression" and "Dosage Growth Defect" were not considered. Mpact entries were considered only if they had the tag “902.01.01.02.01”, which indicates physical interactions. We separately considered the combined TAP-MS data from [36] (data used were those with Purification Enrichment scores of 3.19 and above as defined in [36]), as it overlaps considerably with what has been deposited in Biogrid by [13] and [14]. We considered only one citation for an interaction reported in Collins *et al.* (2007), even if it had been reported by one of, or both of the two original studies. Finally, we considered only interactions associated with PubMed ids, because these can be tracked to their original experimental evidence. We considered different ids as being independent evidence. Although this may inflate the confidence in some interactions, it should not affect the identification of new interactions as described in our screen.

1.5.14. Overlap with previous large-scale studies.

In order to calculate the overlap between our screen and previous results derived from large-scale experiments and a catalogue of manually annotated complexes we computed a number of interactions common to our screen and each of the datasets described below. We first calculated the overlap between DHFR PCA PPI and all interactions reported in a previous study (Figs. 12 & 13; numbers in circle indicate the total number of interactions detected in a particular dataset (since affinity-purification based methods cannot detect homomeric interactions, only heteromeric interactions were used for the analysis)). None of the datasets reports a complete set of interactions between all yeast proteins and thus low overlap between different data sets may be because different datasets cover different sets of yeast proteins. Therefore, we performed a normalization of the number of interactions based on how many proteins are in common between two datasets and how many same interactions could be detected (Fig. 10). Numbers in the PCA circles indicate how many interactions reported in a particular dataset could be detected by our screen (an interaction can be detected only if one of interacting partners showed a signal as bait and another as prey or vice-versa in our final network). Numbers in circles that correspond to a reference dataset indicate how many interactions from our final network could be detected by a reference dataset. When only interactions that could be detected by both experiments are considered, the overlap between DHFR PCA and reference datasets reaches as high as 50%. Reference datasets and criteria for normalization are described below: 1), MIPS catalogue of

manually annotated complexes. Downloaded from ftp://ftpmips.gsf.de/yeast/catalogues/complexcat/complexcat_data_18052006. Complexes detected by large-scale experiments were filtered out from this file and interactions were assigned between all proteins that belong to the same complex. An interaction is considered to be possible if both interacting partners are present in the MIPS catalogue. 2), Krogan *et al. [14]* The core dataset was obtained from supplementary table 7, that lists successfully identified baits and preys obtained from supplementary tables 2 and 3. An interaction is considered to be possible only if one of the proteins is present in the baits list and another is in the preys list. We don't consider a co-occurrence of both proteins in the prey list since the core dataset of this study contained only bait-prey pairs. 3), Gavin *et al. [13]* The network of interactions was obtained as deposited in Biogrid. This study used a statistical framework for deriving a high confidence set of interactions that makes possible interactions between two prey proteins. Therefore for normalization, we considered an interaction to be possible if for a pair of interaction partners, a bait-prey or prey-prey pair exists in the raw purification data (downloaded from http://yeast-complexes.embl.de). 4) Collins *et al. [36]* We used high confidence data with a PE score cutoff of 3.19. Normalization was performed as described above for Gavin *et al.* using a combination of both Krogan and Gavin raw datasets.5), Ito *et al.*, Uetz *et al. [35]* Interactions detected by yeast two-hybrid assays. Interaction was considered possible if one of the interacting partners is among proteins that showed an interaction when tagged with a binding domain and another is

among proteins that showed interactions when tagged with an activation domain.6), Miller et al. [26] Interactions tested using the split-ubiquitin reporter. The data were extracted from supplementary Table 1. An interaction is considered possible if a corresponding pair of Cub-PLV and NubG ORF is present in the dataset.

1.5.15. Clustering of high confidence interactions.

2,534 heteromeric high confidence interactions were clustered as described in [73]. For clustering purposes and during calculation of the association matrix, a value for self-association of a protein was set to 1. Hence, observed homomeric interactions could not be distinguished from those used for clustering purposes. Since two proteins that belong to the same complex may not be interacting but rather kept together by a common interacting partner, an association matrix based on the number of links that connect two proteins can be used for clustering [74]. We found that for the sparse PPI matrix derived from DHFR PCA data, this procedure resulted in a tight and natural clustering of interactions among subunits of known complexes. Briefly, the network was organized into an association matrix with entries for pairs of proteins that range between 0 and 1. Values were calculated as $1/d2$, where d is the shortest path in the network between these two proteins. A hierarchical agglomerative averagelinkage clustering with the uncentered correlation coefficient as the distance matrix was then applied to the association matrix [74]. The data were then visualized using iVici (File S1, http://michnick.bcm.umontreal.ca/ivici/) [75]. Only direct interactions are shown on a complete map (Fig. 15). On the

insets, direct interactions are bright red, while indirect interactions (2 or 3 links between two proteins) are shown as two consecutively darker shades of red, respectively. The clustered network is available in Supplementary File S1. This file can be opened and visualized using iVici [75], a platform independent software available at **(**http://michnick.bcm.umontreal.ca/ivici/).

1.5.16. GO enrichment.

The GO slim map was obtained from SGD 17 February 2007. For every pair of annotation terms associated with a biological process, cellular compartment and molecular function we calculated enrichment/depletion in the number of interactions in our high confidence dataset compared to the number expected by chance. First, we calculated the number of interactions that are detected between proteins associated with a specific pair of GO terms. Next, we constructed 10,000 randomized networks of interactions between the same set of proteins as in our high confidence network and with the same number of interactions per proteins. The randomized networks contain only interactions that could be detected in our screen (as described for the normalized overlap calculations). For each randomized network, a number of interactions between proteins associated with a specific pair of terms were calculated and a z-score was derived by comparison of these numbers with a corresponding number of interactions detected by our analysis. High z-scores correspond to significant enrichment in interactions comparing to random while negative z-scores values correspond to significant depletion.

For pairs of interacting proteins that lack a common GO slim annotation term we calculated a GO semantic score, which is another measure of a functional relationship between proteins. The semantic score takes into account a specificity and hierarchy of a parent term that is common to a pair of proteins. Thus, pairs of proteins that are more closely related have higher scores [76]. The results were compared with semantic scores calculated for random networks generated as described above.

2.Chapter 2: A pathway that controls early M-phase progression by regulation of *CLB2* mRNA stability *via* methylation of hnRNPs

Vincent Messier, Daniel Zenklusen and Stephen W. Michnick†

Département de Biochimie, Université de Montréal.

† Author for correspondence

2.1. Overview

While canonical pathways and regulatory networks provide a representation of molecular interactions in the cell that appears immutable, real regulatory pathways are anything but static. Rather, they reconfigure dynamically as a function of the specific molecular context in which they operate. This was shown initially in yeast and more recently in mammalian cells [77-79]. The cellular response to environmental stress is associated with massive rewiring of protein complexes and of genetic interactions that they tolerate [28]. These lines of mechanistic causality propagate in the protein network resulting in coherent morphological dynamics and cellular responses. Cells have evolved different strategies to coordinate these cellular network rearrangements including: (I) transcriptional programs launched by common transcription factors, (II) post-translational modifications, such as phosphorylation by protein kinases that alter and activate substrates required for cellular responses, (III) active molecule sequestration in subcellular compartments, and (IV) molecule-targeted decay, to name a few [80] (Fig. 17). For example, treatment of yeast with rapamycin has pleiotropic effects on cellular processes. In the cell, rapamycin binds to the ubiquitous and abundant protein FKBP12 and the rapamycin-FKBP12 complex in turn forms a ternary complex with the Target Of Rapamycin (TOR) kinase and inhibit its enzymatic activity [81]. Inhibition of TOR by rapamycin promotes effects through down-regulation of translation by preventing TOR phosphorylation and activation of the protein S6 kinase and in turn phosporylation and activation of the S6

ribosomal protein. Rapamycin also prevents direct phosphorylation of initiation factor (eIF) 4E binding protein 1 (4E-BP1) by TOR, which prevents 4E-BP1 binding to and inhibition of eIF-4E that reduce mRNA translational rates [82, 83]. In parallel, PP2A phosphatase is activated following release of its catalytic subunit, Pph22, from the TAP42 complex when rapamycin-FKBP binds to TOR kinase [84]. The phosphatase catalytic subunit is released when Tap42, a direct substrate of TOR kinase in yeast, is dephosphorylated. The Pph22 activation also occurs during the mitosis transition and under a variety of stresses, such as nitrogen starvation and carbon source limitation, in a TOR dependant manner [84-88]. It is interesting that all of these conditions have drastic effects on cell cycle progression. In this chapter I describe experiments that link TOR and nutrient limitation to a novel mechanism of control of cell cycle progression at the level of mRNA stability.

All cells arise from division of existing cells. The cell reproductive mechanisms are separated into distinct processes leading to cell cycle completion. In budding yeast, first, cell contents are duplicated, and second, they are distributed asymmetrically between the newborn daughter cell and the mother cell. Most of the cell components are replicated continuously throughout the cell cycle, resulting in doubling of cell size by the end of the cycle. The genetic material, conversely, is doubled once per cycle. This occurs during a discrete stage of DNA synthesis or S phase (Fig.18). The separation of mother cell constituent into individual daughter cell occurs in a brief stage

called mitosis or M phase. The S and M phases are separated by two distinct gaps providing additional time for cell growth and for regulatory events to take place. The first gap phase, G1, occurs before S phase, whereas, G2 takes place prior to M phase. During the G1 phase, cells are accumulating mRNAs and proteins in preparation for the subsequent steps leading to mitosis. G1 phase ends when a cell commit for mitosis in a checkpoint called START and move into S phase. Similarly, G2 phase is characterized by a rapid growth and protein accumulation period during which the cell readies itself for mitosis.

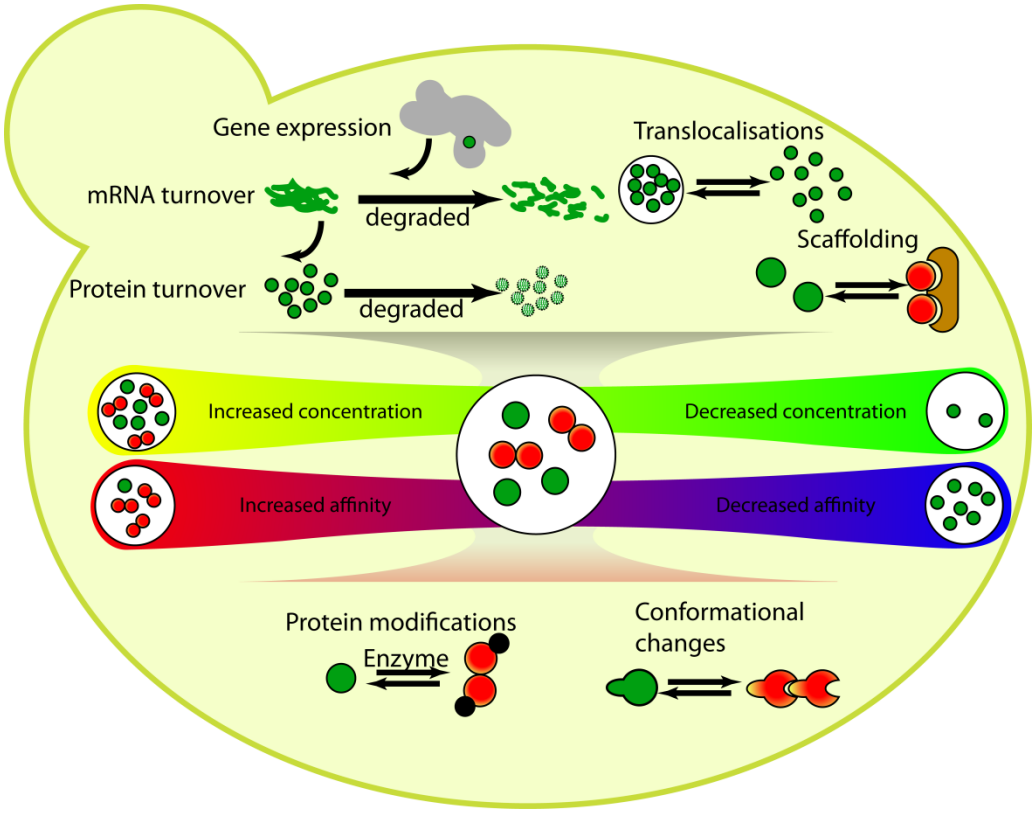

Figure 17: Network representation of a protein interaction regulatory cascade.

Protein, denoted as nodes in a network (red and green circles), represent several entities (gene, mRNA, and protein) and events (transcription, translation, degradation, translocalization, scaffolding, protein post-translational modifications and conformational changes) that simultaneously happen in space and time. Although a series of regulatory events can be represented as a node in the network, the dynamics of the entities and the biological processes that make up the node cannot be symbolize in the same manner.

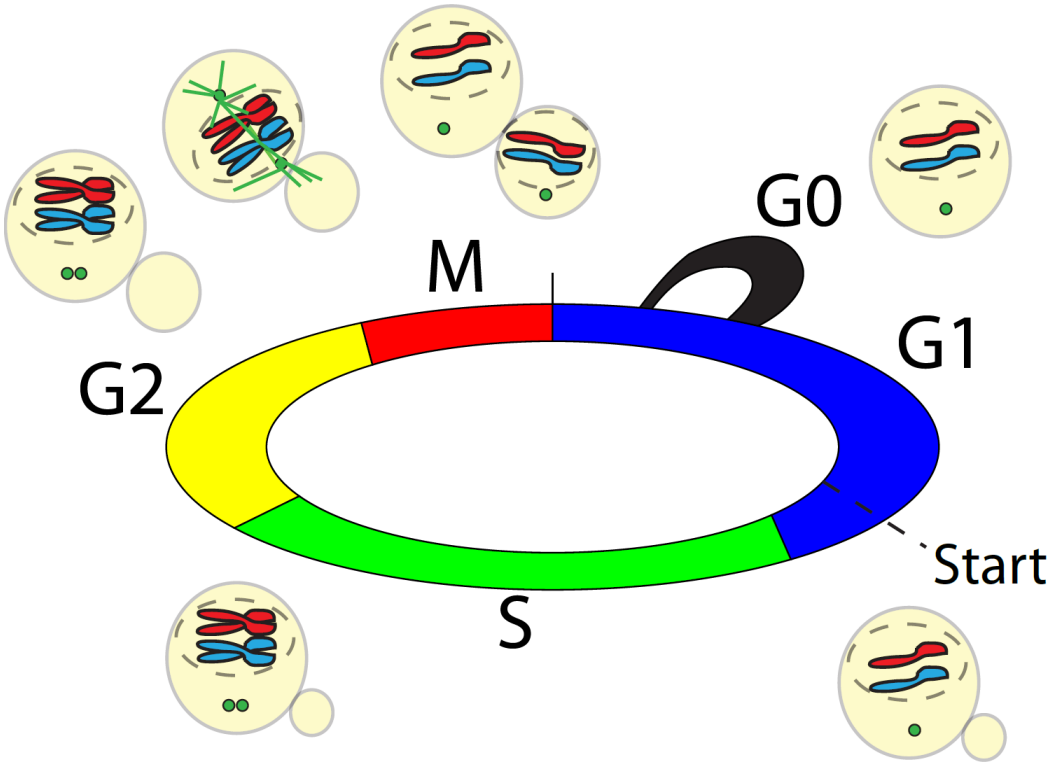

Figure 18: The events of the eukaryotic cell cycle

Cells in G1 regulatory period become committed to either continued division or exit from the cell cycle, entering a prolonged state called G0 [89]. During the G0 stage, cells are maintaining a minimal metabolic and structural state allowing survival in extreme conditions, until stress conditions are alleviated and cell cycle can be resumed. G2 is a regulatory period during which cells ensure DNA integrity and adequate adaptation status to maximize daughter cell viability [90]. The central components that control division are an enzyme family called the cyclin-dependant kinases (Cdks) [6]. These kinases bind to sets of phase specific cyclins, conferring specificity and acute kinase activity for substrates, thereby initiating cell cycle events. These molecules are the limiting factors in cell cycle progression and their regulation is therefore a main focus [91].

The keystone process in the cell cycle oscillation is ultimately the cyclin concentrations, the rate-limiting elements. The quantity of each cyclin is regulated through their synthesis, whose expression is stimulated by various transcription factors such as SCB-binding factor (SBF) and MCB-binding factor (MBF), and through their degradation, promoted by both Cdc20-and Cdh1- Anaphase promoting complex (APC), and Grr1-Skp, Cullin, F-box containing complex (SCF). In fact, activation of SBF and MBF transcription initiates key cell cycle events, including budding, DNA synthesis, and spindle pole body duplication. The SBF complex activates transcription mainly through a *cis*-acting sequence element; called Swi4/6 cell cycle box (SCB). Genes

activated by SBF include those encoding G1 cyclins; CLN1 and CLN2. The MBF complex recognizes the Mlu1 cell cycle box (MCB) element and activates G1-specific transcription of the S-phase cyclin genes, CLB5 and CLB6. In contrast, the APC and SCF complexes are ubiquitin ligases that conjugate ubiquitin onto cyclins and target them for proteolysis. The APC controls metaphase-anaphase transition with its activator Cdc20. Its activity is required for sister chromatids separation. APC with the other activator, Cdh1, is also involved in the G1 phase and controls levels of mitosis regulating proteins. SCF controls G1/S through G2/M transitions. The complex promotes cell cycle progression and cell growth by targeting G1/S cyclins to degradation.

Both cyclin mRNA synthesis and cyclin protein degradation are crucial for the cell cycle progression and delays required for adaptations. However, the molecular mechanisms regulating the cyclin amounts per cell can take a variety of forms. It is not clear whether post-transcriptional regulation could control availability of mRNA for translation. It is now well established that transcription and nuclear mRNA processing and cytosolic stability are coupled, as mRNA maturation, such as capping, splicing and polyadenylation, mRNA export and stability are initiated on nascent transcripts [92, 93]. It was shown that Npl3, Hrp1, Nab2 and other heterogeneous nuclear RNA binding proteins (hnRNPs) specify the translational control, mRNA decay and mRNA localization shuttle between the cytoplasm and the nucleus [94-97]. Their early binding to RNA secondary structure folded co-transcriptionally in the nucleus

is essential for their impact on RNA fate [98]. For example, *CLB2* mRNA is localized to the bud tip in a manner dependent of RNA binding protein Khd1 [99] and is stabilized when associated with Dbf2 kinase [93]. Similarly, *CLN3* mRNA is localized into Whi3 protein foci, inhibiting its translation [100]. Distinctively, *CLB6* mRNA is degraded early in unperturbed S phase, but stabilized during replication stress and DNA damage [101]. In fact, it was shown that Npl3, Nab2, Hrp1 and Tho2 bind to all of the cyclin transcripts [96]. The exact mechanisms by which these cyclin and possibly other transcripts are regulated remains to be determined.

Post-translational modifications of hnRNPs modulate their capacity to bind to and regulate transcripts. For example, specific arginine residue methylation occurring on Nab2, Npl3, Tho2 and Hrp1 by the only arginine methyltransferase in yeast, Hmt1, directly regulates their RNA binding function [94, 95, 102]. This enzyme exists as a functional homohexamer as shown by both affinity chromatography and high-resolution X-ray crystallography [103]. S-Adenosyl-L-methionine (SAM) acts as methyl group donor co-factor for methylation of hnRNPs and other proteins on consensus sequences surrounding an arginine residue (-(G/F)(G/R)GRGG(G/F)-). These modifications may alter their activity, interactions or stability [94-97]. Hmt1 is thought to be regulated at the level of transcription and transcript production is altered under different conditions [104]. There is, however, no evidence that changes in transcript level result in changes in Hmt1 protein synthesis or enzyme activity in the cell. In

this chapter, I present evidence that Hmt1 assembly and induction of catalytic activity is directly controlled by phosphorylation.

Completion of this thesis work required that I adapt and apply recent conceptual and technical advances in the study of cell cycle and RNA measurements. New understanding of cell cycle regulation have resulted from recognition of two things: first, the importance of dissecting populations of cells into those having different lengths of mitotic phases (e.g. newborn daughter *versus* mother cells), second, the use of fluorescent reporters and single cell analysis to define specific regulatory cues in time [105, 106]. Equally, to understand the role of mRNA regulation in cell reproduction control and to set the stage for studying upstream regulatory pathways, requires the ability to follow the spatial and temporal dynamics of single molecules mRNA following transcription [9, 93].

Ultimately, we need strategies to identify upstream regulatory pathways that control mRNA dynamics during the cell cycle. In this chapter I describe a strategy I developed to identify proteins that regulate oscillating mRNAs during mitosis based on proteome-wide measurements of change in protein-protein interactions *in vivo* following perturbations of the cell by rapamycin, a compound known to delay the cell division at G2/M [107]. I then used single mRNA molecule tracking techniques combined with genetic and biochemical analyses, to infer mechanistic details regulating transcript turnover and

translational availability. The same strategy could be applied to identify proteins involved in any cellular process.

Background

In the following manuscript that was submitted in an article format to the journal *Cell*, we used single cell, single molecule FISH to measure absolute quantities of *CLB2* mRNAs in response to a nutrient-limiting mimic stress, treatment with rapamycin. Surprisingly, *CLB2* transcript failed to accumulate in early M-phase in rapamycin-treated cells, leading us to investigate its regulation by hnRNPs and the arginine methyltransferase, Hmt1. By quantitative fluorescence microscopy, we showed that rapamycin causes sequestration of hnRNPs to the nucleus preventing their localization to cytosol, where they are indispensable to mRNA stabilization. Genetic and protein-protein interaction screens, combined with biochemical analysis revealed that the protein kinase Dbf2 competes with phosphatase PP2A (Pph22) for phosphorylation, and cooperative phosphorylation-induced oligomerization and activation of the arginine methyltransferase, Hmt1. Hmt1 in turn methylates hnRNPs to promote their co-transcriptional association with mRNA and nuclear export. Rapamycin inhibition of TOR induces the recruitment of the PP2A phosphatase catalytic subunit Pph22 to Hmt1 and dissociation of Dbf2, favoring dephosphorylation and inactivation of the arginine methyltransferase. Based on these results, we propose a general model of how *CLB2* transcript accumulation and therefore early M-phase transition can occur under normal and stress conditions. We discuss how the unique mechanism of cooperative Hmt1 activation generates a modulator of RNA stability in which small variations in PP2A activity leads to sharp changes in Hmt1 activity while Dbf2 activity remains at constitutive levels. We propose that this may be a general mechanism to modulate the rate of cell cycle progression in response to alterations in environmental conditions.

