uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Jean-Philippe Lambert
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biochemistry)
GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Defining the Budding Yeast Chromatin Associated Interactome

TITRE DE LA THÈSE / TITLE OF THESIS

Daniel Figeys
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Marjorie Brand

Benoît Coulombe (I.R.C.M.)

Lynn Megeney

Laura Trinkle-Mulcahy

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Defining the Budding Yeast

Chromatin Associated Interactome

Jean-Philippe Lambert

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the PhD degree in Biochemistry

Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-79728-0
Our file Notre référence
ISBN: 978-0-494-79728-0

Abstract

The identification of protein-protein as emerged as a powerful tool to characterize biological systems. In budding yeast in particular, systematic mapping of protein-protein interactions for most open reading frames drastically improve our understanding of this model organisms. Still, some classes of proteins, such as DNA binding proteins, remain poorly studied due to a lack of proper tools for their study. In this thesis, I describe the development of a novel affinity purification approach for the characterization of DNA binding proteins and of their associated proteins. The modified chromatin immunopurification (mChIP) approach consist of a single affinity purification step whereby chromatin-bound protein networks are isolated from mildly sonicated and gently clarified cellular extracts using magnetic beads coated with antibodies. The mChIP method was characterized in details demonstrating significant gain in sensitivity toward the determination of DNA binding protein interactome. Moreover, I showed that mChIP purification were successful for multiple class of proteins including numerous non-histone DNA binding proteins. Following the successful development of the mChIP method, I embarked on a large-scale characterization of chromatin associated proteins in budding yeast. In this way, 102 different DNA binding proteins were characterized by mChIP to determine their chromatin associated protein networks. This effort resulted in the detection of 3576 high confidence protein associations with 724 distinct preys. Approximately 75% of the baits had significantly improved interaction coverage using mChIP compared to the classical affinity purification methodologies. I also utilized the

mChIP approach to perform targeted study of histone chaperones and unravel their interconnection. In summary, the mChIP method that I developed now makes it possible to systematically study chromatin associated proteins in budding yeast, breaking down a technical barrier that existed in the field of chromatin research for too long.

Acknowledgements

I would like to take this opportunity to thank my supervisor, Dr. Daniel Figeys, for his help and support throughout my graduate work. It has been a pleasure to interact with him and to learn in his company. Daniel was always supportive of my work, provided me with freedom in my research and as a result enabled me to grow considerably as a scientist. I would also like to thank members of the Figeys lab for their help during my graduate work. In particular, Fred Elisma who has been a great help with everything computer related and Ruijin Tian for numerous helpful discussion regarding mass spectrometry.

I would also like to acknowledge the support of Dr. Kristin Baetz and of her laboratory members throughout my years at the Ottawa Institute of Systems Biology. It has been both a pleasure and a fruitful venture to collaborate with Kristin and her lab members over the years. Another collaborator of many years that I want to acknowledge is Dr. Jeffrey Fillingham for his help, support and guidance. It has been a pleasure of working with Jeff during my graduate work. Finally, I want to thank my fiancée Isabelle, my family and my friends for their support throughout my graduate work.

Table of contents

List of Tables

List of Figures

List of Appendix

List of Abbreviations

α ..... alpha

Δ ..... gene deletion

Mg ..... microgram

µL ..... microliter

µm ..... micrometer

nL ..... nanoliter

~ ..... approximately

Ac ..... acetylation

Ade ..... adenine

AP ..... affinity purification

AP-MS ..... affinity purification coupled to mass spectrometry

cDNA ..... complementary DNA

ChIP ..... chromatin immunoprecipitation

ChIP-chip ..... chromatin immunoprecipitation followed by microarray hybridization