A pathway that controls early M-phase progression by regulation of *CLB2* mRNA stability *via* methylation of hnRNPs

Vincent Messier, Daniel Zenklusen and Stephen W. Michnick†

Département de Biochimie, Université de Montréa.l

† Author for correspondence

2.2. Highlights

- *CLB2* mRNAs fail to accumulate in cells treated with rapamycin
- *CLB2* transcript decay depends on phosphorylation and cooperative activation of Hmt1
- Dbf2 and Pph22 regulate Hmt1 phosphorylation, oligomerization and activation
- Entry into M-phase is sharpened by Hmt1-dependent stabilization of *CLB2* transcripts

2.3. Introduction

Transcriptional programs and therefore mRNA expression patterns are constantly adapting to intrinsic and extrinsic cellular cues including cell cycle control mechanisms. We have little understanding of how posttranscriptional regulation of mRNA contributes to control of mitosis [108, 109]. Temporal and spatial patterns and levels of transcript expression are governed by transcriptional mechanisms, as well as mRNA binding proteins and degradation machinery [110]. While transcription is generally thought to result in linear accumulation of transcripts, the rate of mRNA decay can change exponentially in response to external stress signals [111]. mRNA decay machinery is modulated by decay regulatory proteins that bind to untranslated regions (UTRs) or the open reading frame (ORF) of mRNAs, thereby inducing or preventing mRNA degradation [112, 113]. Most of the mRNA decay machinery is located in the cytosol where mature mRNA degradation is performed (Reviewed in [114]). Early factors, such as the heterogeneous nuclear RNA binding proteins (hnRNPs), bind co-transcriptionally to mRNA to promote mRNA export and stability; their association with nascent mRNA is partially dependent upon post-translational arginine methylation [96]. However, the contribution of mRNA degradation regulators, mRNA degradation machinery and the pathways that link external and internal signals to regulation of mRNA stability remain subjects of intense investigation.

Recent advances in single cell, single molecule fluorescent *in situ* hybridization (FISH) studies have provided insights into mRNA regulatory mechanisms [93, 115, 116]. Notably, a novel mechanism was reported that controls stabilization of the mRNA of G2/M transition regulators, the cyclin Clb2 and Swi5 in the budding yeast *S. cerevisiae* [93]. *CLB2* and *SWI5* mRNA accumulates in the cell starting at early until late M-phase when it is rapidly degraded [117]. Trcek *et al.* demonstrated that the degradation of *CLB2* and *SWI5* transcripts was regulated by co-transcriptional recruitment of the mitotic exit regulating kinase Dbf2 to the *CLB2* and *SWI5* promoters. Dbf2 bound directly to *CLB2* and *SWI5* mRNA and they proposed that the interaction regulates mRNA degradation in a CCR-NOT complex dependent manner. Other proteins are recruited to the *CLB2* locus, including the hnRNP components Hrp1 and Npl3 in a methyltransferase dependent and independent manner, respectively [96], promoting mRNA stability and nuclear mRNA export.

The novel function of Dbf2 in regulating cell cycle-specific transcription begs the question of how such functions might be regulated by external signals to control cell cycle progression. For instance, delay or arrest in early M-phase are induced by various cellular stresses, including protein synthesis inhibition, osmotic shock, DNA damaging agents, Target of rapamycin (TOR) inhibition by rapamycin and nutrient starvation [107, 118-120]. We know little about how M-phase delay or arrest is caused by these

treatments and whether it may occur at the level of mRNA or protein synthesis or degradation.

As with other cellular stresses, rapamycin will delay or arrest different phases of the cell cycle, but the details of how this occurs are not clear [107]. Rapamycin prevents TOR phosphorylation of substrates by first forming a binary complex with the endogenous protein Fkbp12, and this complex then forms a ternary complex with TOR [81]. Inhibition of TOR by rapamycin promotes pleiotropic effects through down-regulation of translation by preventing TOR phosphorylation and activation of the protein S6 kinase and in turn phosporylation and activation of the S6 ribosomal protein. Rapamycin also prevents direct phosphorylation of initiation factor (eIF) 4E binding protein 1 (4E-BP1) by TOR, which prevents 4E-BP1 binding to and inhibition of eIF-4E [82, 83]. In addition, TOR pathway was involved in nitrogen starvation and carbon source limitation sensing and was, therefore, utilized as a tool to mimic these processes. In fact, Gln3 and Ure2, transcription factors responsible for nitrogen starvation and sugar limitations adaptations, translocalizations and gene expression targets upon starvations and rapamycin treatments suggest direct regulation by the TOR pathway [121]. Simultaneously, the translation regulation by Gcn2 and vacuolar organization induced by Npr1 upon starvation are revealed to be targeted by rapamycin-induced TOR inhibition [122, 123]. TOR kinase affects normal G1 cell cycle progression and rapamycin causes reduction of G1 cyclin transcripts and translation in cells, likewise observed

during nitrogen starvation [124, 125]. Interestingly, however, the TOR pathway delays early M-phase cell cycle progression in a translation-independent manner, since the delayed transition is unaffected in cells treated by ribosomal inhibitors or in mutant strains that exhibit reduced initiation of translation [107]. Based on these observations and previous mentioned evidence of Trcek *et al.,* we hypothesized that a TOR-regulated pathway controls G2 cyclin stability post-transcriptionally, *via* mechanisms that regulate mRNA degradation [93].

In this work, we used single cell, single molecule FISH to measure absolute quantities of *CLB2* mRNAs in individual cells in response to rapamycin treatments. Surprisingly, *CLB2* transcript failed to accumulate in early M-phase in rapamycin-treated cells, leading us to investigate its regulation by hnRNPs and the arginine methyltransferase, Hmt1. By quantitative fluorescence microscopy, we showed that rapamycin causes sequestration of hnRNPs to the nucleus preventing their localization to cytosol, where they are indispensable to mRNA stabilization. Genetic and protein-protein interaction screens, combined with biochemical analysis revealed that the protein kinase Dbf2 competes with phosphatase Pph22 for phosphorylation, and cooperative phosphorylation-induced oligomerization and activation of the arginine methyltransferase. Hmt1 in turn methylates hnRNPs to promote their co-transcriptional association with mRNA and nuclear export. Rapamycin inhibition of TOR induces the recruitment of the PP2A phosphatase catalytic subunit Pph22 to Hmt1 and dissociation of Dbf2, favoring dephosphorylation

and inactivation of the arginine methyltransferase. Based on these results, we propose a model of how *CLB2* transcript accumulation and therefore early M-phase transition can occur under normal and stress conditions.

2.4.Results

2.4.1. CLB2 mRNA fail to accumulate in early M-phase in cells treated with rapamycin

We counted single transcripts in individual cells using single cell, single molecule FISH [9]. To achieve this, multiple oligonucleotides complementary to the *CLB2* and *ACT1* mRNAs (actin positive control), labeled with 4-5 fluorescent dyes, were hybridized to paraformaldehyde fixed yeast cells and signal emitted from individual mRNAs detected by wide field fluorescence microscopy [93, 126]. Total cellular mRNA distribution was further monitored using an oligo dT probe detecting all cellular polyA mRNA (Fig. 19A). Individual *CLB2* and *ACT1* mRNAs within individual cells detected as diffraction limited spots were counted [93]. To monitor *CLB2* accumulation over the cell cycle, cells were arrested in G1 by treatment with α-factor and then released into the cell cycle by the removal of α-factor.

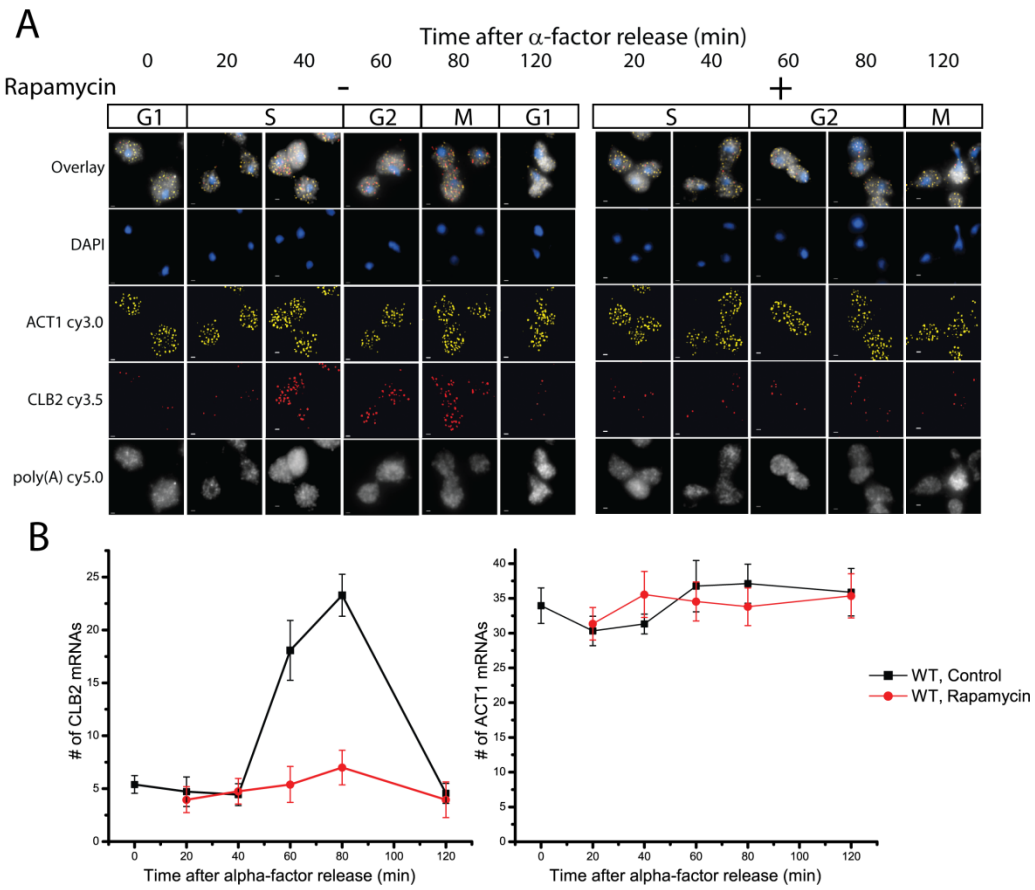

Figure 19: *CLB2* mRNA fail to accumulate in early M-phase in cells treated with rapamycin.

(A) Accumulation of mRNA following release of cells from G1 phase arrest by removal of α-factor. Rapamycin (200 nM) was added to one of the synchronized cultures 20 minutes after α-factor release to prevent premature rapamycin-induced G1 arrest. A mix of cy3 labeled *CLB2* mRNA probes (in yellow), cy3.5 *ACT1* mRNA probes (in red) and poly(A) containing mRNA probes (in gray scale) were combine to reveal single mRNA of individual cells by FISH in cells fixed every 20 minutes following α-factor release or rapamycin treatment. Cell nucleus was stained by DAPI (in blue) contained in cover slip mounting medium. **(B)** Detection and counting of fluorescent mRNAs and single probes was performed for 25 cells treated with rapamycin (200 nM) or ethanol vehicle at each time points for *CLB2* mRNAs (left panel) and *ACT1* mRNAs (right panel). Scale bar: 1 μm.

Cells were treated with 200 nM rapamycin or ethanol vehicle (Control) 20 minutes after α-factor release to prevent rapamycin-induced G1 arrest [107]. Cells were harvested and fixed every 20 minutes for one cell cycle and after 120 minutes following α-factor release (t = 0). Fixed cells were hybridized to mRNA probes and *CLB2, ACT1* and poly(A) mRNA were quantified (Fig. 19B). During the first 40 minutes following α-factor release, control cells harboured an average of 5 *CLB2* mRNA copies per cell, independent of whether they were treated with rapamycin or not. Cells entering early M-phase (60 minutes) accumulated an average of 18 *CLB2* mRNA molecules and continued to accumulate copies into M-phase (80 minutes) reaching a maximum of about 23 molecules per cells before returning to basal copy numbers at 120 minutes. In contrast, rapamycin treated cells did not accumulate new *CLB2* mRNAs, remaining at an average of 5 molecules per cell throughout the recorded period. This conditional build up of transcripts in cells during early M-phase was specific to *CLB2* mRNA as *ACT1* mRNA remained constant in control and rapamycin-treated cells.

2.4.2.Rapamycin causes sequestering of hnRNPs in the nucleus

To investigate the cause of *CLB2* transcript accumulation in cytosol in early M-phase and failure to accumulate in cells treated with rapamycin, we first reasoned that rapamycin could affect mRNA nuclear export and cytoplasmic related mRNA stabilizing activities. It has already been

demonstrated that deletion of genes coding for hnRNP components causes accumulation of mRNA in the nucleus and reduced mRNA stability [127, 128]. Thus we first tested whether correct nuclear and cytosolic localization of hnRNP components occurs in cells treated with rapamycin. We used diploid strains that express one of each the endogenous hnRNP protein Nab2, Npl3 and Hrp1 as fusions to GFP and, simultaneously, express endogenous Nup49 as a fusion to dtTomato as a nuclear membrane marker. Strains were treated with rapamycin or ethanol vehicle for 6 hrs before being adhered on ConA-coated glass bottom 96-well plates and imaged (Fig. 20A). In control cells, the GFP signal intensities for all three hnRNP components were mostly localized within the nucleus with some cytosolic signal. However, cells treated with rapamycin exhibited significantly reduced cytoplasmic GFP signal intensities. We quantified GFP signal intensities over the nucleus, the cytoplasm and the whole cell, and normalized the quantity by area, for 100 cells per strain treated with rapamycin or vehicle control (Fig. 20B). While GFP signal densities stayed relatively unaffected in the nucleus, we observed a significant decrease for each hnRNP in the cytoplasm. The intensities of GFP that we observed in the cytoplasm of rapamycin treated cells were within the range of auto-fluorescence of *wildtype* cells. The nuclear sequestration that we detected in rapamycin treated cells were specific to the hnRNP components Nab2, Npl3 and Hrp1, and did not affect the localization of the mRNA export receptor Mex67, nor did it change the distribution of different karyopherins, including Sxm1, Kap123, Crm1, Los1 and Pse1 (Fig. 21A).

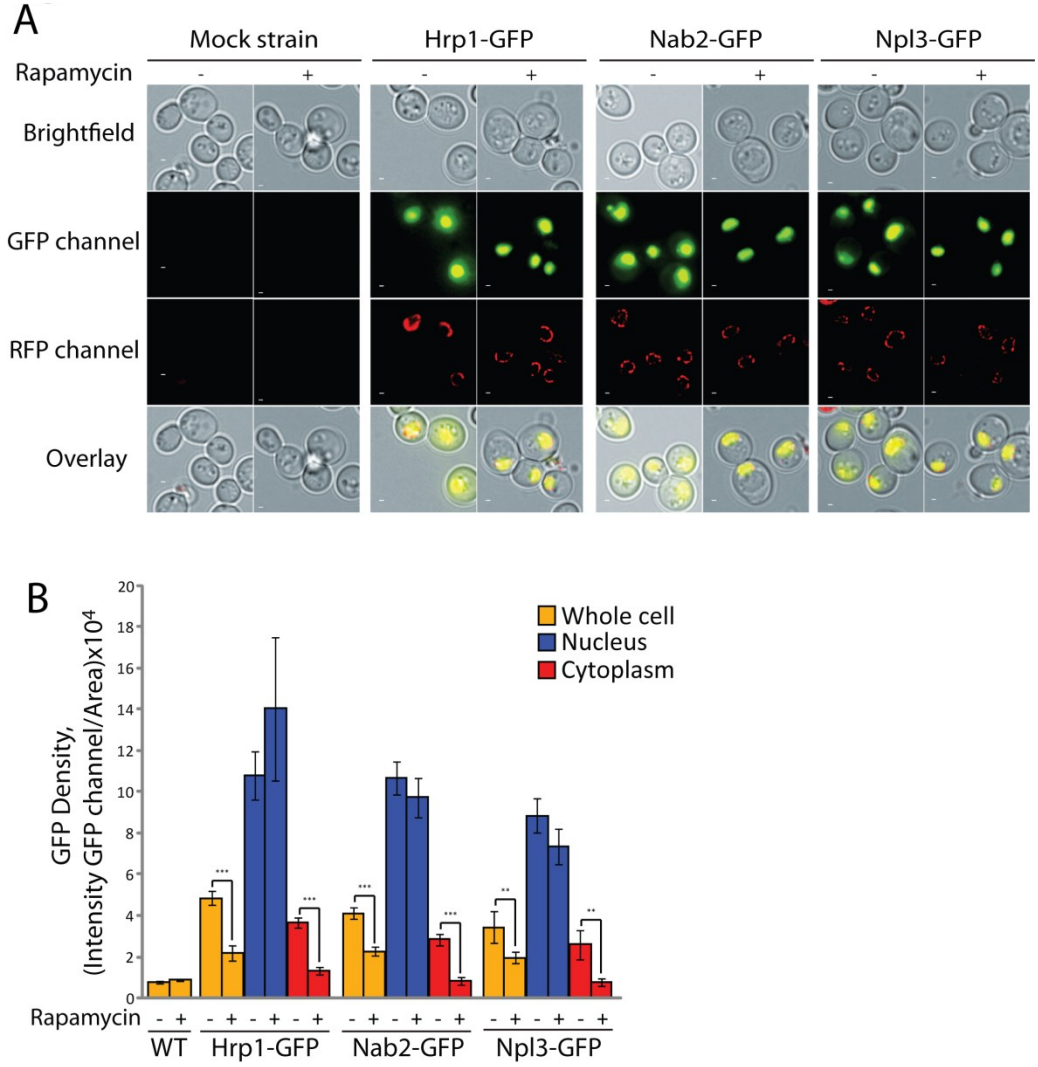

Figure 20: hnRNPs are sequestered in the nucleus in rapamycin treated cells.

(A) Yeast cells endogenously expressing hnRNPs fused to GFP (in green) and Nup49 fused to dTomato (in red), as nuclear membrane marker, were treated with 200 nM rapamycin or ethanol vehicle for 6 hrs. Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and images were overlayed. **(B)** GFP intensities for hnRNPs fused to GFP were quantified and area normalized for 100 cells for each strain treated with rapamycin (200 nM) or ethanol vehicle (in yellow). Cellular GFP signal was further segmented into nucleus (in blue) and cytoplasm (in red) based on Nup49 nuclear membrane marker. Scale bar: 1 µm. See also Figure 21.

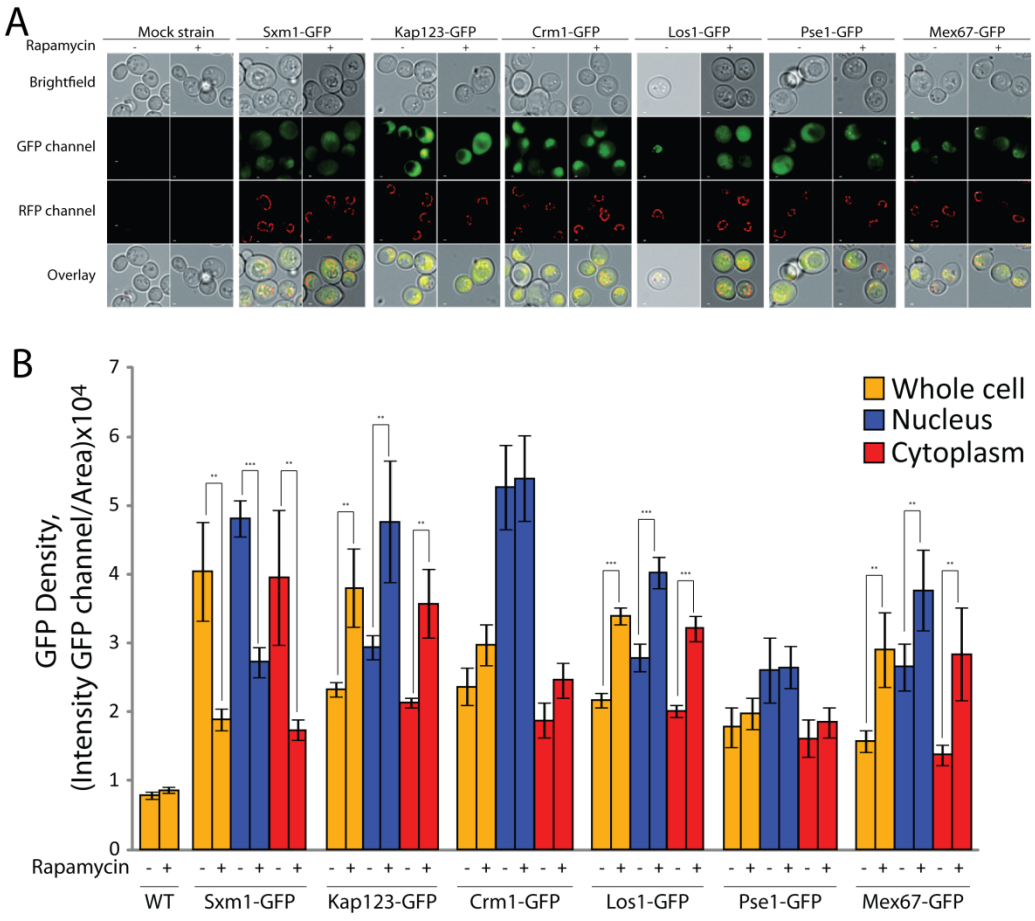

Figure 21: Other karyopherins and mRNA nuclear export elements are homogenously affected in rapamycin treated cells.

(A) Yeast cells endogenously expressing hnRNPs fused to GFP (in green) and Nup49 fused to dTomato (in red), as nuclear membrane marker, were treated with 200 nM rapamycin or ethanol vehicle for 6 hrs. Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and images were overlayed. **(B)** GFP intensities for hnRNPs fused to GFP were quantified and area normalized for 100 cells for each strain treated with rapamycin (200 nM) or ethanol vehicle (in yellow). Cellular GFP signal was further segmented into nucleus (in blue) and cytoplasm (in red) based on Nup49 nuclear membrane marker. Scale bar: 1 μm.

2.4.3.Rapamycin does not affect Hmt1 activity through changes in expression or localization

Both efficient association and localization of the hnRNP components are dependent on specific methylation by the arginine methyltransferase, Hmt1 [95, 129, 130]. Since rapamycin inhibits general translation we reasoned that simply a reduction in Hmt1 expression and therefore reduced methylation activity could account for nuclear sequestering of hnRNP components. Using a yeast strain expressing endogenous Hmt1 fused to GFP, we treated cells with or without rapamycin for 6 hours. Hmt1 expression was quantified as GFP fluorescence intensity by FACS (Fig. 22A). Contrary to expectation, cells treated with rapamycin exhibit approximately an order of magnitude increase in Hmt1 expression per cell.

Although Hmt1 expression increases upon rapamycin treatments, it is possible that Hmt1 activity must be localized to either cytosolic or nuclear compartments in order to act on hnRNPs. We thus examined the effects of rapamycin on Hmt1 localization in the same strains as above, treated with or without rapamycin for 6 hours. Rapamycin caused an increase of Hmt1 in both compartments (Fig. 22B,C). Thus, neither changes in expression nor localization of Hmt1 induced by rapamycin can account for changes in hnRNP localization.

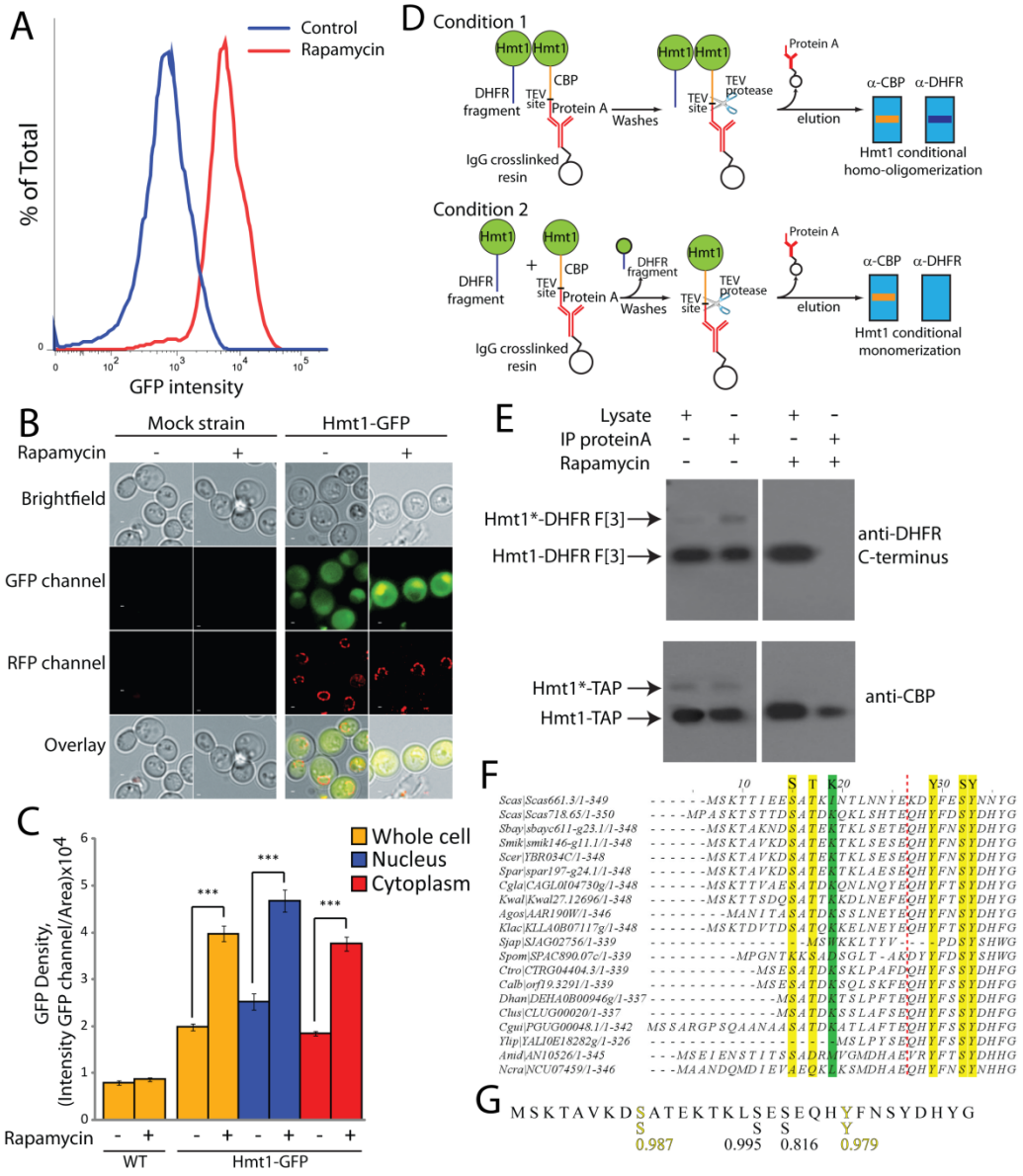

Figure 22: Both modifications and homo-oligomeric affinities are lost on Hmt1 in cell treated with rapamycin.

(A) Yeast cells endogenously expressing Hmt1 fused to GFP were treated with 200 nM rapamycin (in red) or ethanol vehicle (in blue) 6 hrs. GFP intensities per cell in culture treated with rapamycin or ethanol vehicle were record in a LSRII FACS for 30, 000 cells and distribution were displayed in bar graph. **(B)** Yeast cells endogenously expressing Hmt1 fused to GFP (in green) and Nup49 fused to dTomato (in red), as nuclear membrane marker, were treated with 200 nM rapamycin or ethanol vehicle for 6hrs. Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and images were overlayed. **(C)** GFP intensities for Hmt1fused to GFP were quantified and area normalized for 100

cells for each strain treated with rapamycin (200 nM) or ethanol vehicle (in yellow). Cellular GFP signal was further segmented into nucleus (in blue) and cytoplasm (in red) based on Nup49 nuclear membrane marker. **(D)** Flow chart of modified TAP tag purification for Hmt1 homomeric complexes. Diploid strain expressing Hmt1 fused to the TAP tag from the first locus and Hmt1 fused to DHFR F[3] from second locus are treated with rapamycin (200 nM) or ethanol vehicle for 6 hrs. After washes and TEV protease elution, presence in immunoprecipitate of cleaved Hmt1-TAP and Hmt1-DFHR F[3] revealed condition specific oligomerization assembly. **(E)** Diploid yeast cell endogenously expressing Hmt1-TAP on first locus and Hmt1-DHFR F[3] on second locus were treated with 200 nM rapamycin or ethanol vehicle for 6 hrs. For each condition, cells were harvest, lysed and proteins were purified to be applied on IgG resin. IgG immunoprecipitates was eluted by TEV protease cleavage of the Hmt1-TAP. IgG immunoprecipitates and cell lysates were analyzed by immunoblotting for levels of Hmt1-TAP, co-immunoprecipitated and levels of Hmt1-DHFR F[3]. **(F)** Hmt1 N-terminal first 20 amino acids contain conserved serine, threonine and lysine residues across fungi. **(G)** Phosphorylation prediction perform by NetPhos 2.0, target serine 9 and tyrosine 23 to be phosphorylated and conserved across fungi. Scale bar: 1 μm.

2.4.4.Rapamycin prevents homo-oligomerization and activation of Hmt1

Hmt1 arginine methyl transferase activity only occurs when it is a dimer or higher order homo-oligomer [103]. Thus we next tested whether rapamycin could prevent homo-oligomerization of Hmt1. We devised a simple *in vivo* assay to detect oligomeric states based on affinity and immuno-affinity detection of dual immunogenic peptide-tagged Hmt1 (Fig. 22D). We generated a strain in which a TAP tag sequence was integrated 3' at one locus of Hmt1 and an antigenic C-terminal peptide sequence of mouse dihydrofolate reductase (DFHR F[3]) 3' to the Hmt1 coding sequence on the second locus. Following treatment of the strain with rapamycin or vehicle control, we performed IgG immunoprecipitation (against Protein A domain of the TAP tag) of Hmt1-TAP, eluted and ran samples on SDS-PAGE followed by immunoblot with an antibody against either DHFR F[3] or the TAP calmodulin-binding peptide (CBP) (Fig. 22E). We observed that Hmt1-TAP co-immunoprecipitated Hmt1-DHFR F[3] in control, non-treated cells. However, in cells treated with rapamycin, Hmt1-TAP no longer co-immunoprecipitated with Hmt1-DHFR F[3]. These results suggest that rapamycin prevents the homo-oligomerization and therefore activation of Hmt1 activity.

2.4.5.Hmt1 phosphorylation is essential for its homo-oligomerization

We observed that a fraction of Hmt1 migrated more slowly on SDS-PAGE in the control cells (Hmt1*) but this species is not evident in rapamycin-treated cells (Fig. 22E). We reasoned that the slower-migrating species of Hmt1 could be phosphorylated and that perhaps phosphorylation is required for its oligomerization and set out to identify potential sites of phosphorylation. A clue was provided by the crystal structure [103]. Hmt1 was crystallized as an active hexamer. However, the Hmt1 that was crystallized was, in fact, expressed as a mutant in which the first 20 amino acids were deleted and positions 21 and 22 were muted from Gln-His to Asp-Tyr. This mutant Hmt1 was chosen because it was more soluble and produced crystals of higher diffraction quality than *wildtype* Hmt1 [103]. We hypothesized that the first 20 amino acids could contain regulatory elements, perhaps phosphorylation sites that control homo-oligomeric assembly. Since these regulatory elements are necessary for Hmt1 activity, they should be evolutionarily conserved. We performed an alignment of Hmt1 homologues across the *ascomycetes* fungi (Fig. 22F). We identified five conserved potential phosphorylation sites within or around the first 20 amino acids of which Ser 9 and Tyr 23 showed the strongest prediction to be phosphorylated by NetPhos 2.0 [131] (Fig. 22G).

We investigated the possibility that Hmt1 homo-oligomerization depends on phosphorylation, of Ser 9. We adapted the TAP tag immunoprecipitation assay (Fig. 23A) to include a step to deplete phosphorylated species in purified protein complexes by treatment with alkaline phosphatase (AP) following immunoprecipitation of Hmt1-TAP (Fig. 23A). Protein was applied on IgG resin and treated with active alkaline phosphatase or treated with sodium vanadate to inactivate alkaline phosphatase for 1 hour before a final wash. We eluted and separated IgG immunoprecipitate on SDS-PAGE to reveal co-immunoprecipitated Hmt1-DHFR F[3] (immunoblot with anti-DHFR) or Hmt1-TAP (with the CBP antibody) (Fig. 23B). We observed that alkaline phosphatase-treated Hmt1-TAP immunoprecipitates no longer co-purified with Hmt1-DHFR F[3]. However, simultaneous treatment of immunoprecipitates with sodium vanadate prevented alkaline phosphatase-induced monomerization of Hmt1. These results suggest that oligomerization of Hmt1 requires phosphorylation of at least one subunit.

To specifically determine if Ser 9 is required for Hmt1 phosphorylation and sufficient to mediate oligomerization of Hmt1, we generated phosphomimetic (Hmt1S9E) and non-phosphorylatable (Hmt1S9A) mutants. We generated diploid strains as above to express *wildtype* Hmt1 Hmt1wt-TAP, Hmt1S9E or Hmt1S9A from the first locus and Hmt1wt-DFHR F[3], Hmt1S9E-DFHR F[3] or Hmt1S9A-DFHR F[3] from the second locus and performed our immunoprecipitation assay (Fig. 23B).

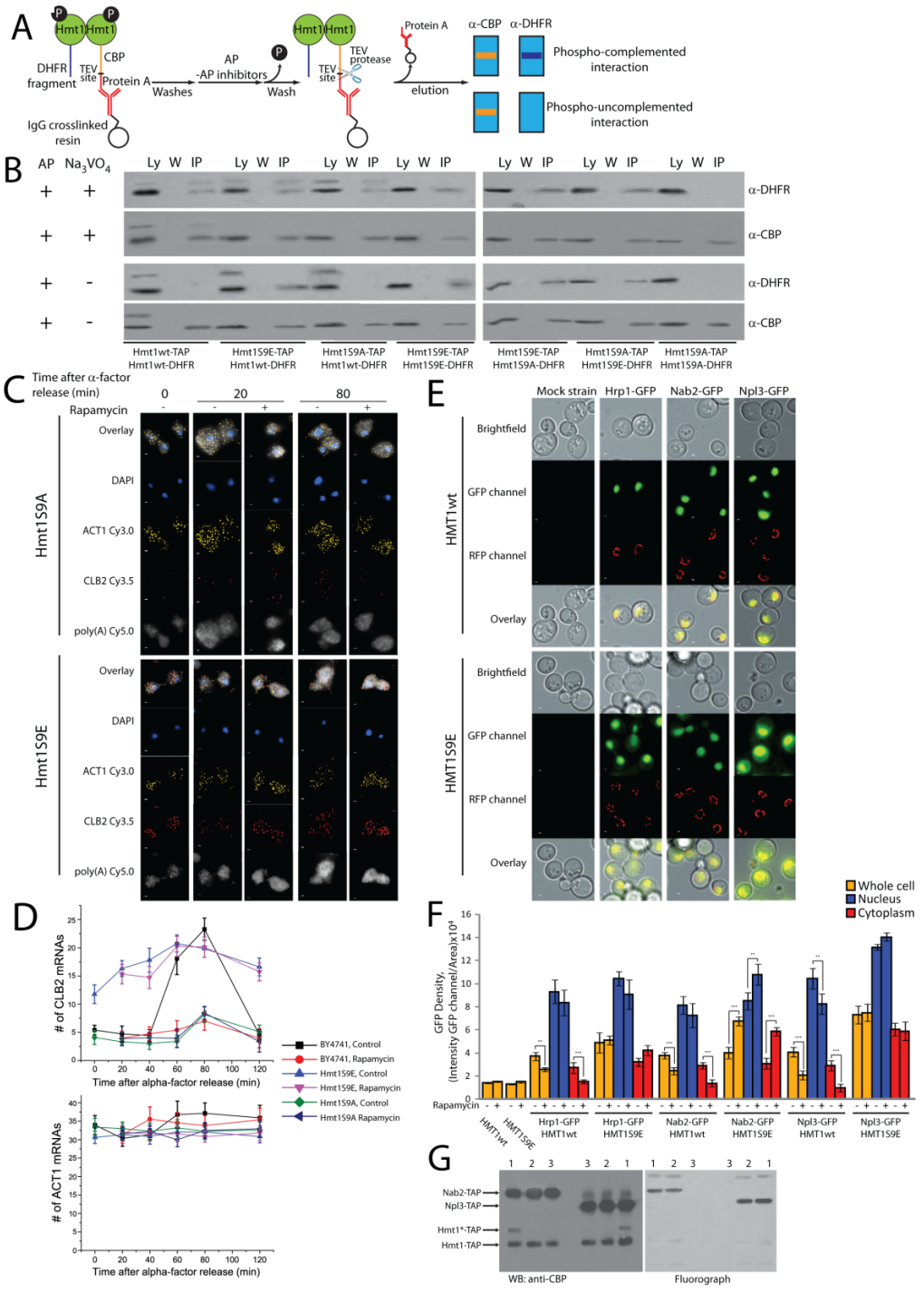

Figure 23: Hmt1 phosphorylation of serine 9 is essential and required for Hmt1 oligomerization, methyltransferase activity and resulting *CLB2* mRNA accumulation in early M-phase and hnRNP nuclear exports.

(A) Flow chart for Hmt1 homomeric immunoprecipitation in diploid yeast cell endogenously expressing Hmt1wt, Hmt1S9E or Hmt1S9A fused to TAP tag on first locus and Hmt1wt, Hmt1S9E or Hmt1S9A fused to DHFR F[3] on the

other locus. IgG immunoprecipitate was treated with alkaline phosphatase or alkaline phosphatase with sodium vanadate before the last wash preceding elution by TEV protease cleavage of Hmt1-TAP. Presence in immunoprecipitate of cleaved Hmt1-TAP and Hmt1-DHFR F[3] revealed phospho-dependent oligomerization **(B)** Diploid yeast cell endogenously expressing Hmt1wt, Hmt1S9E or Hmt1S9A fused to TAP tag on first locus and Hmt1wt , Hmt1S9E or Hmt1S9A fused to DHFR F[3] on the other locus were grown in YPD. For each strain, cells were harvested, lysed and proteins were purified applied to IgG resin. Cell lysates (lane Ly), last washes preceding TEV protease elution of TAP tag protein fusion (lane W) and IgG immunoprecipitates (lane IP) were analyzed by immunoblotting for levels of TAP tag protein fusion, levels and co-immunoprecipitated DHFR F[3] tag protein fusion. **(C)** Strains endogenously expressing Hmt1wt, Hmt1S9E or Hmt1S9A were arrested in G1 phase by alpha-factor and released in cell cycle. Synchronized cultures were treated with rapamycin (200 nM) or ethanol vehicle 20 minutes after alpha-factor released to prevent premature rapamycin induced G1 arrest. A mix of cy3 labeled *CLB2* mRNA probes (in yellow), cy3.5 *ACT1* mRNA probes (in red) and poly(A) containing mRNA probes (in gray scale) were combine to revealed single mRNA of single cell by FISH in cell fixed every 20 minutes for each culture. Cell nucleus was stained by DAPI (in blue) contain in cover slip mounting media. **(D)** Detection and counting of fluorescent mRNAs and single probes was performed for 25 cells treated with rapamycin (200 nM) or ethanol vehicle in each time points per strains for *CLB2* mRNAs (in upper panel) and *ACT1* mRNAs (in lower panel). **(E)** Yeast cells endogenously expressing hnRNPs fused to GFP (in green), Nup49 fused to RFP (in red), as nuclear membrane marker, and overexpressing Hmt1wt or Hmt1S9E were treated with rapamycin (200 nM). Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and overlayed. **(F)** GFP intensities for hnRNPs in Hmt1wt or Hmt1S9E overexpressing background strains were quantified and area normalized for 100 cells for each strain treated with rapamycin (200 nM) or ethanol vehicle (in yellow). Cellular GFP signal was further segmented into nucleus (in blue) and cytoplasm (in red) based on Nup49 nuclear membrane marker. **(G)** Methylation assay was perform with TAP-tagged proteins immunoprecipitated with IgG sepharose beads from asynchronous cultures, fractionated on SDS-PAGE and immunoblotted with anti-CBP (left panel). Npl3-TAP and Nab2-TAP immunoprecipitates were treated with enrich Hmt1wt-TAP (lane 1), Hmt1S9E-TAP (lane 2) or Hmt1S9A-TAP (lane 3) immunoprecipitates in the presence of [methyl-^{3}H]SAM, fractionated on SDS-PAGE and detected by autoradiography (right panel). Scale bar: 1 μm. See also Figure 24.

We observed that both alkaline phosphatase treated and sodium vanadate-inactivated alkaline phosphatase treated Hmt1S9E-TAP immunoprecipitates co-purified with Hmt1wt-DHFR F[3], which demonstrates that phosphomimetic mutation of Ser 9, by itself, is sufficient to elicit oligomerization of Hmt1 in the absence of phosphorylated protein species.

In contrast, alkaline phosphatase treated Hmt1S9A-TAP immunoprecipitates no longer interacted with Hmt1wt-DHFR F[3], but, interestingly, both Hmt1wt-DHFR F[3] forms co-purified in sodium vanadate inhibited alkaline phosphatase treated Hmt1S9A-TAP immunoprecipitate. Therefore, only a third or less of phosphorylated Hmt1 is required to mediate the homo-oligomerization.

We were able to co-purify both Hmt1S9A-DHFR F[3] and Hmt1S9E-DHFR F[3] in Hmt1S9E-TAP immunoprecipitates both in the presence and in the absence of alkaline phosphatase. In addition, Hmt1S9A-TAP immunoprecipitates co-immunoprecipitated Hmt1S9E-DHFR F[3], both in the presence and absence of alkaline phosphatase activity. These results suggest that both unphosphorylated and phosphorylated Ser 9 Hmt1 subunits can oligomerize with a phosphorylated Hmt1 subunit. Thus phosphorylation of at least one Hmt1 monomer may be sufficient to induce dimerization and in turn, nucleate an enzymatically active hexamer.

Finally, we observed that Hmt1S9A-TAP immunoprecipitates did not co-purify Hmt1S9A-DHFR F[3], both in presence and absence of alkaline phosphatase. This result demonstrates that Ser 9 is required for Hmt1 phosphorylation and essential for Hmt1 oligomerization but that Hmt1S9A is a constitutive monomer in cells that express this mutant alone.