ChIP-seq ..... chromatin immunoprecipitation followed by sequencing

COMPASS ..... complex proteins associated with Set1

DNA ..... deoxyribonucleic acid

DNase ..... deoxyribonuclease I

E-MAP ..... epistatic miniarray profile

ERα ..... estrogen receptor alpha

ESI ..... electrospray ionization

EtBr ..... ethidium bromide

FACT ..... facilitates chromatin transcription

GFP ..... green fluorescent protein

GST ..... glutathione S-transferase

H2B K123 ..... histone H2B lysine 123

H3K4 ..... histone H3 lysine 4

H3K56 ..... histone H3 lysine 56

HPLC ..... high performance liquid chromatograph

HU ..... hydroxyurea

HIS ..... histidine

IgG ..... Immunoglobulin G

KanMX ..... kanamycin resistance cassette

KAT ..... lysine acetyltransferase

KDAC ..... lysine deacetylase

KMT ..... lysine methylase

LacO ..... lactose operator

LC ..... liquid chromatography

LC-MS/MS ..... liquid chromatography coupled to mass spectrometry

Leu ..... leucine

Lys ..... lysine

MALDI ..... Matrix-assisted laser desorption/ionization

MBF ..... Mlu1 cell cycle box binding factor

mChIP ..... modified chromatin immunopurification

MCM ..... minichromosome maintenance complex

Me ..... methylation

MNase ..... micrococcal nuclease S7

MS ..... mass spectrometry

MS/MS ..... product ion scan

m/z ..... mass-to-charge ratio

NEG ..... negative

Ni-NTA ..... nickel nitrilotriacetic acid

ORF ..... open reading frame

PCR ..... polymerase chain reaction

PiCh ..... proteomics of isolated chromatin segments

PTM ..... post-translational modification

QUICK ..... Protein interaction screening by quantitative immunoprecipitation combined with knockdown

RNAPII ..... RNA polymerase II

RNA ..... ribonucleic acid

SAX ..... strong anion exchange

SBF ..... SCB binding factor

SDL ..... synthetic dosage lethality

SDS-PAGE ..... sodium dodecyl sulphate polyacrylamide gel electrophoresis

SGA ..... synthetic genetic array

SGD ..... *Saccharomyces* Genome Database

SILAC ..... stable isotope labeling with amino acids in cell culture

TAP ..... tandem affinity purification

TCA ..... trichloroacetic acid

TORC1 ..... target of rapamycin complex 1

TORC2 ..... target of rapamycin complex 2

TRP ..... tryptophan

Ub ..... ubiquitin

Ura ..... uracil

WB ..... Western blot

WT ..... wild type

Y2H ..... yeast-two hybrid

YPD ..... yeast extract, peptone, dextrose media

Chapter 1

Introduction

Part of section 1.4 to 1.5 of this chapter is published as:

Jean-Philippe Lambert, Kristin Baetz and Daniel Figeys. *Of proteins and DNA - proteomic role in the field of chromatin research*. Molecular BioSystems, 6 (1), 2010, 30-37.

Permission to reprint this work was obtained from *The Royal Society of Chemistry*.

Authors contributions:

JPL wrote the paper.

KB edited the manuscript.

DF edited the manuscript.

1.1 Introduction to proteomics

The term proteome was introduced in 1995 and represents the "protein complement expressed by a genome" (Wilkins et al., 1996). The study of a proteome with regards to its composition, post-translational modifications (PTM), expression level, localization or interconnection is referred to as proteomic. While the term proteomic is recent, it has its roots in the development of proper separation and detection methods for proteins which originated decades earlier. For instance, the advents of gel electrophoresis (Laemmli, 1970) and subsequently of high resolution two-dimensional gel electrophoresis (O'Farrell, 1975) were not only essential to the development of proteomic, but remain an integral part of most protein laboratories. Formerly, proteins were visualized with protein dyes such as coomassie brilliant blue (Fazekas de St Groth et al., 1963) but with time more sensitive detection reagents such as silver nitrate (Merril et al., 1981; Shevchenko et al., 1996) or SYPRO stains (Steinberg et al., 1996) were developed. This improved sensitivity toward proteins helped probe at a greater depth many organisms (reviewed in (Rabilloud, 2000)). Still in many ways, proteomic as we know it started with drastic improvements in the delivery of protein samples to mass spectrometers (Andersen and Mann, 2000) enabling the identification of protein in gel bands (or spots) following electrophoresis.