2.4.6.Phosphorylation of Hmt1 is required to regulate *CLB2* mRNA stability

Based on our observations thus far, we hypothesized that rapamycin causes a reduction of *CLB2* mRNA stability through a mechanism requiring Hmt1 complex dephosphorylation and disassembly, with consequent loss of methyltransferase activity. This results in nuclear sequestering Hmt1-methylation dependent hnRNPs. Based on this hypothesis, we predicted that Hmt1S9E, would be a constitutively active homo-oligomer that would promote *CLB2* mRNA stability that is insensitive to rapamycin treatment. In contrast, we predicted that Hmt1S9A would be an inactive, constitutive monomer and that *CLB2* mRNA would fail to accumulate in early M-phase, whether cells were treated or not with rapamycin. Therefore, as above, *CLB2*, *ACT1* and poly(A) tail containing transcripts were counted in single cells, by single molecule FISH in strains of different mutant backgrounds [9] (Fig. 23C, Fig. 24A).

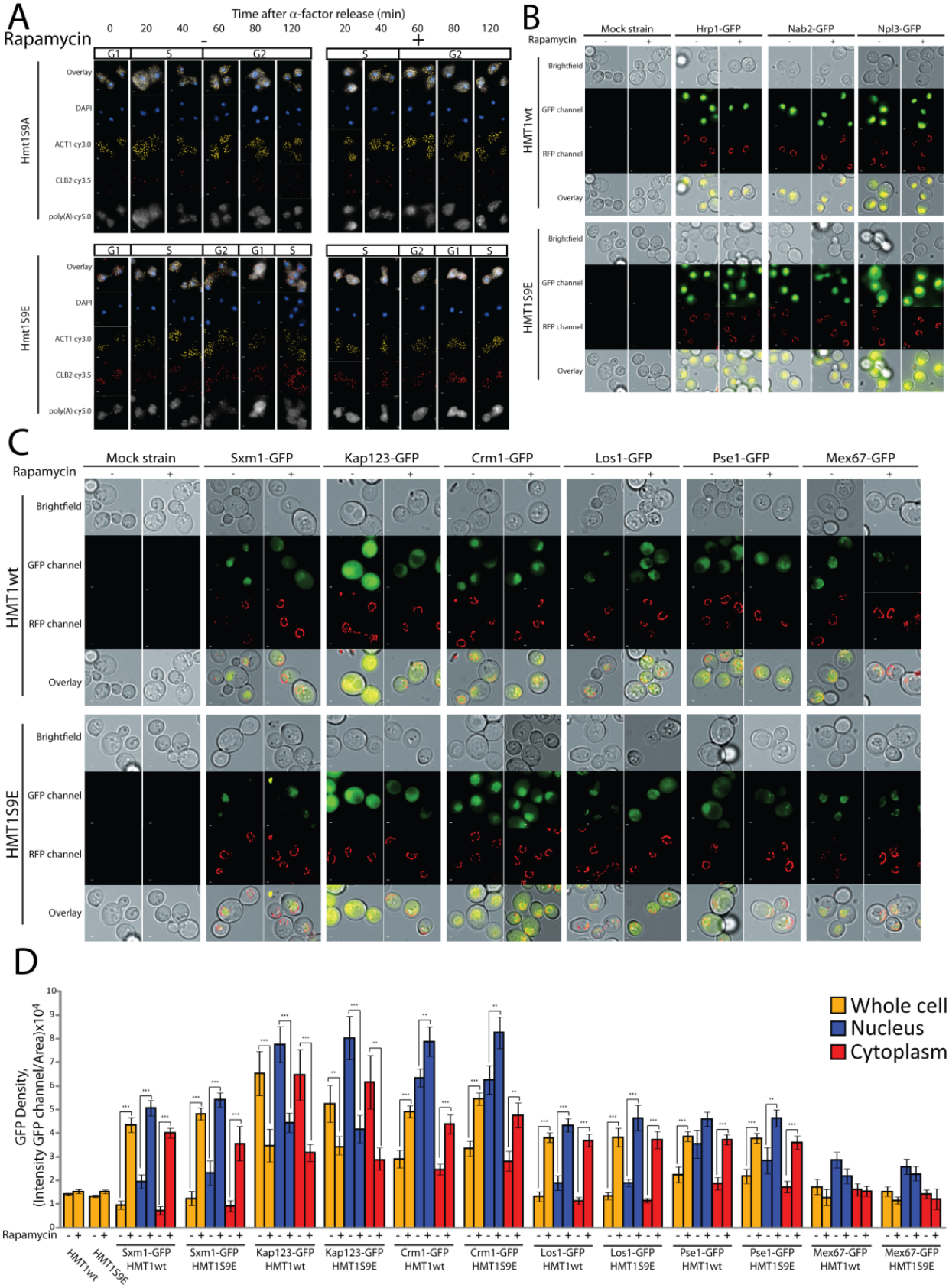
A
Time after α-factor release (min)
Rapamycin
Hmt1S9A
Hmt1S9E
Overlay
DAPI
ACT1 cy3.0
CLB2 cy3.5
poly(A) cy5.0
G1
S
G2
B
Mock strain
Hrp1-GFP
Nab2-GFP
Npl3-GFP
HMT1wt
HMT1S9E
Brightfield
GFP channel
RFP channel
C
Mock strain
Sxm1-GFP
Kap123-GFP
Crm1-GFP
Los1-GFP
Pse1-GFP
Mex67-GFP
D
GFP Density, (Intensity GFP channel/Area)x10^4
Whole cell
Nucleus
Cytoplasm
HMT1wt
HMT1S9E
Sxm1-GFP HMT1wt
Sxm1-GFP HMT1S9E
Kap123-GFP HMT1wt
Kap123-GFP HMT1S9E
Crm1-GFP HMT1wt
Crm1-GFP HMT1S9E
Los1-GFP HMT1wt
Los1-GFP HMT1S9E
Pse1-GFP HMT1wt
Pse1-GFP HMT1S9E
Mex67-GFP HMT1wt
Mex67-GFP HMT1S9E

Figure 24: Hmt1 phosphorylation of serine 9 is essential and required for Hmt1oligomerisation, methyltransferase activity *CLB2* mRNA accumulation in early M-phase and hnRNP nuclear exports.

(A) Strains endogenously expressing Hmt1wt, Hmt1S9E or Hmt1S9A were arrested in G1 phase by alpha-factor and released in cell cycle. Synchronized cultures were treated with rapamycin (200 nM) or ethanol vehicle 20 minutes after alpha-factor released to prevent premature rapamycin induced G1 arrest. A mix of cy3 labeled *CLB2* mRNA probes (in yellow), cy3.5 *ACT1* mRNA probes (in red) and poly(A) containing mRNA probes (in gray scale) were combine to revealed single mRNA of single cell by FISH in cell fixed every 20 minutes for each culture. Cell nucleus was stained by DAPI (in blue) contain in mounting media. **(B)** Yeast cells endogenously expressing hnRNPs fused to GFP (in green), Nup49 fused to RFP (in red), as nuclear membrane marker, and overexpressing Hmt1wt or Hmt1S9E were treated with rapamycin (200 nM) or ethanol vehicle. Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and overlayed. **(C)** Yeast cells endogenously expressing karyopherins and mRNA nuclear export elements fused to GFP (in green), Nup49 fused to RFP (in red), as nuclear membrane marker, and overexpressing Hmt1wt or Hmt1S9E were treated with rapamycin (200 nM) or ethanol vehicle. Cells were adhered on ConA-coated glass bottom 96-well plates, imaged and overlayed. (**D**) GFP intensities for karyopherins and mRNA nuclear export elements in Hmt1wt or Hmt1S9E overexpressing background strains were quantified and area normalized for 100 cells for each strains treated with rapamycin (200 nM) or ethanol vehicle (in yellow). Cellular GFP signal was further segmented into nucleus (in blue) and cytoplasm (in red) based on Nup49 nuclear membrane marker. Scale bar: 1 μm.

Cells endogenously expressing Hmt1S9E or Hmt1S9A were synchronized with α-factor and treated with rapamycin or vehicle 20 minutes after α-factor release. Cells were harvested and fixed every 20 minutes for 80 minutes and at 120 minutes following α-factor release. Fixed cells were hybridized to mRNA probes and *CLB2, ACT1* and poly(A) mRNA were quantified and compared to *wildtype* strain (Fig. 23D). Cells expressing Hmt1S9E exhibited elevated *CLB2* mRNA throughout the cell cycle, in both rapamycin or vehicle treated cells. We counted an average of 12 *CLB2* molecules during G1. The *CLB2* mRNA accumulated until late M-phase to a maximum of 20 transcripts, which is comparable to quantities in vehicle treated *wildtype* strain in the same cell cycle phase. However, cells expressing Hmt1S9A failed to accumulate *CLB2* mRNA throughout the cell cycle in both rapamycin and vehicle treated cells. We counted around 6 *CLB2* transcripts in these cells, independent of cell cycle phase and rapamycin treatment. These observations were specific for *CLB2* transcripts as *ACT1* mRNA remained constant in control and rapamycin-treated cells in both mutant strains, similar to results in *wildtype* cells (see Fig. 22B).

2.4.7.hnRNP nuclear sequestering is dependent on Hmt1 phosphorylation

Since *CLB2* mRNA is stabilized in Hmt1S9E expressing cells, we investigated whether Hmt1S9E expression alone is sufficient to prevent rapamycin-induced nuclear sequestration of hnRNP proteins. We generated

diploid strains harbouring an expression vector for *wildtype HMT1* (Hmt1wt) or HMT1S9E that simultaneously expressed one of each hnRNP components (Nab2, Npl3 and Hrp1) fused to GFP and Nup49 fused to dtTomato as a nuclear membrane marker. Strains were treated with rapamycin or ethanol vehicle for 6 hrs before being adhered onto ConA-coated glass bottom 96-well plates and imaged (Fig. 23E, Fig. 24B). In rapamycin-treated Hmt1wt cells, the GFP signal for all three hnRNP components were sequestered in the nucleus. However, Hmt1S9E cells treated with rapamycin exhibited significant cytoplasmic GFP signal (Fig. 23F). These Hmt1S9E associated effects were specific to Nab2, Npl3 and Hrp1. Other karyopherins, including Sxm1, Kap123, Crm1, Los1, Pse1 and Mex67 (Fig. 24C), showed similar levels of expression in cytosol and nucleus that changed proportionately in both compartments following rapamycin treatment, in strains expressing either Hmt1wt or Hmt1S9E (Fig. 24D).

2.4.8.Methyltransferase activity of Hmt1 is regulated by its phosphorylation

To determine whether Hmt1S9E and Hmt1S9A could methylate substrates *in vitro*, we performed a methylation assay for yeast purified Nab2 and Npl3 with radiolabeled methyl donor, S-Adenosyl-L-[methyl-^{3}H]methionine (^{3}H-SAM) in the presence of yeast purified Hmt1wt-TAP,

Hmt1S9E-TAP or Hmt1S9A-TAP. The expression and stability of each TAP fusion protein was verified by western blot analysis of the methylation reactions using the CBP antibody (Fig. 23G). Proteins were resolved on SDS-PAGE, and labeling was revealed by fluorography. Both purified Nab2 and Npl3 were substrates of Hmt1wt and Hmt1S9E *in vitro*. However, Hmt1S9A did not methylate these substrates. These results confirm that Hmt1wt-TAP and Hmt1S9E-TAP are functional methyltransferases, whereas the catalytic activity of monomeric Hmt1S9A is abolished.

2.4.9.Rapamycin-dependant Hmt1 protein-protein interactions include potential kinases and phosphatases

We thus far determined that hnRNP localization and *CLB2* mRNA stability depends on phosphorylation of Hmt1. We next needed to identify a kinase that could act on Hmt1. Equally, we hypothesized that there must exist a non-canonical TOR pathway with rapamycin-dependent phosphatase activity that dephosphorylates and inactivates Hmt1. Thus, we reasoned that we needed to identify kinases and phosphatases that interact with Hmt1 in a rapamycin-dependant manner. To identify potential candidates as well as any other rapamycin-dependent Hmt1 protein-protein interactions, we used the DHFR survival-selection Protein-fragment Complementation Assay (PCA) [20, 24, 34, 64]. This assay has been used to access the protein-protein interaction network

of *S. cerevisiae* for approximately 92% of its proteins, *in vivo* and expressed at endogenous levels [132]. We, performed the DHFR PCA screen between strains endogenously expressing Hmt1 fused to DHFR fragment F[1,2] and complementary fragment DHFR F[3] fused to a collection of 5,250 ORFs [132]. Diploid yeast strains were printed onto plates containing methotrexate DHFR PCA selective medium and sublethal rapamycin concentration (2 nM) or vehicle. Clones in which an interaction occurs between Hmt1 and another protein results in complementary fragments of DHFR, to which they are fused, being brought together in space where they refold and reconstitute methotrexate-resistant DHFR activity. Thus, simple readout of an interaction between two proteins is survival and proliferation of cells. Colonies were allowed to grow for 5 days at 30°C and resulting plate images recorded with a high-resolution digital camera. Images were analyzed to extract pixel intensity over colony area in both rapamycin and vehicle-containing selective media (Table S1). We calculated a fold-change in interaction for each tested Hmt1 pair-wise interaction of the collection based on the Log_2 ratio of colony intensities of rapamycin *versus* vehicle treated cells (Fig. 25A). We observed 34 proteins whose interaction with Hmt1 changed by at least 8-fold. These proteins were highly enriched in biological processes such as RNA 3'-end processing ($p<10^{-6}$), post-translational protein modification ($p<10^{-4}$) and mRNA processing ($p<10^{-3}$) (Table S2).

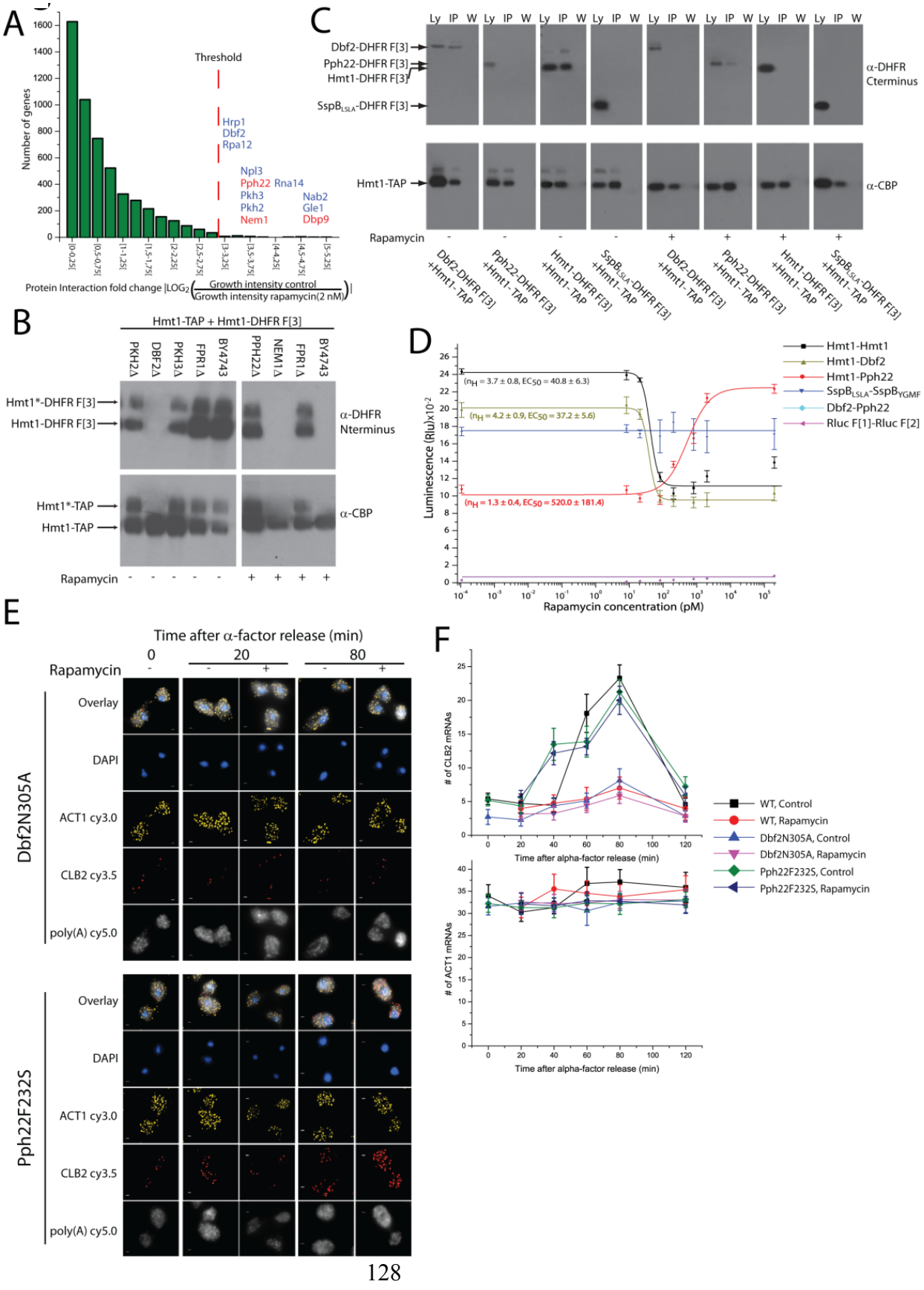
A
Number of genes
Threshold
Hrp1
Dbf2
Rpa12
Npl3
Pph22 Rna14
Pkh3
Nab2
Pkh2
Gle1
Nem1
Dbp9
Protein Interaction fold change |LOG2(Growth intensity control / Growth intensity rapamycin(2 nM))|
B
Hmt1-TAP + Hmt1-DHFR F[3]
PKH2Δ
DBF2Δ
PKH3Δ
FPR1Δ
BY4743
PPH22Δ
NEM1Δ
FPR1Δ
BY4743
Hmt1*-DHFR F[3]
Hmt1-DHFR F[3]
α-DHFR Nterminus
Hmt1*-TAP
Hmt1-TAP
α-CBP
Rapamycin
C
Ly IP W
Dbf2-DHFR F[3]
Pph22-DHFR F[3]
Hmt1-DHFR F[3]
SspB$_{LSLA}$-DHFR F[3]
α-DHFR Cterminus
Hmt1-TAP
α-CBP
Rapamycin
Dbf2-DHFR F[3] +Hmt1-TAP
Pph22-DHFR F[3] +Hmt1-TAP
Hmt1-DHFR F[3] +Hmt1-TAP
SspB$_{LSLA}$-DHFR F[3] +Hmt1-TAP
D
Luminescence (Rlu)x10^{-2}
(n$_H$ = 3.7 ± 0.8, EC$_{50}$ = 40.8 ± 6.3)
(n$_H$ = 4.2 ± 0.9, EC$_{50}$ = 37.2 ± 5.6)
(n$_H$ = 1.3 ± 0.4, EC$_{50}$ = 520.0 ± 181.4)
Rapamycin concentration (pM)
Hmt1-Hmt1
Hmt1-Dbf2
Hmt1-Pph22
SspB$_{LSLA}$-SspB$_{YGMF}$
Dbf2-Pph22
Rluc F[1]-Rluc F[2]
E
Time after α-factor release (min)
0
20
80
Rapamycin
Dbf2N305A
Pph22F232S
Overlay
DAPI
ACT1 cy3.0
CLB2 cy3.5
poly(A) cy5.0
F
# of CLB2 mRNAs
# of ACT1 mRNAs
Time after alpha-factor release (min)
WT, Control
WT, Rapamycin
Dbf2N305A, Control
Dbf2N305A, Rapamycin
Pph22F232S, Control
Pph22F232S, Rapamycin

Figure 25: Dbf2 kinase and Pph22 phosphatase regulate Hmt1 phosphorylation state and related activity on *CLB2* accumulation during early M-phase differentially in rapamycin treated cells.

(A) Diploid yeast strains endogenously expressing Hmt1 fused to DHFR F[1,2] PCA fragment and complementary PCA fragment DHFR F[3] fused to a collection of 5250 (91%) ORFs were screen on solid-agar DHFR PCA selective media supplemented with sub-lethal rapamycin (2 nM) dose or ethanol vehicle. Plates were incubated 120 hrs at 30°C, photographed and analyzed with ImageJ, measuring integrated pixel intensity over each colony area. The fold change of protein interaction reported by the DHFR PCA colonies growth in each condition was computed and the distributions of fold changes were displayed in a histogram. 12 of the 34 strongest changes were presented on top of the appropriate fold change range for both reduced (in blue) and increased (in red) interaction intensities in rapamycin treated reporter strains. **(B)** Diploid yeast cell endogenously expressing Hmt1-TAP on first locus and Hmt1-DHFR F[3] on second locus in PKH2Δ, DBF2Δ, PKH3Δ, FPR1Δ or BY4743 background strains were treated with ethanol vehicle for 6 hrs. In parallel, diploid yeast cell endogenously expressing Hmt1-TAP on first locus and Hmt1-DHFR F[3] on second locus in PPH22Δ, NEM1Δ, FPR1Δ or BY4743 background strains were treated with rapamycin (200 nM) for 6 hrs. For each strain, cells were harvest and lysed to be applied on IgG resin. IgG immunoprecipitates was eluted by TEV protease cleavage of the Hmt1-TAP. IgG immunoprecipitates and cell lysates were analyzed by immunoblotting for levels of Hmt1-TAP, co-immunoprecipitated and levels of Hmt1-DHFR F[3]. **(C)** Diploid yeast cell endogenously expressing Hmt1-TAP and Dbf2, Pph22 or Hmt1 fused to DHFR F[3] or overexpressing a control $SspB_{LSLA}$-DHFR F[3] were treated with rapamycin (200 nM) or ethanol vehicle for 6 hrs. For each strain, cells were harvest, lysed and proteins were purified to be applied on IgG resin. IgG immunoprecipitates were eluted by TEV protease cleavage of the Hmt1-TAP. Cell lysates (lane Ly), IgG immunoprecipitates (lane IP) and last wash (lane W) were analyzed by immunoblotting for levels of Hmt1-TAP, co-immunoprecipitated and levels of DHFR F[3] fused protein. **(D)** *In vivo* analysis of the steady state level of Hmt1 complex with itself with Dbf2 and Pph22 reported by Rluc PCA as a function of rapamycin incubation of 6 hrs. Protein interaction reported by Rluc PCA were controlled over spontaneously interacting partner $SspB_{LSLA}$-$SspB_{YGMF}$, non-interacting partner Dbf2-Pph22 and spontaneous Rluc PCA fragments reconstitution. Error bars, s.e.m. (n=3). **(E)** Strains endogenously expressing Pph22F232S and Dbf2N305A were arrested in G1 phase by alpha-factor and released in cell cycle. Synchronized cultures were treated with rapamycin (200 nM) or ethanol vehicle 20 minutes after alpha-factor released to prevent premature rapamycin induced G1 arrest. A

mix of cy3 labeled *CLB2* mRNA probes (in yellow), cy3.5 *ACT1* mRNA probes (in red) and poly(A) containing mRNA probes (in gray scale) were combine to revealed single mRNA of single cell by FISH in cell fixed every 20 minutes for each culture. Cell nucleus was stained by DAPI (in blue) contain in cover slip mounting media. **(F)** Detection and counting of fluorescent mRNAs and single probes was performed for 25 cells treated with rapamycin (200 nM) or ethanol vehicle in each time points per strains for *CLB2* mRNAs (in upper panel) and *ACT1* mRNAs (in lower panel). Scale bar: 1 μm. See also Figure 26.

Among these, the hnRNP proteins Nab2, Npl3 and Hrp1 showed reduced interaction with Hmt1 in reporter strains treated with rapamycin, along with several kinases, including Dbf2, Pkh2 and Pkh3. We also observed interaction of Hmt1 with two phosphatases, Nem1 and Pph22, but in these cases, the interactions increased in rapamycin. We thus next tested these kinases and phosphatases as potential regulators of Hmt1 phosphorylation.

2.4.10.Hmt1 phosphorylation and oligomerization are regulated by Dbf2 kinase and Pph22 phosphatase

We first tested Dbf2, Pkh2 and Pkh3 as potential Hmt1 kinases and Nem1 and Pph22 potential Hmt1 phosphatase in a genetic screen for both their requirements for Hmt1 phosphorylation and oligomerization. To do so, we generated diploid strains as above to express Hmt1wt-TAP, from one locus and Hmt1wt-DFHR F[3] from the other locus in knockout strains *DBF2*Δ, *PKH2*Δ, *PKH3*Δ, *FPR1*Δ and BY4743 *wildtype* strains treated with vehicle for 6 hrs or *NEM1*Δ, *PPH22*Δ, *FPR1*Δ and BY4743 background strains treated with rapamycin for 6 hrs and performed our immunoprecipitation assay (Fig. 25B). The *FPR1* gene encodes the yeast FKBP12 proteins, which interacts with rapamycin to form a ternary inhibitory complex with TOR kinase [81]. Thus, the *FPR1*Δ strain is insensitive to rapamycin, thus serving as a control for TOR pathway activation. We observed that Hmt1-TAP phosphorylation was

abolished in the *DBF2*Δ strain and Hmt1-TAP immunoprecipitates no longer co-purify Hmt1-DHFR F[3]. This result is specific to the *DBF2*Δ strain as *PKH2*Δ, *PKH3*Δ and *FPR1*Δ all retained Hmt1-TAP phosphorylation and Hmt1-DHFR F[3] co-immunoprecipitating in Hmt1-TAP immunoprecipitates similar to the *wildtype* strain. Repeating the experiments with the phosphatase candidates, we found that rapamycin treated *PPH22*Δ strains retained Hmt1-TAP phosphorylation and co-purified Hmt1-DHFR F[3] in Hmt1-TAP immunoprecipitates. The *FPR1*Δ strain showed similar results, insensitive to rapamycin. These results suggest that Dbf2 and Pph22 are directly or indirectly essential for Hmt1 phosphorylation and dephosphorylation, respectively.

We verified the Hmt1-Dbf2 and Hmt1-Pph22 physical interactions by Hmt1-TAP immunoprecipitation. We generated diploid strains as above to express Hmt1wt-TAP and DFHR F[3] fused to Dbf2, Pph22, Hmt1 and a control protein from *Haemophilus influenzae*, $SspB_{LSLA}$, and performed our immunoprecipitation assay from strains treated with rapamycin (200 nM) or vehicle (Fig. 25C). We observed that Dbf2-DHFR F[3] and Hmt1-DHFR F[3] co-immunoprecipitated Hmt1-TAP immunoprecipitates in cells treated with vehicle only and that Pph22-DHFR F[3] interacted with Hmt1-TAP immunoprecipitates only in cells treated with rapamycin. These results were specific for Hmt1, Dbf2 and Pph22 as $SspB_{LSLA}$-DHFR F[3] did not immunoprecipitate with Hmt1-TAP in cells treated or not with rapamycin. These results and those of the DHFR PCA suggest that both Pph22 and Dbf2

interact directly with Hmt1 in cells treated with rapamycin and vehicle, respectively.

To determine whether Dbf2 kinase could phosphorylate Hmt1 *in vitro*, we performed a kinase assay for yeast purified Hmt1wt, Hmt1S9E and Hmt1S9A with radiolabeled [γ-^{32}P]ATP in the presence of insect cell purified Dbf2-Mob1 or Dbf2N305A-Mob1 kinase dead mutant [133]. Proteins were resolved on SDS-PAGE, and ^{32}P incorporation was revealed by fluorography (Fig. 26A). Hmt1wt was phosphorylated *in vitro* by Dbf2-Mob1. Absence of ^{32}P incorporation in Dbf2N305-Mob1 kinase dead mutant showed specific phosphorylation by Dbf2 and not a co-purified kinase during protein purification both in insect cells and yeast cells. In contrast, Hmt1S9E and Hmt1S9A did not incorporate ^{32}P in both Dbf2-Mob1 and Dbf2N305A-Mob1 kinase dead mutant. These results indicate that Hmt1 can be a direct substrate of Dbf2 and that Ser 9 is required for the unique Dbf2 phosphorylation site on Hmt1.

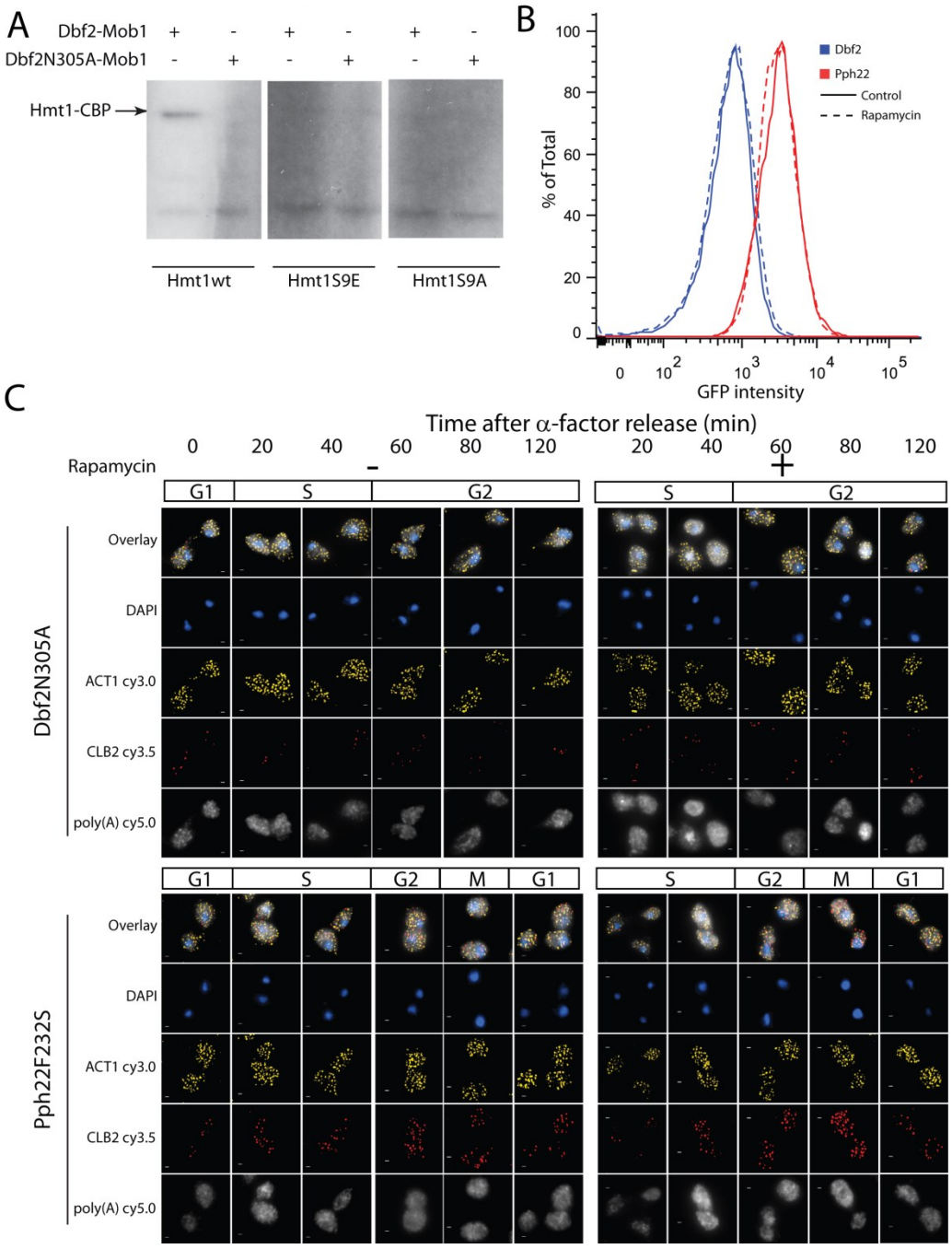

Figure 26: Dbf2 kinase and Pph22 phosphatase regulate Hmt1 phosphorylation state and related activity on *CLB2* accumulation during early M-phase differentially in rapamycin treated cells.

(A) Dbf2 kinase assay was perform with Hmt1wt, Hmt1S9E and Hmt1S9A mutants TAP-tagged proteins immunoprecipitated with IgG sepharose beads from asynchronous cultures. Each Hmt1 species were treated with insect purified Cdc15 activated Dbf2-Mob1 kinase complex or Dbf2N305A-Mob1 kinase-dead mutant in the presence of [γ-^{32}P]ATP, fractionated on SDS-PAGE

and detected by autoradiography.(**B**) Yeast cells endogenously expressing Dbf2 or Pph22 fused to GFP were treated with 200 nM rapamycin (dotted line) or ethanol vehicle (plain line) 6 hrs. GFP intensities per cell in culture treated with rapamycin or ethanol vehicle were record in a LSRII FACS for 30, 000 cells and distribution were displayed in bar graph. (**C**) Strains endogenously expressing Pph22F232S and Dbf2N305A were arrested in G1 phase by alpha-factor and released in cell cycle. Synchronized cultures were treated with rapamycin (200 nM) or ethanol vehicle 20 minutes after alpha-factor released to prevent premature rapamycin induced G1 arrest. A mix of cy3 labeled *CLB2* mRNA probes (in yellow), cy3.5 *ACT1* mRNA probes (in red) and poly(A) containing mRNA probes (in gray scale) were combine to revealed single mRNA of single cell by FISH in cell fixed every 20 minutes for each culture. Cell nucleus was stained by DAPI (in blue) contain in mounting media. Scale bar: 1 μm.

2.4.11.Rapamycin induces cooperative disassembly of Dbf2-Hmt1 and Hmt1-Hmt1 complexes but graded assembly of Hmt1-Pph22

We measured the levels of the Hmt1-Hmt1, Hmt1-Dbf2 and Hmt1-Pph22 complex with another *in vivo* PCA based on the *Renilla reniformis* luciferase (Rluc) as a reporter that allows for direct detection of interactions among proteins expressed endogenously without significantly altering their binding kinetics [21, 134]. Direct measurements of the steady state levels of the Hmt1-Hmt1 and Hmt1-Dbf2 protein complexes showed a switch-like dissociation of the complex in cellular response to rapamycin with Hill coefficients of 3.7 ± 1.7 and 3.6 ± 0.4, respectively, and half-maximum effective concentration (EC_{50}) of 34.8 pM and 36.1 pM (Fig. 25D). The Hmt1-Pph22 protein complex revealed a non-cooperative binding in response to rapamycin with Hill coefficient of 1.4 ± 0.5 and half-maximum effective concentration (EC_{50}) of 538.4 pM. These protein complex dynamics were specific as the constitutive $SspB_{LSLA}$-$SspB_{YGMF}$ interaction did not change over all rapamycin concentrations tested. As controls for spontaneous complementation of the Rluc fragments, we tested the fragments alone or fused to non-interacting protein, Dbf2-Pph22, and observed only background luminescence [21, 134].

We reasoned that changes in abundance of Dbf2 and Pph22 in response to rapamycin could contribute to changes in measured steady-state interactions. Using yeast strains expressing endogenous Dbf2 or Pph22 fused to GFP, we treated cells with or without rapamycin for 6 hours. The cells were analyzed by FACS for integral GFP intensities per cell (Fig. 26B). The quantities of both Dbf2 and Pph22 were constant in cells treated with rapamycin or vehicle. These results suggest that other molecular events trigger change in Dbf2 and Pph22 binding to Hmt1.

2.4.12.Dbf2 activity increases and Pph22 activity decreases *CLB2* mRNA stability

To establish the roles of Dbf2 kinase and Pph22 phosphatase activities on *CLB2* mRNA stability, we reasoned that a strain harbouring a Dbf2N305A kinase-dead mutant would result in *CLB2* mRNA destabilization and a strain expressing a Pph22F232S temperature-sensitive mutant should allow for *CLB2* mRNA accumulation in rapamycin treated cells. Therefore, *CLB2*, *ACT1* and poly(A) tail containing transcripts were counted in individual cells by single molecule FISH in strains exogenously expressing these mutants. We synchronized the Dbf2N305A or Pph22F232S strains with α-factor and treated them with rapamycin or vehicle for 20 minutes after α-factor release. Dbf2N305A mutant cells were grown at 30°C, while Pph22F232S mutant cells

were grown at the restrictive temperature of 37°C. Cells were harvested and fixed every 20 minutes for one cell cycle and after 120 minutes following α-factor release (t = 0) (Fig. 25E, Fig. 26C). Fixed cells were hybridized to mRNA probes and *CLB2, ACT1* and poly(A) mRNA were quantified and compared to *wildtype* strain (Fig. 25F). We observed that the Dbf2N305A mutant strain failed to accumulate *CLB2* transcript even in the absence of rapamycin treatment. In contrast, the Pph22F232S mutant strain grown at the restrictive temperature accumulated *CLB2* mRNA quantities comparable to the *wildtype* strain treated with vehicle throughout all of the cell cycle. Interestingly, the Pph22F232S mutant strain resumed accumulating *CLB2* mRNA when treated with rapamycin.

2.4.13.Failure to accumulate *CLB2* mRNA correlates with early M-phase progression delay

We envisioned that failure to accumulate *CLB2* mRNA would affect normal cell cycle progression. We analyzed the DNA content in synchronized strains treated with rapamycin or vehicle added 20 minutes after α-factors release. We used *cdc15-2* mutant strain background and the same mutant strains that express genomic mutations Hmt1S9E, Hmt1S9A, Dbf2N305A or Pph22F232S at restrictive temperatures (i.e., limiting cell cycle from G1 to telophase, the point of arrest of *cdc15-2* mutants), fixed them and stained either

DNA, with propidium iodide (PI), or mitotic spindle, with Tub1 specific immunostaining [135]. Cells were analyzed for the onset of budding (Fig. 27A-B) and subsequent spindle dynamics (Fig. 27C-F). The DNA content of cells were studied by FACS for integral PI intensities per cell (Fig. 27G). We observed a 30 minute early M-phase transition delay in the rapamycin-treated *cdc15-2* mutant background strain, hindering metaphase-anaphase spindle emergence in comparison to vehicle-treated cells. In contrast, spindle assembly was accelerated in Hmt1S9E whether or not the strain was treated with rapamycin. Both Dbf2N305A and Hmt1S9A exhibited delayed progression through early M-phase after both rapamycin and vehicle treatments. The rapamycin-treated Pph22F232S strain progressed through mitosis with comparable rates to the vehicle-treated *cdc15-2* background strain.

2.5.Discussion

Our understanding of cell cycle regulation have advanced recently due to careful dissection of cell populations into those having different lengths of mitotic phases (e.g. newborn daughter *versus* mother cells) and applications of fluorescent reporters and single cell analysis to define specific regulatory cues in time [105, 136].

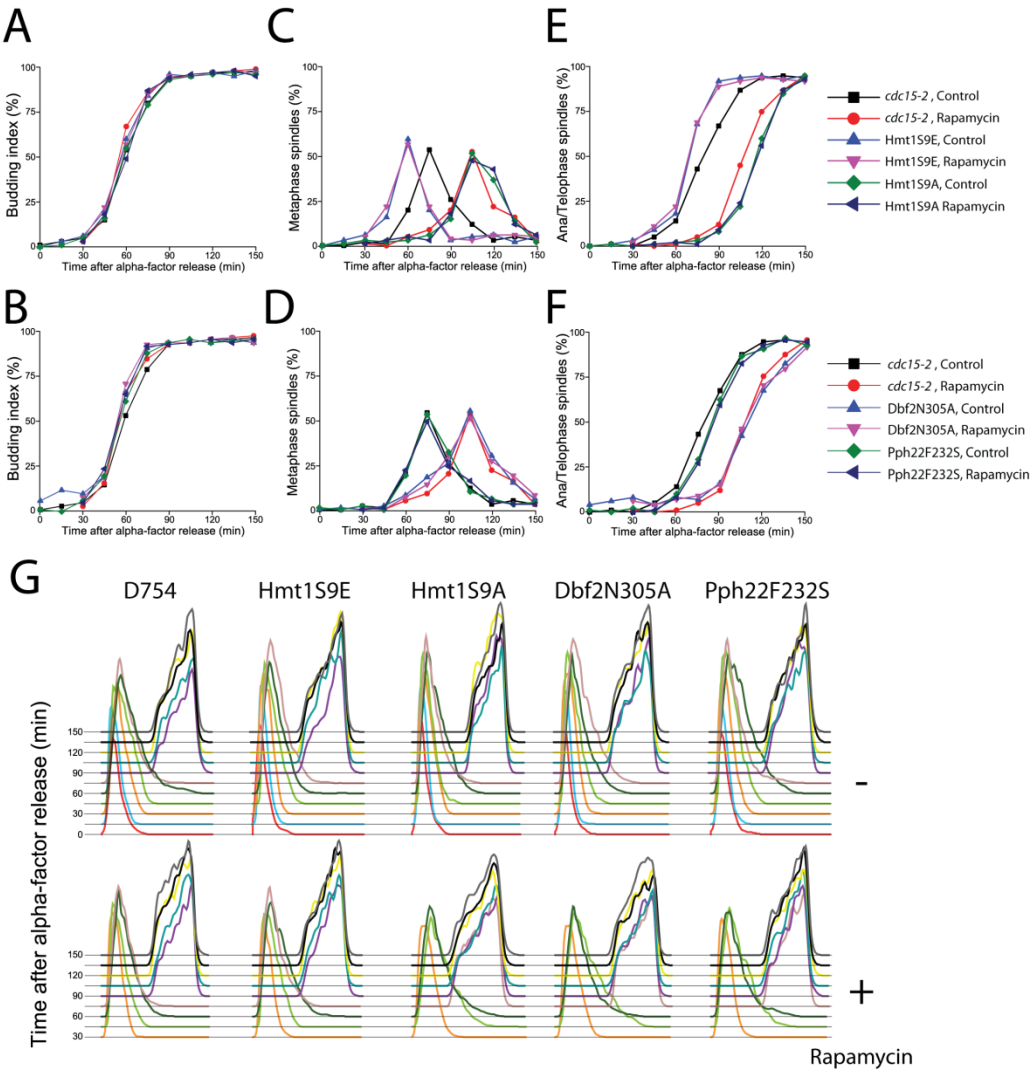

Figure 27: Hmt1 regulations are necessary for M phase progression and induced early M phase delayed in rapamycin treated cells.

Temperature sensitive strains *cdc15-2* endogenously expressing Hmt1wt, Hmt1S9E, Hmt1S9A, Dbf2N305A and Pph22F232S were arrested in G1 phase by alpha-factor and released in cell cycle. Synchronized cultures were treated with rapamycin (200 nM) or ethanol vehicle 20 minutes after alpha-factor released to prevent premature rapamycin induced G1 arrest. Cells were harvest, fixed and the proportion of budded cells (A-B), cells with metaphase spindles (C-D), anaphase/telophase spindles (E-F) and cells DNA content was stained with propidium iodide (PI) (G) measured by FACS analysis.

Equally, the need to understand the role of mRNA regulation in general and cell cycle progression specifically has been advanced by measurement of the spatial and temporal dynamics of single molecules of mRNA following transcription [93]. These developments have set the stage for dissecting regulatory pathways that may control cell cycle progression by regulating mRNA dynamics.

Here we demonstrate a strategy to identify proteins that regulate oscillating mRNAs during mitosis based on proteome-wide measurements of changes in protein-protein interactions *in vivo* [132, 134]. The same strategy could be applied to identify proteins involved in any cellular process. Here we first used the strategy to screen for interactions of Hmt1 with kinases and phosphatases by treating cells with rapamycin, a compound known to delay the cell division in early M-phase. These efforts resulted in our identification of Dbf2 and Pph22. We then characterized the dynamics of Hmt1 assembly and activation. Next, since asynchrony in the cell population would prevent us from detecting small differences in the dynamics of *CLB2* transcript degradation, we measured mRNA decay in single cells and extended the methods to allow us to study mRNA dynamics in large populations of cells [93]. Results of these efforts allow us to propose a general model by which normal progression or stress-dependent delay of early M-phase is mediated by phosphorylation, oligomerization and activation of the arginine methyltransferase Hmt1 (Fig. 28A).