Key to this progress was the development of "soft" ionization techniques which enables large macromolecules to be effectively delivered in the gas phase as charged molecules prior to their analysis with mass spectrometer. One such technique is called electrospray ionization (ESI) (Fenn et al., 1989), see Figure 1-1 for a schematic of the ESI

apparatus, and was the ionization technique utilized during my graduate work. The other common ionization technique is the matrix-assisted laser desorption/ionization (MALDI) but it was not utilized for my work and will not be discussed here. ESI functions by generating a fine liquid aerosol of the sample by delivering it through a very small nozzle, or tip, while a potential is applied. An electric field between the nozzle and the mass spectrometer entrance is maintained to direct the charged droplets toward the mass spectrometer. As the charged droplets travel, they shrink as solvent molecules evaporates until too many charges are present concurrently in a single droplet causing an "explosion" producing desolvated charged molecules, or "quasi molecular" ions, see (Steen and Mann, 2004) for complete review of the ESI process. Since mass spectrometers are concentration sensitive devices, gain in sensitivity were obtained by miniaturizing the electrospray process (nanoESI) through the use of very fine tips (with an opening of 1-10μm) and reduced flow rate (<500nL/min) (Wilm and Mann, 1996). The use of nanoESI enabled the analysis of protein samples in the femtomolar range (Wilm et al., 1996) on par with the detection limit of common protein staining procedure for SDS-PAGE gels (Shevchenko et al., 1996).