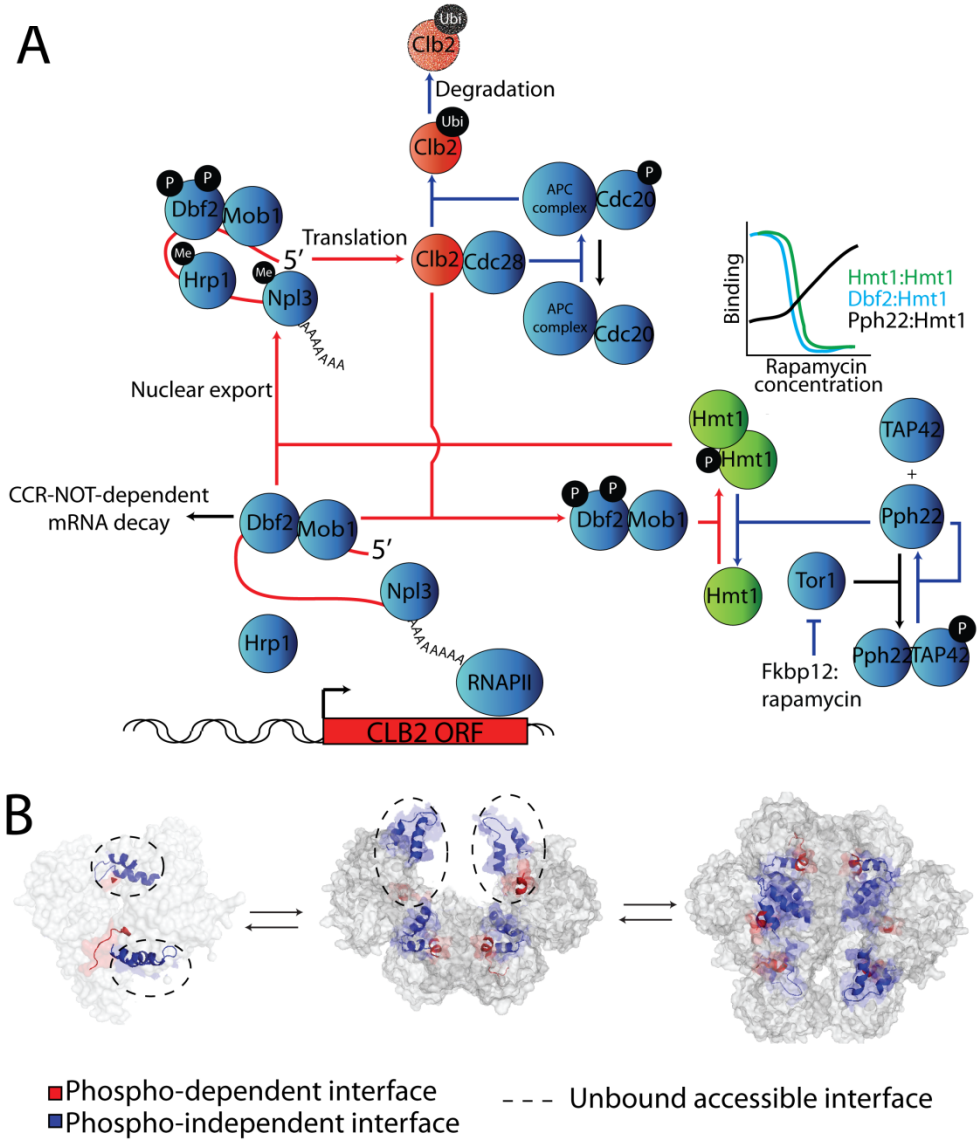

Figure 28:Proposed Models

(A) Model of mechanism regulating *CLB2* mRNA accumulation during early M-phase. (B) Model of phospho-dependent Hmt1 cooperative assembly.

Our model predicts that during early M-phase Dbf2 kinase phosphorylates Hmt1 and induces its cooperative assembly and activation. Active Hmt1 then methylates hnRNP components, which transport and stabilize nascent message of the cyclin *CLB2*. A positive feedback loop is propagated by *CLB2* mRNA being translated into cyclin B2, thus associating with Cdk1 to promote entry into mitosis. Cdk1-Clb2 upregulates Cdc5 kinase activity [137, 138]. The activated polo-kinase, Cdc5, phosphorylates and activates Cdc15 kinase and both kinases phosphorylate Dbf2-Mob1, increasing Dbf2 kinase activity during early M-phase and as a result increasing Hmt1 activity [139-144]. Feedback is abolished by proteasome-mediated proteolysis of Clb2 through activation of the APC complex at the end of M-phase. Simultaneously, activated PP2A dephosphorylates Hmt1, leading to reduction in *CLB2* mRNA stability [145]. As discussed below, however, given the cooperative activation of Hmt1 by Dbf2, *versus* graded deactivation by Pph22, constitutive Dbf2 kinase activity could activate Hmt1 throughout the cell cycle if there were small decreases in Pph22 activity. This could provide a means by which progression of the cell cycle can be tuned.

In addition to proposing a mechanism for normal mitosis progression, our results can explain how early M-phase attenuation may occur in cells treated with rapamycin. It is known that PP2A is activated following release of its catalytic subunit Pph22 from the TAP42 complex when rapamycin-FKBP binds to TOR kinase during the mitosis transition and under a variety of stress

[84-87]. In our model, activated PP2A would shift Hmt1 equilibrium from its active hexameric form to its inactive monomeric state, resulting ultimately in destabilization of *CLB2* mRNA and therefore attenuation of entry into M-phase. This basic circuit could also potentially be the target of a variety of stress-responsive kinases and phosphatases. Our genome-wide screen for Hmt1 revealed protein-protein interactions with alternative kinases and phosphatase including the MAP kinases Pkh2 and Pkh3, which are implicated in cell wall integrity [146]. These enzymes have been directly implicated in mRNA decay and mRNA inclusion in P-bodies [147]. Both of these phenomena may be mediated, at least in part, by concerted Pkh2 and Pkh3 association with Hmt1 and ribonucleoprotein complexes during their transport and stabilization. Furthermore, we observed an interaction between Hmt1 and Nem1, a phosphatase implicated in nuclear membrane biogenesis, conserved from yeast to humans and required for sporulation in yeast [148]. A connection between these two processes remains to be discovered. It would also be interesting to investigate whether Nem1 dephosphorylation of Hmt1 is required for sporulation [149].

Our findings are consistent with recent evidence for roles of Dbf2 in stabilizing *CLB2* transcripts during early M-phase, the associations of Dbf2, Hmt1 and hnRNP components with gene loci and the functions of hnRNP proteins in nuclear export and stabilization of mRNA in the cytoplasm [93, 96]. It was suggested by Trcek, *et al.* that Dbf2 kinase activity is dispensable to

CLB2 transcript stability, which appears contradictory to our results. We can resolve this discrepancy. Trcek, *et al.* performed their experiments by using another mutation of the "kinase-dead" Dbf2 mutant. It has, however, been reported that this mutant has residual kinase activity *in vitro* [133, 150]. Our experiments were performed with the kinase dead mutation of Dbf2 integrated into the genome and therefore expressed at endogenous levels. We know that the activity of Dbf2 during early M-phase is minimal, not reaching a peak until later in M phase [139-141]. We thus argue that the conclusions of Trcek, *et al.* are correct in that Dbf2 does bind to the *CLB2* locus and regulates mRNA accumulation, but Dbf2 activity is not dispensable to regulation of mRNA stability by hnRNPs.

Our observation that inactivation of Hmt1 activity does not lead to the retention of *CLB2* mRNAs in the nucleus, but to a failure to accumulate in the cytoplasm, indicates that Hmt1 methylation of hnRNP proteins does not primarily regulate *CLB2* mRNA nuclear export, but rather, primes the mRNA for regulation in the cytoplasm (e.g. degradation). Consistent with these observations, two studies have shown that arginine methylation by Htm1 acts as a signal for the co-transcriptional assembly and maturation of mRNPs [96, 151]. Nuclear priming of later mRNA regulatory events by co-transcriptional assembly of hnRNP components is a common step prior to, as examples, the deposition of the exon-exon junction complex for cytoplasmic mRNA quality control or in assembling mRNPs for cytoplasmic mRNA localization [152,

153]. Presently, we do not know the protein composition of a *CLB2* mRNP complex, but it is likely that its composition changes as a result of Htm1 methylation. Consistent with our observation that different hnRNP components fail to accumulate in the cytoplasm when Hmt1 is inhibited, the lack of methylation of Nab2, Npl3 and Hrp1 may prevent their binding to *CLB2* mRNA. Alternatively, it is possible that the association of other proteins to *CLB2* mRNA is dependent on their methylation by Hmt1. Interestingly, a recent study showed that Htm1 activity is required for the association of the Nab2 and Hpr1 with the Ccr4-Not complex, a multiprotein complex implicated in transcription as well as cytoplasmic mRNA degradation and, therefore, a prime candidate to be actively involved in the regulated degradation of *CLB2* mRNA [128]. Further studies will show which of the components on the *CLB2* mRNP are actively implicated in regulated degradation of *CLB2* mRNA in the cytoplasm.

We demonstrated that Ser 9 is required for Hmt1 phosphorylation and that dimerization rapidly nucleates unphosphorylated subunits, resulting in generation of the active hexameric Hmt1 complex (Fig. 28B). Thus, only one of two subunits needs to be phosphorylated in order to induce dimers and tetramers. Our results suggest that the phospho-dependent dimerization interface allows phosphorylated Hmt1 to assemble into a dimer. The Hmt1 dimer subsequently assembles with non-phosphorylated subunits, resulting in a hexameric complex through a second, phospho-independent interaction

interface. Our observations are consistent with gel filtration data showing higher stability of Hmt1 in its dimeric and hexameric forms [103]. Although the canonical consensus sequence for Dbf2 kinase (R-X-X-S) is not consistent with the sequence preceding Ser 9, the pattern is similar to the consensus for Dbf2 human homologs, which can accommodate both a R and K at position -3 [154]. Alternatively, another serine could be phosphorylated by Dbf2 in a Ser 9-dependant manner [133]. There is, however, only one Dbf2 kinase consensus sequence in Hmt1, Ser 45, and based on examination of the crystal structure, the side chain of this residue is buried and therefore inaccessible [103].

Recent studies have demonstrated that the sharpness and coherence of cell cycle transitions are regulated by transcriptional and biochemical positive feedback loops and notably those in which at least one step in the loop is ultrasensitive can result in bistable transitions [106, 136, 155-158]. The model we describe is a positive feedback loop in which Dbf2-mediates cooperative assembly and activation of Hmt1, contributing to the control of the early M-phase transition and exit from mitosis. The ultrasensitivity was evident in rapamycin-induced disassembly of Dbf2-Hmt1 and Hmt1-Hmt1 complexes. This regulatory design enables a sharp tunable response to stresses that delay the cell cycle, while maintaining to a near constant, the requirement for a threshold level of active Clb2-Cdk1 complexes to allow cell cycle progression. The Hmt1 phosphorylation-dependent cooperative assembly adjusts the rate of development of Clb2-Cdk1 activity in response to both stress and cell cycle

cues by restraining *CLB2* mRNA availability. These predictions are consistent with the accelerated early M-phase cycling transitions in *CLB2* mRNA in the constitutively active Hmt1S9E expressing mutant strain. The constitutively inactive Hmt1S9A mutant simply showed no or very low levels of *CLB2* mRNA, resulting in a dramatic disruption of early M-phase transition. The effect of the Pph22 phosphatase inactive mutant resulted in recovery of the sharpness of early M-phase transition in rapamycin-treated cells. These results are consistent with the non-cooperative behavior of the Pph22-Hmt1 interaction that we observed. This model does not exclude the possibility that Hmt1 can be dephosphorylated by other phosphatases.

Our results suggests that the competition between Pph22 and Dbf2 for Hmt1 phosphosite could easily happen without any change in the quantities of these enzymes as occurs for Dbf2 during M phase. Specifically, we showed that first, Hmt1 oligomerization and activation involves substoichiometric quantities of Hmt1. Only one subunit of Hmt1 need be phosphorylated in order to generate an active hexamer. Second, the self-oligomerization of Hmt1 and binding to Dbf2 is highly cooperative. In contrast, the association of Pph22 with Hmt1 is simply graded. This means that a small change in Pph22 activity would lead to a large shift in the equilibrium of Hmt1 towards the phosphorylated state. Any small perturbation of Pph22, such as via inhibition of the TOR pathway by rapamycin, can induce a dramatic decrease in Hmt1 activity and thus mRNA stabilization.

Finally, our results suggest that yeast have evolved a mechanism to adapt to a nutrient limited environment, mimicked by rapamycin-induced TOR inhibitions, by controlling the stability of mRNAs. Specifically, it is notable that cells respond to rapamycin without delaying cell cycle progression in the Pph22 phosphatase dead mutant strain supporting our model for early M-phase attenuation by rapamycin. Thus, the dephosphorylation of Hmt1 by Pph22 dominantly leads to a decrease in the rate of early M-phase progression. The cells thus benefit from an extended early M-phase transition, allowing them to adapt to starvation by having the time to synthesize the necessary materials to generate viable daughter cells. We know that the stabilization of many RNAs by hnRNPs are under the control of Hmt1 including ribosomal and other cyclin and signalling protein mRNAs [96, 124, 159]. The current model for G1 arrest in rapamycin-treated and nutrient starved yeast, implicating Cln3 and CLN3 mRNA down-regulations, is consistent with our Hmt1 regulation of mRNA stabilisation by hnRNPs. In fact, there are no convincing observations that would place the G1-cyclin tuning outside of a post-transcriptional regulation. However, there are ample evidences that both the G1-cyclin mRNA and ribosome rRNAs are regulated through their stability upon rapamycin-induced TOR inhibition and nitrogen starvation [124, 125, 159]. It is, also, described that untranslated regions, such as 5'UTR, play a role as switching for more stable 5'UTR stabilized the transcripts, that would be predicted by our hnRNP-dependant model [124]. Synthesis of ribosomal RNA is a limiting factor in

regulation of growth and cyclins and signalling proteins regulate responses to nutritional and other stresses. It is conceivable and worth investigating whether the Pph22-Dbf2 competition could be modulated under different conditions and at different phases of the cell cycle. Pph22 and Dbf2 are conserved from yeast to humans and thus it would be interesting to explore whether the same mechanism of cell cycle control is also conserved.

2.6.Acknowledgement

We thank Tatjana Trcek and Robert H Singer (Albert Einstein College of Medicine) for *CBL2* and *ACT1* FISH probes, and sharing data prior to publication, Raymond J Deshaies and Tiffany KJ Chang for providing us with Dbf2 expressing insect cell lines, and Damien D'Amours for the gift of D754 strain and Tub1 immunostaining reagents. Tatjana Trcek, Robert H Singer and Damien D'Amours for valuable discussions and comments. DZ is supported by Canadian Institutes of Health Research MOP-BMB-232642, the Canadian Foundation for Innovation and holds a FRSQ Chercheur Boursier Junior I. The authors acknowledge support from CIHR grants MOP-GMX-152556 and MOP-GMX- 231013 and NSERC of Canada grant 194582 (to SWM).

2.7.Experimental Procedures

2.7.1.Yeast strains, plasmids, growth conditions and buffers

Growth conditions

All strains used in this study (Table S3) were derived from BY4741, BY4742, BY4743 and D754 genotypes. Each of the strains from the GFP library and TAP tagged library were grown overnight at 30°C in 400 µl SC-His. The GFP strains were mated to BY4742 harbouring endogenously expressed Nup49-dtTomato in YPD for 6 hours and resulting diploids were selected on agar containing SC-His + G418 plates. The TAP-tagged strains were mated to BY4742 harbouring endogenously expressed DHFR F[3] tagged protein in YPD for 6 hours and resulting diploids were selected on agar containing SC-His + hygromycin B plates. Cells auto-fluorescence and non specific protein affinity to IgG resin was measured using BY4743 diploid strain. The onset of budding and spindle formation was compared between *cdc15-2* background strain and other mutant strains, as previously described in [160]. Unless otherwise specified, strains were grown in YPD supplemented with 200 nM rapamycin (2 mM rapamycin stock solution in ethanol, Bioshop) or ethanol vehicle (001 % v/v) for 6 hrs.

Cell using α-factor

Cell synchronization in G1 was obtained by treatment of exponentially growing Mat*a* cells with 10 μM α-factor mating pheromone for 2 h in acidified YPD (pH 3.5). The G1-arrested cells were washed with distilled water three times, released into fresh YPD. Cells were treated with 200 nM rapamycin or ethanol vehicle (Control) 20 minutes after α-factor release to prevent rapamycin induced G1 arrest [107]. Cells were collected at the indicated times after α-factor release for subsequent assay.

Plasmids generation by gateway system and mutated by site-directed mutagenesis

Hmt1, Dbf2, and Pph22 expression vectors were created using Gateway cloning technology. The BP reaction was performed on Hmt1, Dbf2 and Pph22 expression clone of yeast expression Gateway collection (YSC3867, OpenBiosystems). The expression clones (100 ng) were mix to the entry vector (100 ng). The BP reaction between the two plasmids were carried out in TRIS-EDTA buffer (pH 8) with one unit of BP clonase II mix (Invitrogen) at room temperature for a hour. A 2 μL aliquot was transform in DH5α. The *HMT1*, *DBF2* and *PPH22* entry clone transformants were selected on kanamycin-

containing broth (LB) agar plates. The entry clones were used in the LR reaction with the p413GPDpromoter–TAP destination vector [161]. The entry clones (100 ng) were mix to the destination vector (100 ng). The LR reaction between the two plasmids was carried out in TRIS-EDTA buffer (pH 8) with one unit of LR clonase II mix (Invitrogen) at room temperature for an hour. A 2 μL aliquot was transform in DH5α. The Hmt1wt, Dbf2 and Pph22 expression clone transformants were selected on ampicillin-containing broth (LB) agar plates. To create the Hmt1S9E, Hmt1S9A, Dbf2N305A and Pph22F232S mutant expression vector, site-directed mutagenesis was carried out by using standard site-directed mutagenesis protocol (Quick Change Site-Directed Mutagenesis Kit- Stratagene # 200519) following the manufacturer instructions on the p413GPDpromoter-ORF-TAP expression clone. The presence of correct mutation was verified by sequencing. Hmt1wt and Hmt1S9E expression vectors were transformed as describe in [59], and transformants were selected on agar-containing SC-His plates.

Homologous recombination generated strains

The strains BY4741 (*MAT***a** *his3Δ leu2 Δ met15 Δ ura3 Δ*), BY4742 (*MATα his3 Δ leu2 Δ lys2 Δ ura3 Δ*) and D754 (*MAT***a** *his3Δ leu2Δ met15Δ ura3Δ ade2Δ trp1Δ can1Δ rad5Δ ade1* [PSI+] *cdc15-2*) were transformed as described in [59] with PCR products amplified from the templates described in

[132, 134] to create specific homologous recombination DNA fragments required to generate Rluc PCA fragments and DHFR F[3] tagged strains (Suppl. Table S3). Whereas *hmt1*Δ, *dbf2*Δ and *pph22*Δ and strains were transformed with the lithium acetate protocol with PCR products amplified from template Hmt1, Dbf2 and Pph22 mutant fused to TAP cDNA expression vector (Suppl. Table S3), respectively, to create specific homologous recombination DNA cassettes required to replace KANMX locus by mutated ORF-TAP. The transformation protocol was adapted to miniaturized volume as follows: 8 µl of PCR product was mixed with 20 µl of chemically-competent yeast, mixed with 72 µl polyethyleneglycol (PEG) buffer and incubated for 30 minutes at room temperature. Cells were then heat shocked with 10 µL DMSO for 15 minutes at 42°C in a water bath. The transformation buffer was replaced with 250 µl YPD and cells were left to recover for 4 hours at 30°C. Transformants were selected on antibiotic containing YPD agar plates, for PCA fragments homologous recombination, or on SC-His agar plates, for mutant ORF-TAP replacement of KANMX locus, and allowed to grow for 4 days at 30°C. In all cases the BY4741 (*MAT***a**) strain was transformed with the Rluc F[1]-NAT1 cassettes and BY4742 (*MATα*) with the DHFR F[3]-HPH or Rluc F[2] cassettes. The transformed strains were grown on YPD agar plates plus appropriate antibiotic agar plates (100 µg/mL nourseothricin for *MAT***a** transformed strains (WERNER BioAgents, Jena, Germany) or 250 µg/mL hygromycin B for *MATα* transformed strains (Wisent Corporation, Quebec, Canada). Identification of successfully recombinant clones was performed as

follows: Putative recombinant clones were picked by hand, cell lysis was performed by heat treatment and confirmation of the correct location of a genome insertion was determined by PCR using the cassette 3' diagnostic primers described in (Suppl. Table S4). The yeast lysates were placed in PCR tubes and mixed with the ORF specific and cassette specific diagnostic primers. Annealing of the diagnostic primers was performed at 56°C and 35 PCR cycles were carried out with the regular Taq DNA polymerase (Bioshop Canada, Montreal, Canada), with elongation cycles of 1.5 minutes at 72°C, in a 50 µl total reaction volume. Aliquots of 30 µl from each PCR product were mixed with brilliant blue tainted glycerol and gel electrophoresed on ethidium bromide stained 1% agarose gels. Correct recombination was confirmed based on sizes of PCR products. The confirmed strains were glycerol-stocked in pre-assigned positions of 96-well plates and stored at -80°C.

2.7.2.FISH probes and procedures

Probes are described in Suppl. Table S5. For each gene, 5 probes were used, where each probe was labelled with > 90% labelling efficiency. Design, synthesis and labelling of probes were performed as described previously (Femino et al., 2003; Zenklusen et al., 2008). *ACT1* and *CLB2* mRNA were quantified by counting individual FISH revealed mRNA foci after subtracting

background intensities and four time single probe intensity and applying hat kernel to increase contrast.