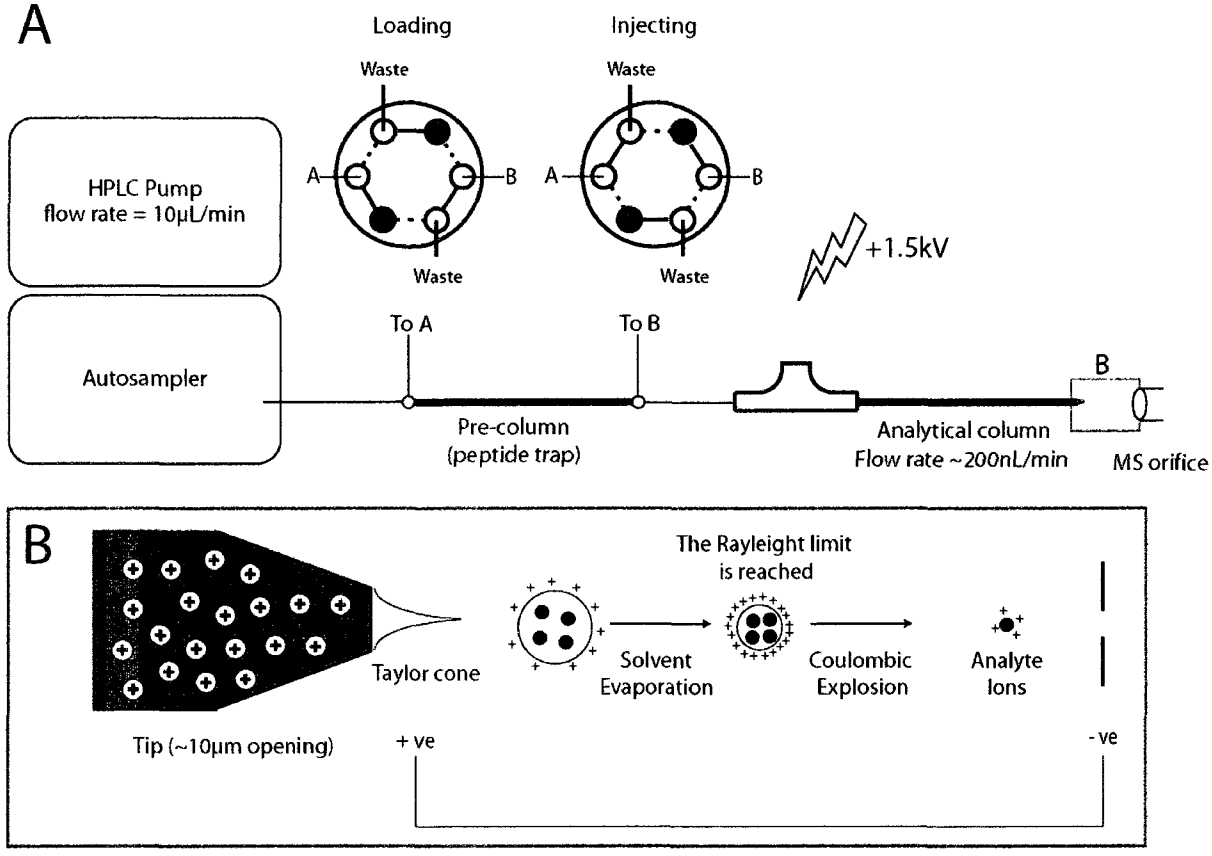


Figure 1-1: Principle of electrospray ionization and LC-MS/MS system. (A) Schematic of the LC-MS/MS system utilized during my graduate studies. A microcapillary pump coupled to an autosampler delivers the peptide mixture in a mobile phase composed of 95% water, 5% acetonitrile and 0.1% formic acid at 10 µL/min to a pre-column containing C_{18} material. After 10 minutes of desalting, the flow rate is reduced to 200 nL/min with a switch valve and the peptide directed to the analytical column by an increasing gradient of acetonitrile for electrospray ionization into a linear ion trap. **(B)** Model of electrospray ionization. A fine liquid aerosol of positively charged peptides is generated through a 10 µm tip while a potential is applied. An electric field between the tip and the mass spectrometer entrance is maintained to direct the charged droplets toward the mass spectrometer. As the charged droplets travel, they shrink as solvent molecules evaporates, resulting in a reduction in pH until too many charges are present concurrently in a single droplets causing an "explosion" producing desolvated charged molecules which can be analyzed in the mass spectrometer.

Concurrently with the development of nanoESI, significant advances occurred in protein separation systems. In particular, numerous interfaces were developed to couple liquid chromatography (LC) devices (such as high performance liquid chromatography (HPLC)) to mass spectrometers (MS) (Deterding et al., 1991) (Figure 1-1A). This required the miniaturization of LC systems to 1) align their flow rate with nanoESI functional range but also to 2) improve their separation capacity and thus minimize the complexity of the sample introduced in the MS (Deterding et al., 1991). Early examples of LC-MS uses for proteomic analysis resulted in the identification of a few proteins per experiments (Haynes et al., 1998; Yates et al., 1997). The experimental workflow utilized in these experiments includes the extraction of proteins from a sample of interest, their digestion into peptides by the trypsin protease, the fractionation of peptides by LC and subsequent analysis by MS. Peptide analysis begins by a MS scan to measure the mass-to-charge ratio (m/z) of peptides present at a given time during the LC-MS/MS analysis (Figure 1-2A). Then starting with the most intense peptide, five or more product ion scan experiments (MS/MS) are performed to acquire MS/MS spectrum of individual peptide (Figure 1-2B). The process described above is usually referred to as shotgun proteomics, in reference to the similarities with shotgun DNA sequencing (Anderson, 1981).

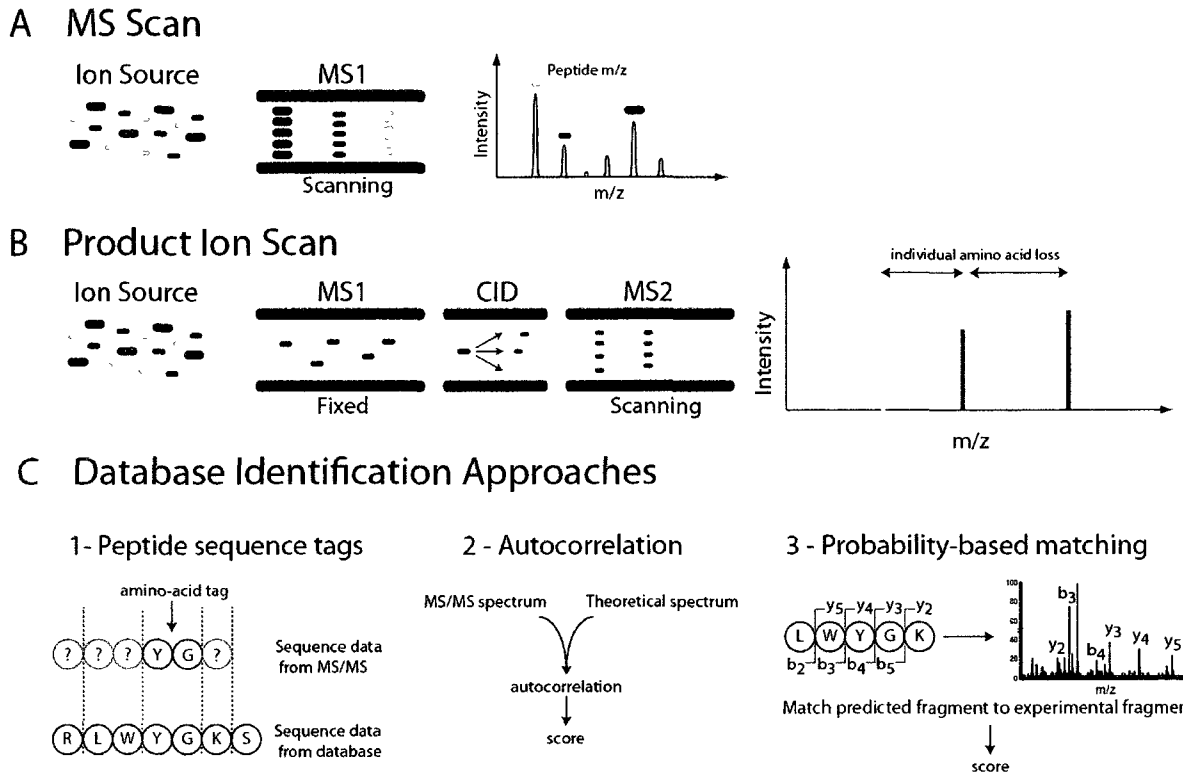


Figure 1-2: Methodologies of mass spectrometry based identification of peptides. (A) Following the introduction of a charged peptide in the gas phase into the mass spectrometer, it is directed to quadrupoles or an ion trap where it is scanned to determine its mass-to-charge ration (m/z). Subsequently, a charged peptide of interest present in the MS scan can be isolated, submitted to collision induced dissociation (CID) and the produced fragments measured to generate a MS/MS spectrum **(B)**. **(C)** Numerous bioinformatic approaches have been reported to match MS/MS spectrum to particular peptides. The peptide sequence tags approach relies on the fact that most MS/MS spectrum contains short amino acids sequence that can be easily identified in conjunction with masses observed before and after the amino-acid tag to determine the identity of a peptide (Mann and Wilm, 1994). MS/MS spectrum can also be compared to theoretical spectrum by autocorrelation to determine their overlap and report the best matching peptide sequences with scores (Eng et al., 1994). The most common approach for the analysis MS/MS spectrum relies on probability-based matching algorithm in which theoretical predicted peptide fragment from every sequences in the selected database are first computed. Subsequently, experimental tandem MS data (MS/MS) are matched to the predicted fragments in a top-down fashion beginning with the most intense ions and the best match reported with a score. One such algorithm is Mascot (Perkins et al., 1999) which was utilized in my graduate work.

A critical aspect of shotgun proteomics are the algorithms needed to properly assign the data generated by mass spectrometry to protein sequences. While a large variety of algorithms have been reported in the last 20 years (Figure 1-2C), Mascot remains one of the most commonly used software to analyze shotgun proteomics data (Perkins et al., 1999) and is the software I utilized during my work. The underlying mechanism employed by the Mascot software is to first compute theoretical predicted peptide fragments from every sequence in the selected database. Subsequently, experimental tandem MS data (MS/MS) are matched to the predicted fragments in a top-down fashion beginning with the most intense ions (reviewed in (Steen and Mann, 2004)). The integration of these bioinformatic tools in proteomics experiments is now common and has enabled high-throughput proteomics studies to be performed. For instance, analysis of haploid and diploid *Saccharomyces cerevisiae* cells by LC-MS/MS enabled the detection of approximately ~4300 proteins, representing the most complete study of this organism to date (de Godoy et al., 2008). Thus, proteomic is now a mature field which complements efficiently other "omics" field (e.g. genomic) as well as more traditional biochemical approaches resulting in an increased scope of life science studies.

1.2 Development of systems biology

The development of high-throughput tools enabled the emergence of systems biology, a shift away from the reductionist scientific approach toward larger *system-wide* studies. At its roots, systems biology attempts to fully comprehend a system through large-

scale quantitative high-throughput analysis and subsequent modeling of its structure, interconnections and dynamic behaviours (Kitano, 2002; Sauer et al., 2007). An elegant example of such a study is the report by Ishii et al. regarding *Escherichia coli* cells (Ishii et al., 2007). Complementary analysis of the transcriptome, proteome and metabolome of *E. coli* wild-type cells as well as 24 single-gene disruptant related to glycolysis and pentose phosphate pathway were performed (Ishii et al., 2007). Integration of this wealth of information enabled many conclusions to be drawn with regards to the *E. coli* system. In particular, the overall stability and robustness of the *E. coli* metabolic network proved to be higher than expected when challenged with various disturbances (Ishii et al., 2007).