2.7.3.Strains, plasmids, growth conditions, buffers and primers

Suppl. Table 3 and 4 describe strains, plasmids, media and buffers used in this study, their synchronization protocols and growth conditions. Primers used to generate strains and diagnostic primers are listed in Suppl. Table 5.

2.7.4.Single cell, single molecule Fluorescence *in situ* Hybridization (FISH)

Suppl. Table 5 describes FISH probes used in this study and the FISH procedures.

2.7.5.Fluorescence-Activated Cell Sorter (FACS) analysis of strains expressing GFP-tagged proteins

Proteins were quantified by fluorescence activated cell sorting (FACS) fluorometry from strains harboring a protein fused to GFP as described

previously [37]. GFP-tagged yeast cells were illuminated using a 488 nm laser (Coherent Sapphire, solid state, 20 mW, 488 nm, Becton Dickinson, San Jose, CA, USA), and fluorescence was collected through a 505 nm long-pass and HQ530/30 nm band-pass filters and 550 nm long-pass and HQ 575/26 nm band-pass filters (Becton Dickinson, San Jose, CA, USA) on a LSR-II cytometer. Thresholds of 20,000 and 15,000 for the FSC and SSC gates (FlowJo software, Tree Star; Ashland, OR), using HQ488/10 nm band-pass filters (Dickinson, San Jose, CA, USA), respectively, were used to exclude cellular debris. Cells were collected at a flow rate of 1,000-1,500 cells s^{-1} with a total accumulation of 30,000 cells run^{-1}. Rapamycin (Bioshop) was used at a final concentration of 200 ng/ml and added to cells, growing in SC medium containing all amino acids at 30°C with shaking at OD_{600} = 0.05, 6 hrs before analysis.

2.7.6.Fluorescence-Activated Cell Sorter (FACS) analysis of cell DNA content

For FACS analysis of DNA content, cells were synchronized and fixed at 20 minute intervals with ethanol for 1 h, washed with 50 mM sodium citrate in 50 mM Tris-HCl pH 7.5, and suspended in 50 mM Tris-HCl pH 7.5. Cells were treated with 1 mg/ml of RNaseA at 37°C for 1 h, then treated with 40

μg/ml of proteinase K at 55°C for 1 h, and washed and suspended in PBS buffer. The resulting cells were stained with 100 μg/ml propidium iodide (PI) at room temperature for 1 h. The DNA content of cells was measured on a BD LSR II flow cytometer. PI-stained yeast cells were illuminated using a 488 nm laser (Coherent Sapphire, solid state, 20 mW, 488 nm, Becton Dickinson, San Jose, CA, USA), and fluorescence was collected through HQ630/30 nm band-pass filters (Becton Dickinson, San Jose,CA, USA) on the LSR-II cytometer. Rapamycin (Bioshop) was used at a final concentration of 200 ng/ml.

2.7.7.Fluorescent protein localization

Diploid yeast cells endogenously expressing GFP-tagged protein and Nup49-dtTomato, as nuclear membrane marker, were grown overnight in SC-His + G418, centrifuged and diluted in SC with all amino acids supplemented with 200 nM rapamycin or ethanol (vehicle) to OD_{600} = 0.01. Cells were allowed to grow for 6 hours at 30°C with shaking and were transferred to a ConA-coated glass-bottom plate, gently centrifuged at 1500 rpm for 5 min and imaged.

2.7.8.Microscopy

All microscopy was carried out using filters from Chroma Technology for GFP- (41001), dtTomato- (49008), DAPI- (31000V2), cy3.0- (SP102V1), cy3.5- (SP103V1) and cy 5.0- (49006) specific excitation and emission spectra. Sample visualisations were made on a fully robotized Nikon TE2000E epifluorescence-inverted microscope equipped with a high-resolution coolsnap HQ2 camera (Photometrics). Sample illumination was accomplished with a X-cite XL120Q mercury lamp (X-cite) and magnified by an oil immersion violet corrected 60x objective (VC60XH, Nikon) and 1.5x internal optic magnification included on TE2000E for a 90x total magnification. Images were capture and analyzed using NIS-elements V3.1 (Nikon).

2.7.9.Hmt1 DHFR PCA large-scale screen

The DHFR screen procedures are described in previous literature [132]. To test for Hmt1 differential protein interactions (rapamycin *versus* ethanol vehicle control), each of 5,250 *MAT*α strains harbouring ORF-DHFR F[3]::HPH fusions were individually mated with the *MAT****a*** strain endogenously expressing the Hmt1-DHFR F[1,2]::NAT1 fusion, on YPD plates. Diploid strains were selected on SC-Lys-Met for 48 hours at 30°C. These diploid strains

were printed on SC+MTX with a sublethal rapamycin concentration (2 nM) or ethanol (vehicle) to select for positive DHFR PCA reconstitution in each condition. Colonies were allowed to grow for 21 days at 30°C. Individual plates were photographed with a high-resolution (4 Mega pixel, Canon) digital camera and resulting images were converted to 8-bit grayscale images and analyzed with ImageJ 1.36b software (National Institutes of Health, USA), measuring pixel intensities of the colonies over a constant area..

2.7.10.TAP-tagged protein immunoprecipitation

Asynchronous cell cultures were grown to log phase before centrifugation, zymolase treatment and cell lysis in immunoprecipitation (IP) buffer, containing protease inhibitors. Immunoprecipitation was performed with 2 mL of rabbit IgG sepharose beads (Sigma). Samples were rotated for 1 hour at 4°C followed by 3 washing steps with 2 mL IP buffer and elution after AcTEV protease (Invitrogen) cleavage at room temperature for 1 hour. Next, the samples were resolved by SDS-PAGE and detected by immunoblot analysis with rabbit-generated anti-C-terminal-DHFR antibodies (Sigma) for DHFR F[3] epitope recognition and rabbit-generated anti-calmodulin binding protein antibodies (Upstate) for TAP-tagged epitope visualisation. For alkaline phosphatase experiments, cell extracts and immunoprecipitates were prepared as described above. Before elution, immunoprecipitates were either treated with

alkaline phosphatase (New England Biolabs) or simultaneously treated with sodium vanadate (10 mM, Bioshop) and alkaline phosphatase in 2 mL IP buffer supplemented with NEBuffer 3 for 1 hour at 30°C and washed a third time with IP buffer. The elution and immunoblot was performed as described above.

2.7.11.Hmt1 methyltransferase assay

Methyltransferase activity assays were performed in a 15 µl reaction of 1x methyltransferase buffer in addition to enzyme and substrate proteins. For the methylation of Nab2-TAP and Npl3-TAP by *wildtype* Hmt1, S9E mutant Hmt1 or S9A mutant Hmt1, reactions were performed in the reaction mix described above with approximately 0.7 µM substrate and with 0.01 µM enzyme, as quantified from lysate resolved by SDS-PAGE and coomassie stained. Reactions were incubated at 30°C for 1 hour. All reactions were terminated by the addition of SDS-PAGE sample buffer and boiling. Samples were resolved by 12% SDS-PAGE followed by fluorography. Autoradiographs were exposed for 2 weeks at -80 °C.

2.7.12.Rluc PCA luminescence detection

Rluc PCA signals were directly acquired in reporter cell cultures at OD_{600} = 0.6. Cells were grown overnight in SC-Met-Lys to make a pre-culture. From the pre-culture, fresh cultures were started at an OD_{600} of 0.05 and allowed to grow up to 0.2 OD_{600} at 30°C with shaking. For each sample, rapamycin dilutions (or equivalent ethanol volumes) were prepared and added to each culture to grow 6 hrs at 30°C with shaking. For each sample, cells equivalent to 0.6 OD_{600} were centrifuged, supernatant was discarded and cells were resuspended in 180 µl of fresh SC-Met-Lys medium with rapamycin dilution (or equivalent ethanol volume) and transferred into white 96-well flat bottom plates (Greiner bio-one # 655075). The luciferase substrate benzyl-coelenterazine (Nanolight #301) was diluted from the stock (2 mM in absolute ethanol) using PBS at pH 7.2 containing 1 mM EDTA. An LMax II^{384} Luminometer (Molecular Devices, Sunnyvale, CA, USA) was used to measure the protein-protein interaction signal. Using the internal injectors of the Luminometer, 20 µl was added to each sample of substrate (to a final concentration of 10 µM) and incubated for 60 seconds. After incubations, the Rluc PCA signal was integrated for 10 seconds.

2.7.13.Purification of Dbf2-Mob1 kinase complex

FlagHis6HA3Dbf2 (Dbf2) was co-immunoprecipitated with His6Mob1TEVmyc9 (Mob1) from Hi5 insect cells as previously described [143]. To activate Dbf2-Mob1, the protein complex bound to anti-FLAG M2 beads (Sigma) was incubated with baculovirus-expressed Cdc15His6 in the presence of Cdc15 kinase buffer for 30 min at room temperature. The beads were washed three times with Cdc15 washing buffer to remove ATP and Cdc15His6. Dbf2-Mob1 was then eluted from the beads with 1 μg/ml FLAG peptide (Sigma) in Dbf2 kinase buffer (DKB) for four hours at 4°C. The same procedure was used to produce Dbf2N305A-Mob1, the kinase-inactive mutant of Dbf2. The eluted active or inactive Dbf2-Mob1 was used for subsequent assays.

2.7.14.Protein kinase assay

TAP-tagged Hmt1wt, Hmt1S9E and Hmt1S9A, bound to 20 μl of beads, were dialyzed with DKB and incubated with Dbf2-Mob1 or Dbf2N305A-Mob1 (~13 ng of Dbf2) and 2 μCi [γ-^{32}P]ATP for 30 min at room temperature. Kinase reactions were stopped by addition of SDS-PAGE sample buffer, fractionated on SDS-PAGE and detected by autoradiography.

3.Chapter 3: Discussion

In studying the protein interactome of *S. cerevisiae*, we have discovered unexplored functional territories that allowed me to systematically elucidate some of the basic underlying mechanisms that allow the cell cycle to adapt to changing nutrient conditions. The course I took provides a new avenue to identify key regulators in adaptation to stress conditions or indeed any cellular process. Further, our findings reveal how a cell has evolved to delay and arrest cell cycle progression by supervising propagation of complex mRNA post-transcriptional regulation.

3.1.Implications and future directions

Most protein interaction screens performed, so far, have provided only static views of the PPI network. Some attempts to describe dynamic behavior of such networks have been described but allow one only to address a restricted number of genes at a time and only those already known to be involved in a specific cellular adaptation [162]. In our case study, we performed an unbiased PPI screen of Hmt1 for all reporter strains in our collection in cells treated with rapamycin. In absence of other information about potential kinases, phosphatases or regulators of these events, a biased set of reporters I would hardly have had any chance of success in identifying the rapamycin-induced regulators of Hmt1 Ser9 phosphorylation. In one fell swoop, my strategy

identified a potential set of regulators, potential kinases and phosphatases, along with proteins responsible for RNA export, maturation and turnover.

Interestingly, following our discovery of the interaction of Hmt1 with Dbf2, this same kinase was discovered to play a role in stabilizing *CLB2* transcripts during entry into M phase. Further, we knew that Dbf2, Hmt1 and hnRNP components associate with specific gene loci, including that of CLB2 and the functions of hnRNP proteins in nuclear export and stabilization of mRNA in the cytoplasm [93, 96]. We know that the activity of Dbf2 in entry into M phase is minimal, not reaching a peak until later in M phase [139, 140]. We were thus able to argue that Dbf2 activity is not dispensable to regulation of mRNA stability by hnRNPs. It is known, furthermore, that PP2A is activated following release of its catalytic subunit Pph22 from the TAP42 complex when rapamycin-FKBP binds to TOR kinase and upon starvation [84]. In our model, activated PP2A dephosphorylates and shifts Hmt1 equilibrium from its active hexameric form to its inactive monomeric state, resulting ultimately in destabilization of *CLB2* mRNA and therefore attenuation of entry into M phase.

Rapamycin-induced inhibition of TOR kinase and afferent signalling pathways has been subjects of great interest because of its potent immunosuppresive and antineoplasic potentials. Until recently, all therapeutic activities of rapamycin were attributed to translational regulation, through S6K and S6 proteins, necessary for efficient cap-dependant translation, and other

afferent signalling pathways [81, 163-166]. Our results suggest a novel mechanism of rapamycin action, the Dbf2/Pph22 competition for Hmt1 phosphorylation state, implicated in regulating transcript stabilities upon TOR kinase inhibitions. The Dbf2 kinase is cell cycle regulated, with basal activity throughout the cell cycle and was shown to bind mRNA co-transcriptionally, such as *CLB2* transcripts [93]. Likewise, we demonstrated that a Pph22 phosphatase dead mutant strain treated with rapamycin contains phosphorylated and activated Hmt1 hexamers. Pph22 could be one of the several phosphatases that act on Hmt1 [84, 167]. The basic circuit of Hmt1 control could potentially be the target of drug design, since both the Dbf2 and Hmt1 have human homologs as does Pph22 and other phosphatases that may act on Hmt1 homologs. It is conceivable that compounds targeting these enzymes could illicit similar consequences to cell cycle progression upon inhibition as rapamycin, but with less pleiotropic effects.

We know little about both the type and extent of the re-organization the protein interactome undergoes in response to different environmental changes. This lack of knowledge is mainly due to the colossal size of the interaction space and the proportional efforts needed to map it. For example, in order to discover interactions among the ~6,000 *S. cerevisiae* proteins, one needs to screen a space of 36,000,000 possible interactions. Thus, in order to tackle this problem we must reduce the size of the search space. On one hand, even measuring the dynamics of one interaction, Hmt1 homo-oligomerization,

provided information about the spectrum of protein network adaptations to rapamycin. On the other hand, a proteome-wide analysis of self-interactions under different conditions could reveal whole spaces of cellular processes never imagined. This also provides a solution to the problem of the size of a dynamic interactome screen of hetero-interactions. If we measure only self-interactions, the size of the space to probe scales linearly - not exponentially - with the number of proteins involved. In other words, we could probe the inner variability of the nodes in the network to identify the regions that change most under a specific set of conditions. We expect that self-interactions can be used to probes such changes, because the quantity of a self-interaction integrates information about both: (i) protein effective concentration (the higher the concentration of a protein the more it prone to self-interacts), and (ii) protein structural state (a structural change or a post-translational modification inducing a change in oligomeric state would result in a change in signal). Importantly also, we would expect this strategy to cover a significant part of the proteome. We indeed know that up to at least 50% of proteins form homo-oligomers under one condition [168, 169] and it has been known for several decades that allosteric as well as homo-oligomeric changes are frequently used to control protein function [170]. Therefore, about 50% or more of proteins probed could theoretically yield a self-interaction signal, and most importantly a change in such a signal. These interaction dynamics data could be combined with transcriptomic, protein turnover and protein localization information, potentially revealing new insights into protein regulatory networks. This approach could be

generalized to identify and mechanistically link genes, including those of unknown function, to any cellular process.

Post-transcriptional regulation of mRNA stability is a vast field of unexplored potential mechanisms for regulating specific cellular response. Recent results of the Encyclopedia of DNA Elements (ENCODE) project have shown that human genome produces a considerable amount of non-coding RNA. It is conceivable that RNA molecules have the capacity to regulate gene expression, transcript availability for translation and protein functions [171-173]. To explore this new continent, it will be necessary to design new approaches to systematically detect protein-RNA and RNA-RNA interactions in intact cells to better understand the role of transcription regulation and non-coding RNA in adaptation to intrinsic and extrinsic stresses and indeed all cellular processes. I hope that the strategy I have laid out in this thesis proves useful to all those prepared to explore this new world.

Chapter 4: Bibliography

1. Imoto, Y., et al., *The cell cycle, including the mitotic cycle and organelle division cycles, as revealed by cytological observations.* J Electron Microsc (Tokyo), 2011. **60 Suppl 1**: p. S117-36.
2. Ball, D.A., et al., *Oscillatory dynamics of cell cycle proteins in single yeast cells analyzed by imaging cytometry.* PLoS One, 2011. **6**(10): p. e26272.
3. Orlova, D.Y., et al., *Arrangement of nuclear structures is not transmitted through mitosis but is identical in sister cells.* J Cell Biochem, 2012.
4. Kiyomitsu, T. and I.M. Cheeseman, *Chromosome- and spindle-pole-derived signals generate an intrinsic code for spindle position and orientation.* Nat Cell Biol, 2012. **14**(3): p. 311-7.
5. Reed, S.I., *Ratchets and clocks: the cell cycle, ubiquitylation and protein turnover.* Nat Rev Mol Cell Biol, 2003. **4**(11): p. 855-64.
6. Morgan, D.O., *Principles of CDK regulation.* Nature, 1995. **374**(6518): p. 131-4.
7. Orlando, D.A., et al., *Global control of cell-cycle transcription by coupled CDK and network oscillators.* Nature, 2008. **453**(7197): p. 944-7.
8. Stine, W.B., Jr., et al., *In vitro characterization of conditions for amyloid-beta peptide oligomerization and fibrillogenesis.* J Biol Chem, 2003. **278**(13): p. 11612-22.
9. Zenklusen, D., D.R. Larson, and R.H. Singer, *Single-RNA counting reveals alternative modes of gene expression in yeast.* Nat Struct Mol Biol, 2008. **15**(12): p. 1263-71.
10. Ito, T., et al., *Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins.* Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1143-7.
11. Ito, T., et al., *A comprehensive two-hybrid analysis to explore the yeast protein interactome.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4569-74.
12. Ho, Y., et al., *Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry.* Nature, 2002. **415**(6868): p. 180-3.

13. Gavin, A.C., et al., *Proteome survey reveals modularity of the yeast cell machinery.* Nature, 2006. **440**(7084): p. 631-6.
14. Krogan, N.J., et al., *Global landscape of protein complexes in the yeast Saccharomyces cerevisiae.* Nature, 2006. **440**(7084): p. 637-43.
15. Tu, B.P., et al., *Cyclic changes in metabolic state during the life of a yeast cell.* Proc Natl Acad Sci U S A, 2007. **104**(43): p. 16886-91.
16. Tu, B.P., et al., *Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes.* Science, 2005. **310**(5751): p. 1152-8.
17. Bolon, D.N., et al., *Specificity versus stability in computational protein design.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12724-9.
18. Levchenko, I., et al., *Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag.* Mol Cell, 2003. **12**(2): p. 365-72.
19. Flynn, J.M., et al., *Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10584-9.
20. Pelletier, J.N., F.X. Campbell-Valois, and S.W. Michnick, *Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12141-6.
21. Stefan, E., et al., *Quantification of dynamic protein complexes using Renilla luciferase fragment complementation applied to protein kinase A activities in vivo.* Proc Natl Acad Sci U S A, 2007. **104**(43): p. 16916-21.
22. Remy, I. and S.W. Michnick, *A highly sensitive protein-protein interaction assay based on Gaussia luciferase.* Nat Methods, 2006. **3**(12): p. 977-9.
23. Cramer, P., D.A. Bushnell, and R.D. Kornberg, *Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution.* Science, 2001. **292**(5523): p. 1863-76.
24. Remy, I. and S.W. Michnick, *Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5394-9.
25. Diehn, M., et al., *Large-scale identification of secreted and membrane-associated gene products using DNA microarrays.* Nat Genet, 2000. **25**(1): p. 58-62.
26. Miller, J.P., et al., *Large-scale identification of yeast integral membrane protein interactions.* Proc Natl Acad Sci U S A, 2005. **102**(34): p. 12123-8.
27. Zamyatnin, A.A., Jr., et al., *Assessment of the integral membrane protein topology in living cells.* Plant J, 2006. **46**(1): p. 145-54.
28. Bandyopadhyay, S., et al., *Rewiring of genetic networks in response to DNA damage.* Science, 2010. **330**(6009): p. 1385-9.
29. Pena-Castillo, L. and T.R. Hughes, *Why are there still over 1000 uncharacterized yeast genes?* Genetics, 2007. **176**(1): p. 7-14.

30. Enquist-Newman, M., et al., *Dad1p, third component of the Duo1p/Dam1p complex involved in kinetochore function and mitotic spindle integrity.* Mol Biol Cell, 2001. **12**(9): p. 2601-13.
31. Pu, S., et al., *Identifying functional modules in the physical interactome of Saccharomyces cerevisiae.* Proteomics, 2007. **7**(6): p. 944-60.
32. Monod, J., *paper presented at the Proceedings of the Eleventh Nobel Symposium.* Sodergarn, Lidingo, County of Stockholm, 1968.
33. Michnick, S.W., et al., *Universal strategies in research and drug discovery based on protein-fragment complementation assays.* Nat Rev Drug Discov, 2007. **6**(7): p. 569-82.
34. Pelletier, J.N., et al., *An in vivo library-versus-library selection of optimized protein-protein interactions.* Nat Biotechnol, 1999. **17**(7): p. 683-90.
35. Uetz, P., et al., *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. **403**(6770): p. 623-7.
36. Collins, S.R., et al., *Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae.* Mol Cell Proteomics, 2007. **6**(3): p. 439-50.
37. Newman, J.R., et al., *Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise.* Nature, 2006. **441**(7095): p. 840-6.
38. Hughes, E.G., *Using evidence from randomized trials in fertility practice.* Hum Fertil (Camb), 2007. **10**(1): p. 7-12.
39. Kim, H., et al., *A global topology map of the Saccharomyces cerevisiae membrane proteome.* Proc Natl Acad Sci U S A, 2006. **103**(30): p. 11142-7.
40. Damelin, M. and P.A. Silver, *In situ analysis of spatial relationships between proteins of the nuclear pore complex.* Biophys J, 2002. **83**(6): p. 3626-36.
41. Alber, F., et al., *Determining the architectures of macromolecular assemblies.* Nature, 2007. **450**(7170): p. 683-94.
42. Medalia, O., et al., *Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography.* Science, 2002. **298**(5596): p. 1209-13.
43. Nickell, S., et al., *A visual approach to proteomics.* Nat Rev Mol Cell Biol, 2006. **7**(3): p. 225-30.
44. Cid, V.J., et al., *Orchestrating the cell cycle in yeast: sequential localization of key mitotic regulators at the spindle pole and the bud neck.* Microbiology, 2002. **148**(Pt 9): p. 2647-59.
45. McCusker, D., et al., *Cdk1 coordinates cell-surface growth with the cell cycle.* Nat Cell Biol, 2007. **9**(5): p. 506-15.
46. Park, H.O. and E. Bi, *Central roles of small GTPases in the development of cell polarity in yeast and beyond.* Microbiol Mol Biol Rev, 2007. **71**(1): p. 48-96.
47. Pruyne, D., et al., *Mechanisms of polarized growth and organelle segregation in yeast.* Annu Rev Cell Dev Biol, 2004. **20**: p. 559-91.