An alternative definition of systems biology refers to the *systematic* study of biological systems components (Snyder and Gallagher, 2009). To date, budding yeast remains the model organism that has been studied the most in a systematic manner. While historical reasons did favour the emergence of this model organism, its genetic simplicity ensures its continual utilization (Snyder and Gallagher, 2009). *S. cerevisiae* possesses a small genome (12 million base pairs; ~6000 genes) but more importantly it can efficiently perform homologous recombination enabling facile genetic manipulations on a genome wide scale. As such, large-scale characterization of the budding yeast open reading frames (ORF) was successfully performed by systematic deletion (Winzeler et al., 1999) or transposon insertion (Ross-Macdonald et al., 1999). A subsequent study furthered the systematic characterization of ORFs by constructing an almost complete collection of gene deletion mutants encompassing over 96% of budding yeast predicted genes (Giaever et al., 2002). Building on the success of yeast knockout projects, a procedure termed synthetic

genetic array (SGA) was developed to quickly generate strains lacking pairs of genes (Tong et al., 2001). The analysis of the fitness of double mutant strains generated by SGA helps to uncover functional relationships between genes (Tong et al., 2001; Tong et al., 2004) and has now been applied to approximately one third of all budding yeast genes (Costanzo et al., 2010). An extension of the SGA procedure named epistatic miniarray profile (E-MAP) was also developed to study in a more quantitative fashion genetic interactions that can be both negative (aggravating) and positive (alleviating) with regards to cellular fitness (Collins et al., 2007b). E-MAP studies have been applied with success to chromosome biology (Collins et al., 2007b), protein phosphorylation (Fiedler et al., 2009), plasma membrane biology (Aguilar et al., 2010) and DNA damage repair (Nagai et al., 2008).

While analysis of gene knockout in budding yeast was extremely informative, systematic analysis of protein overexpression also provided a great tool for gene characterization. For instance, Spoko et al. prepared an ordered array containing 5280 strains in which a single ORF was conditionally overexpressed (Sopko et al., 2006). It was observed that 769 genes were toxic when overexpressed and that these genes were biased toward functions involved in cellular regulations (Sopko et al., 2006). Additionally, the overexpression array was successfully coupled with the SGA technology to systematically identify interaction between one gene deletion and one gene overexpression (Sopko et al., 2006) (termed synthetic dosage lethality (SDL) (Kroll et al., 1996)). This novel technology was successfully employed to identify new substrates for the Pho85 kinase whose substrates are all regulated by phosphorylation (Sopko et al., 2006). SGA and SDL screens were determined to be complementary in nature (Baetz et al., 2006) and have been

combined in the study of kinetochore (Measday et al., 2005) and of transcription factor (Berthelet et al., 2010). These systematic genetic studies, and many others, in budding yeast provided a wealth of information which has been critical in bringing down the number of uncharacterized *S. cerevisiae* ORFs to ~10% (Christie et al., 2009), a feat not reproduced in other model organisms.

Study of *S. cerevisiae* proteome also benefited from large-scale systematic efforts. As was previously described, the creation of specific *S. cerevisiae* collections enabled a wide-array of genetic studies. Along the same line, the yeast community has greatly benefited from the creation of collections where most ORFs have been tagged with green fluorescent protein (GFP) or with tandem affinity purification (TAP) tag under the control of their endogenous promoters (Howson et al., 2005). For instance, study of 4156 GFP tagged yeast ORF by microscopy enabled the localization of ~75% of the yeast proteome in mid-log phase (Huh et al., 2003). Remarkably, 70% of the proteins without previous localization information were successfully studied as part of this systematic study (Huh et al., 2003). Similarly, 4,251 TAP-tagged budding yeast ORFs were studied by comprehensive Western blot analysis to determine their expression level (Ghaemmaghami et al., 2003). This analysis revealed that ~80% of the yeast proteome is expressed during normal cell growth and that protein expression ranges from ~50 to over a million copies per cell (Ghaemmaghami et al., 2003). A subsequent study also utilizing the TAP-tagged collection was able to systematically define protein half-life (Belle et al., 2006). To do so, TAP-tagged strains growing in mid-log phase were treated with cycloheximide, an inhibitor of protein biosynthesis, and the TAP-tagged protein level determined by Western blot at multiple

time points (Belle et al., 2006). The mean protein half-life was determined to be ~43 minutes while 161 proteins were determined to be extremely unstable with a half-life of less than 4 minutes. By integrating the half-life data with previously reported mRNA abundance and ribosome abundance, it was recognized that most ORFs were either optimized for production (maximum production and maintenance efficiency) or for regulation (low production and fast degradation) (Belle et al., 2006). These large-scale proteomic studies emphasize the value of studying biological systems at the protein level, a focus of my thesis work described in chapter 2 and 3.