48. Shintani, T. and D.J. Klionsky, *Autophagy in health and disease: a double-edged sword.* Science, 2004. **306**(5698): p. 990-5.
49. Lee, J.A., et al., *ESCRT-III dysfunction causes autophagosome accumulation and neurodegeneration.* Curr Biol, 2007. **17**(18): p. 1561-7.
50. Coudreuse, D.Y., et al., *Wnt gradient formation requires retromer function in Wnt-producing cells.* Science, 2006. **312**(5775): p. 921-4.
51. Nothwehr, S.F., S.A. Ha, and P. Bruinsma, *Sorting of yeast membrane proteins into an endosome-to-Golgi pathway involves direct interaction of their cytosolic domains with Vps35p.* J Cell Biol, 2000. **151**(2): p. 297-310.
52. Reddy, J.V. and M.N. Seaman, *Vps26p, a component of retromer, directs the interactions of Vps35p in endosome-to-Golgi retrieval.* Mol Biol Cell, 2001. **12**(10): p. 3242-56.
53. Hierro, A., et al., *Functional architecture of the retromer cargo-recognition complex.* Nature, 2007. **449**(7165): p. 1063-7.
54. Deloche, O. and R.W. Schekman, *Vps10p cycles between the TGN and the late endosome via the plasma membrane in clathrin mutants.* Mol Biol Cell, 2002. **13**(12): p. 4296-307.
55. Yen, W.L., et al., *Atg27 is required for autophagy-dependent cycling of Atg9.* Mol Biol Cell, 2007. **18**(2): p. 581-93.
56. Subramaniam, R., et al., *Direct visualization of protein interactions in plant cells.* Nat Biotechnol, 2001. **19**(8): p. 769-72.
57. Ramseyer, J., et al., *The use of affinity chromatography in purification of cyclic nucleotide receptor proteins.* Biochim Biophys Acta, 1976. **446**(2): p. 358-70.
58. Goldstein, A.L. and J.H. McCusker, *Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae.* Yeast, 1999. **15**(14): p. 1541-53.
59. Gietz, R.D. and R.A. Woods, *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method.* Methods Enzymol, 2002. **350**: p. 87-96.
60. Gavin, A.C., et al., *Functional organization of the yeast proteome by systematic analysis of protein complexes.* Nature, 2002. **415**(6868): p. 141-7.
61. Xia, Y., L.J. Lu, and M. Gerstein, *Integrated prediction of the helical membrane protein interactome in yeast.* J Mol Biol, 2006. **357**(1): p. 339-49.
62. Brideau, C., et al., *Improved statistical methods for hit selection in high-throughput screening.* J Biomol Screen, 2003. **8**(6): p. 634-47.
63. Remy, I. and S.W. Michnick, *Visualization of biochemical networks in living cells.* Proc Natl Acad Sci U S A, 2001. **98**(14): p. 7678-83.
64. Remy, I., I.A. Wilson, and S.W. Michnick, *Erythropoietin receptor activation by a ligand-induced conformation change.* Science, 1999. **283**(5404): p. 990-3.

65. Benton, R., et al., *Atypical membrane topology and heteromeric function of Drosophila odorant receptors in vivo.* PLoS Biol, 2006. **4**(2): p. e20.
66. Falcon, S. and R. Gentleman, *Using GOstats to test gene lists for GO term association.* Bioinformatics, 2007. **23**(2): p. 257-8.
67. R. Ihaka, R.G., *R : A Language for Data Analysis and Graphics.* Journal of Computational and Graphical Statistics, 1996. **5**.
68. Ihmels, J., S. Bergmann, and N. Barkai, *Defining transcription modules using large-scale gene expression data.* Bioinformatics, 2004. **20**(13): p. 1993-2003.
69. Itzhaki, Z., et al., *Evolutionary conservation of domain-domain interactions.* Genome Biol, 2006. **7**(12): p. R125.
70. Stein, A., R.B. Russell, and P. Aloy, *3did: interacting protein domains of known three-dimensional structure.* Nucleic Acids Res, 2005. **33**(Database issue): p. D413-7.
71. Finn, R.D., M. Marshall, and A. Bateman, *iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions.* Bioinformatics, 2005. **21**(3): p. 410-2.
72. Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.
73. Rives, A.W. and T. Galitski, *Modular organization of cellular networks.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1128-33.
74. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.
75. Tarassov, K. and S.W. Michnick, *iVici: Interrelational Visualization and Correlation Interface.* Genome Biol, 2005. **6**(13): p. R115.
76. Lord, P.W., et al., *Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation.* Bioinformatics, 2003. **19**(10): p. 1275-83.
77. Luscombe, N.M., et al., *Genomic analysis of regulatory network dynamics reveals large topological changes.* Nature, 2004. **431**(7006): p. 308-12.
78. Mani, K.M., et al., *A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas.* Mol Syst Biol, 2008. **4**: p. 169.
79. Wang, K., et al., *Genome-wide identification of post-translational modulators of transcription factor activity in human B cells.* Nat Biotechnol, 2009. **27**(9): p. 829-39.
80. Jothi, R., et al., *Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture.* Mol Syst Biol, 2009. **5**: p. 294.
81. Heitman, J., N.R. Movva, and M.N. Hall, *Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast.* Science, 1991. **253**(5022): p. 905-9.

82. Brunn, G.J., et al., *Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin.* Science, 1997. **277**(5322): p. 99-101.
83. Chung, J., et al., *Rapamycin-FKBP specifically blocks growth-dependent activation of and signaling by the 70 kd S6 protein kinases.* Cell, 1992. **69**(7): p. 1227-36.
84. Di Como, C.J. and K.T. Arndt, *Nutrients, via the Tor proteins, stimulate the association of Tap42 with type 2A phosphatases.* Genes Dev, 1996. **10**(15): p. 1904-16.
85. Boisnard, S., et al., *H2O2 activates the nuclear localization of Msn2 and Maf1 through thioredoxins in Saccharomyces cerevisiae.* Eukaryot Cell, 2009. **8**(9): p. 1429-38.
86. Santhanam, A., et al., *PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p.* Eukaryot Cell, 2004. **3**(5): p. 1261-71.
87. Travesa, A., A. Duch, and D.G. Quintana, *Distinct phosphatases mediate the deactivation of the DNA damage checkpoint kinase Rad53.* J Biol Chem, 2008. **283**(25): p. 17123-30.
88. Duvel, K., et al., *Multiple roles of Tap42 in mediating rapamycin-induced transcriptional changes in yeast.* Mol Cell, 2003. **11**(6): p. 1467-78.
89. Tiainen, M., et al., *Terminally differentiated skeletal myotubes are not confined to G0 but can enter G1 upon growth factor stimulation.* Cell Growth Differ, 1996. **7**(8): p. 1039-50.
90. Mailand, N., et al., *Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability.* EMBO J, 2002. **21**(21): p. 5911-20.
91. Hartley, R.S., R.E. Rempel, and J.L. Maller, *In vivo regulation of the early embryonic cell cycle in Xenopus.* Dev Biol, 1996. **173**(2): p. 408-19.
92. Komili, S. and P.A. Silver, *Coupling and coordination in gene expression processes: a systems biology view.* Nat Rev Genet, 2008. **9**(1): p. 38-48.
93. Trcek, T., et al., *Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast.* Cell, 2011. **147**(7): p. 1484-97.
94. Green, D.M., et al., *Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p.* J Biol Chem, 2002. **277**(10): p. 7752-60.
95. Wong, C.M., et al., *Yeast arginine methyltransferase Hmt1p regulates transcription elongation and termination by methylating Npl3p.* Nucleic Acids Res, 2010. **38**(7): p. 2217-28.
96. Yu, M.C., et al., *Arginine methyltransferase affects interactions and recruitment of mRNA processing and export factors.* Genes Dev, 2004. **18**(16): p. 2024-35.

97. Najbauer, J., et al., *Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins.* J Biol Chem, 1993. **268**(14): p. 10501-9.
98. Giorgi, C. and M.J. Moore, *The nuclear nurture and cytoplasmic nature of localized mRNPs.* Semin Cell Dev Biol, 2007. **18**(2): p. 186-93.
99. Hasegawa, Y., K. Irie, and A.P. Gerber, *Distinct roles for Khd1p in the localization and expression of bud-localized mRNAs in yeast.* RNA, 2008. **14**(11): p. 2333-47.
100. Gari, E., et al., *Whi3 binds the mRNA of the G1 cyclin CLN3 to modulate cell fate in budding yeast.* Genes Dev, 2001. **15**(21): p. 2803-8.
101. Palou, G., et al., *Cyclin regulation by the s phase checkpoint.* J Biol Chem, 2010. **285**(34): p. 26431-40.
102. McBride, A.E., et al., *Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions.* J Biol Chem, 2005. **280**(35): p. 30888-98.
103. Weiss, V.H., et al., *The structure and oligomerization of the yeast arginine methyltransferase, Hmt1.* Nat Struct Biol, 2000. **7**(12): p. 1165-71.
104. Bradley, P.H., et al., *Coordinated concentration changes of transcripts and metabolites in Saccharomyces cerevisiae.* PLoS Comput Biol, 2009. **5**(1): p. e1000270.
105. Oikonomou, C. and F.R. Cross, *Rising cyclin-CDK levels order cell cycle events.* PLoS One, 2011. **6**(6): p. e20788.
106. Oikonomou, C. and F.R. Cross, *Frequency control of cell cycle oscillators.* Curr Opin Genet Dev, 2010. **20**(6): p. 605-12.
107. Nakashima, A., et al., *The yeast Tor signaling pathway is involved in G2/M transition via polo-kinase.* PLoS One, 2008. **3**(5): p. e2223.
108. Hollien, J. and J.S. Weissman, *Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response.* Science, 2006. **313**(5783): p. 104-7.
109. Romero-Santacreu, L., et al., *Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae.* RNA, 2009. **15**(6): p. 1110-20.
110. Bertrand, E., et al., *Localization of ASH1 mRNA particles in living yeast.* Mol Cell, 1998. **2**(4): p. 437-45.
111. Munchel, S.E., et al., *Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay.* Mol Biol Cell, 2011. **22**(15): p. 2787-95.
112. Guhaniyogi, J. and G. Brewer, *Regulation of mRNA stability in mammalian cells.* Gene, 2001. **265**(1-2): p. 11-23.
113. Schoenberg, D.R. and L.E. Maquat, *Regulation of cytoplasmic mRNA decay.* Nat Rev Genet, 2012. **13**(4): p. 246-59.

114. Houseley, J. and D. Tollervey, *The many pathways of RNA degradation.* Cell, 2009. **136**(4): p. 763-76.
115. Larson, D.R., et al., *Real-time observation of transcription initiation and elongation on an endogenous yeast gene.* Science, 2011. **332**(6028): p. 475-8.
116. Lionnet, T., et al., *Nuclear physics: quantitative single-cell approaches to nuclear organization and gene expression.* Cold Spring Harb Symp Quant Biol, 2010. **75**: p. 113-26.
117. Spellman, P.T., et al., *Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization.* Mol Biol Cell, 1998. **9**(12): p. 3273-97.
118. Michea, L., et al., *Cell cycle delay and apoptosis are induced by high salt and urea in renal medullary cells.* Am J Physiol Renal Physiol, 2000. **278**(2): p. F209-18.
119. Bulavin, D.V., et al., *Initiation of a G2/M checkpoint after ultraviolet radiation requires p38 kinase.* Nature, 2001. **411**(6833): p. 102-7.
120. Ahn, S.H., A. Acurio, and S.J. Kron, *Regulation of G2/M progression by the STE mitogen-activated protein kinase pathway in budding yeast filamentous growth.* Mol Biol Cell, 1999. **10**(10): p. 3301-16.
121. Hardwick, J.S., et al., *Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 14866-70.
122. Staschke, K.A., et al., *Integration of general amino acid control and target of rapamycin (TOR) regulatory pathways in nitrogen assimilation in yeast.* J Biol Chem, 2010. **285**(22): p. 16893-911.
123. Martin, Y., et al., *Npr1 Ser/Thr protein kinase links nitrogen source quality and carbon availability with the yeast nitrate transporter (Ynt1) levels.* J Biol Chem, 2011. **286**(31): p. 27225-35.
124. Barbet, N.C., et al., *TOR controls translation initiation and early G1 progression in yeast.* Mol Biol Cell, 1996. **7**(1): p. 25-42.
125. Gallego, C., et al., *The Cln3 cyclin is down-regulated by translational repression and degradation during the G1 arrest caused by nitrogen deprivation in budding yeast.* EMBO J, 1997. **16**(23): p. 7196-206.
126. Zenklusen, D. and R.H. Singer, *Analyzing mRNA expression using single mRNA resolution fluorescent in situ hybridization.* Methods Enzymol, 2010. **470**: p. 641-59.
127. Hector, R.E., et al., *Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export.* EMBO J, 2002. **21**(7): p. 1800-10.
128. Kerr, S.C., et al., *The Ccr4-Not complex interacts with the mRNA export machinery.* PLoS One, 2011. **6**(3): p. e18302.
129. Henry, M.F. and P.A. Silver, *A novel methyltransferase (Hmt1p) modifies poly(A)+-RNA-binding proteins.* Mol Cell Biol, 1996. **16**(7): p. 3668-78.
130. Shen, E.C., et al., *Arginine methylation facilitates the nuclear export of hnRNP proteins.* Genes Dev, 1998. **12**(5): p. 679-91.

131. Blom, N., S. Gammeltoft, and S. Brunak, *Sequence and structure-based prediction of eukaryotic protein phosphorylation sites.* J Mol Biol, 1999. **294**(5): p. 1351-62.
132. Tarassov, K., et al., *An in vivo map of the yeast protein interactome.* Science, 2008. **320**(5882): p. 1465-70.
133. Mah, A.S., et al., *Substrate specificity analysis of protein kinase complex Dbf2-Mob1 by peptide library and proteome array screening.* BMC Biochem, 2005. **6**: p. 22.
134. Malleshaiah, M.K., et al., *The scaffold protein Ste5 directly controls a switch-like mating decision in yeast.* Nature, 2010. **465**(7294): p. 101-5.
135. St-Pierre, J., et al., *Polo kinase regulates mitotic chromosome condensation by hyperactivation of condensin DNA supercoiling activity.* Mol Cell, 2009. **34**(4): p. 416-26.
136. Skotheim, J.M., et al., *Positive feedback of G1 cyclins ensures coherent cell cycle entry.* Nature, 2008. **454**(7202): p. 291-6.
137. Simpson-Lavy, K.J. and M. Brandeis, *Phosphorylation of Cdc5 regulates its accumulation.* Cell Div, 2011. **6**: p. 23.
138. Mortensen, E.M., et al., *Cdc28-dependent regulation of the Cdc5/Polo kinase.* Curr Biol, 2005. **15**(22): p. 2033-7.
139. Johnston, L.H., et al., *The product of the Saccharomyces cerevisiae cell cycle gene DBF2 has homology with protein kinases and is periodically expressed in the cell cycle.* Mol Cell Biol, 1990. **10**(4): p. 1358-66.
140. Toyn, J.H. and L.H. Johnston, *The Dbf2 and Dbf20 protein kinases of budding yeast are activated after the metaphase to anaphase cell cycle transition.* EMBO J, 1994. **13**(5): p. 1103-13.
141. Hotz, M., et al., *Spindle pole bodies exploit the mitotic exit network in metaphase to drive their age-dependent segregation.* Cell, 2012. **148**(5): p. 958-72.
142. Xu, S., et al., *Phosphorylation and spindle pole body localization of the Cdc15p mitotic regulatory protein kinase in budding yeast.* Curr Biol, 2000. **10**(6): p. 329-32.
143. Mah, A.S., J. Jang, and R.J. Deshaies, *Protein kinase Cdc15 activates the Dbf2-Mob1 kinase complex.* Proc Natl Acad Sci U S A, 2001. **98**(13): p. 7325-30.
144. Lee, S.E., et al., *Order of function of the budding-yeast mitotic exit-network proteins Tem1, Cdc15, Mob1, Dbf2, and Cdc5.* Curr Biol, 2001. **11**(10): p. 784-8.
145. Wang, Y. and T.Y. Ng, *Phosphatase 2A negatively regulates mitotic exit in Saccharomyces cerevisiae.* Mol Biol Cell, 2006. **17**(1): p. 80-9.
146. Roelants, F.M., et al., *Pkh1 and Pkh2 differentially phosphorylate and activate Ypk1 and Ykr2 and define protein kinase modules required for maintenance of cell wall integrity.* Mol Biol Cell, 2002. **13**(9): p. 3005-28.
147. Luo, G., et al., *Nutrients and the Pkh1/2 and Pkc1 protein kinases control mRNA decay and P-body assembly in yeast.* J Biol Chem, 2011. **286**(11): p. 8759-70.

148. Kim, Y., et al., *A conserved phosphatase cascade that regulates nuclear membrane biogenesis.* Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6596-601.
149. Enyenihi, A.H. and W.S. Saunders, *Large-scale functional genomic analysis of sporulation and meiosis in Saccharomyces cerevisiae.* Genetics, 2003. **163**(1): p. 47-54.
150. Visintin, R. and A. Amon, *Regulation of the mitotic exit protein kinases Cdc15 and Dbf2.* Mol Biol Cell, 2001. **12**(10): p. 2961-74.
151. Chen, Y.C., et al., *Protein arginine methylation facilitates cotranscriptional recruitment of pre-mRNA splicing factors.* Mol Cell Biol, 2010. **30**(21): p. 5245-56.
152. Huttelmaier, S., et al., *Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1.* Nature, 2005. **438**(7067): p. 512-5.
153. Shen, Z., A. St-Denis, and P. Chartrand, *Cotranscriptional recruitment of She2p by RNA pol II elongation factor Spt4-Spt5/DSIF promotes mRNA localization to the yeast bud.* Genes Dev, 2010. **24**(17): p. 1914-26.
154. Hao, Y., et al., *Tumor suppressor LATS1 is a negative regulator of oncogene YAP.* J Biol Chem, 2008. **283**(9): p. 5496-509.
155. Buchler, N.E. and F.R. Cross, *Protein sequestration generates a flexible ultrasensitive response in a genetic network.* Mol Syst Biol, 2009. **5**: p. 272.
156. Tsai, T.Y., et al., *Robust, tunable biological oscillations from interlinked positive and negative feedback loops.* Science, 2008. **321**(5885): p. 126-9.
157. Santos, S.D. and J.E. Ferrell, *Systems biology: On the cell cycle and its switches.* Nature, 2008. **454**(7202): p. 288-9.
158. Holt, L.J., A.N. Krutchinsky, and D.O. Morgan, *Positive feedback sharpens the anaphase switch.* Nature, 2008. **454**(7202): p. 353-7.
159. Pestov, D.G. and N. Shcherbik, *Rapid cytoplasmic turnover of yeast ribosomes in response to rapamycin inhibition of TOR.* Mol Cell Biol, 2012. **32**(11): p. 2135-44.
160. D'Amours, D., F. Stegmeier, and A. Amon, *Cdc14 and condensin control the dissolution of cohesin-independent chromosome linkages at repeated DNA.* Cell, 2004. **117**(4): p. 455-69.
161. Alberti, S., A.D. Gitler, and S. Lindquist, *A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae.* Yeast, 2007. **24**(10): p. 913-9.
162. Schlecht, U., et al., *Multiplex assay for condition-dependent changes in protein-protein interactions.* Proc Natl Acad Sci U S A, 2012. **109**(23): p. 9213-8.
163. Li, H., et al., *Nutrient regulates Tor1 nuclear localization and association with rDNA promoter.* Nature, 2006. **442**(7106): p. 1058-61.

164. Martin, D.E., A. Soulard, and M.N. Hall, *TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1.* Cell, 2004. **119**(7): p. 969-79.
165. Noda, T. and Y. Ohsumi, *Tor, a phosphatidylinositol kinase homologue, controls autophagy in yeast.* J Biol Chem, 1998. **273**(7): p. 3963-6.
166. Guertin, D.A., et al., *Functional genomics identifies TOR-regulated genes that control growth and division.* Curr Biol, 2006. **16**(10): p. 958-70.
167. Huber, A., et al., *Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis.* Genes Dev, 2009. **23**(16): p. 1929-43.
168. Levy, E.D., et al., *3D complex: a structural classification of protein complexes.* PLoS Comput Biol, 2006. **2**(11): p. e155.
169. Kuhner, S., et al., *Proteome organization in a genome-reduced bacterium.* Science, 2009. **326**(5957): p. 1235-40.
170. Monod, J., J.P. Changeux, and F. Jacob, *Allosteric proteins and cellular control systems.* J Mol Biol, 1963. **6**: p. 306-29.
171. Banfai, B., et al., *Long noncoding RNAs are rarely translated in two human cell lines.* Genome Res, 2012. **22**(9): p. 1646-57.
172. Natarajan, A., et al., *Predicting cell-type-specific gene expression from regions of open chromatin.* Genome Res, 2012. **22**(9): p. 1711-22.
173. Dong, X., et al., *Modeling gene expression using chromatin features in various cellular contexts.* Genome Biol, 2012. **13**(9): p. R53.