1.3 Method for the Identification of protein-protein interaction

As Anuj Kumar and Michael Snyder previously stated "*... no protein is an island entire of itself - or at least, very few proteins are*" (Kumar and Snyder, 2002). Whether proteins form homodimers, heterodimers or are part of large multisubunits complex, protein-protein interactions are critical to most protein function (Lamond and Mann, 1997). As such, a significant fraction of the proteomic field has been devoted to the development of appropriate techniques for the identification of protein-protein interactions and to their utilization. A popular method applied to the characterization of protein-protein interaction is the yeast-two-hybrid (Y2H) method (Fields and Song, 1989) (reviewed in (Fields, 2009)). The principle behind Y2H is to fragment a transcriptional activator, such as GAL4, into a DNA-binding fragment and an activation domain fragment which remain functional in the activation of a reporter gene when in close proximity. Then libraries of baits and preys

cDNA are fused to either fragment and transformed into yeast cells. Following mating of the appropriate cells, subsequent selection for the reporter gene and colony visualization, it is possible to detect protein-protein interaction on a large-scale efficiently (Ito et al., 2000). Y2H has been successfully applied to the majority of the yeast genome (Ito et al., 2001; Uetz et al., 2000; Yu et al., 2008). Y2H has been demonstrated to be efficient at detecting transient or condition specific binary interaction but not particularly adapted to characterizing protein complex (Yu et al., 2008).

The identification of protein-protein interaction has traditionally been performed through biochemical purification relying on physical characteristics of the bait under study. The advent of cloning technique for protein tagging has permitted the emergence of efficient affinity purification reagents such as the glutathione S-transferase (GST) (Smith and Johnson, 1988), FLAG (Brizzard et al., 1994) and TAP (Rigaut et al., 1999) tags. Standardization of the affinity purification reagents made protein complex purification a common tool in many laboratories around the world. It was quickly realized that the mass spectrometer was a detector of choice to couple with affinity purification to enable both sensitive and universal protein detection (Lamond and Mann, 1997). The first application of affinity purification and mass spectrometry (AP-MS) approach was for the characterization of the budding yeast U1 small nuclear ribonucleoprotein complex (Neubauer et al., 1997). Histidine-tagged Snp1 was purified by nickel nitrilotriacetic acid (Ni-NTA) affinity chromatography, the associated proteins resolved on SDS-PAGE gels, in-gel digested with trypsin and analysed by MS/MS (Neubauer et al., 1997). This effort resulted in the identification of 20 subunits of the yeast U1 small nuclear ribonucleoprotein (Neubauer et

al., 1997). Following this success, numerous other groups began utilizing AP-MS to map protein-protein interactions (Cooper et al., 2000; Rout et al., 2000; Stevens and Abelson, 1999).

Originally, AP-MS employed a single affinity step to remove proteins non-specifically associated with the protein complex of interest. In 1999, Rigaut et al. introduced a novel affinity purification procedure based on two orthogonal affinity steps considerably reducing background contamination (Rigaut et al., 1999). The tandem affinity purification (TAP) is a tag containing two protein A sequences, a TEV cleavage site and a calmodulin binding peptide sequence (Rigaut et al., 1999) (Figure 1-3A). TAP purifications are performed by first using immunoglobulin (IgG) coated sepharose beads to bind the protein A sequences in the TAP-tagged bait. Then the beads are extensively washed and the TAP tag cleaved with the TEV protease releasing the partially purified protein complex. The second affinity purification step utilizes calmodulin beads which can bind to the calmodulin binding peptide in the presence of calcium. Following washing of the calmodulin resin, a calcium chelator such as ethylene glycol tetraacetic acid is introduced to elute the purified protein complex for MS analysis (Figure 1-3A). The capacity to produce large amount of protein complexes of high purity was extremely beneficial when TAP was developed since mass spectrometers duty cycle was still high (i.e. slow scan speed). As such, removal of background contaminants enabled the mass spectrometer to focus on the relevant proteins and thus facilitated the detection of protein-protein interaction.

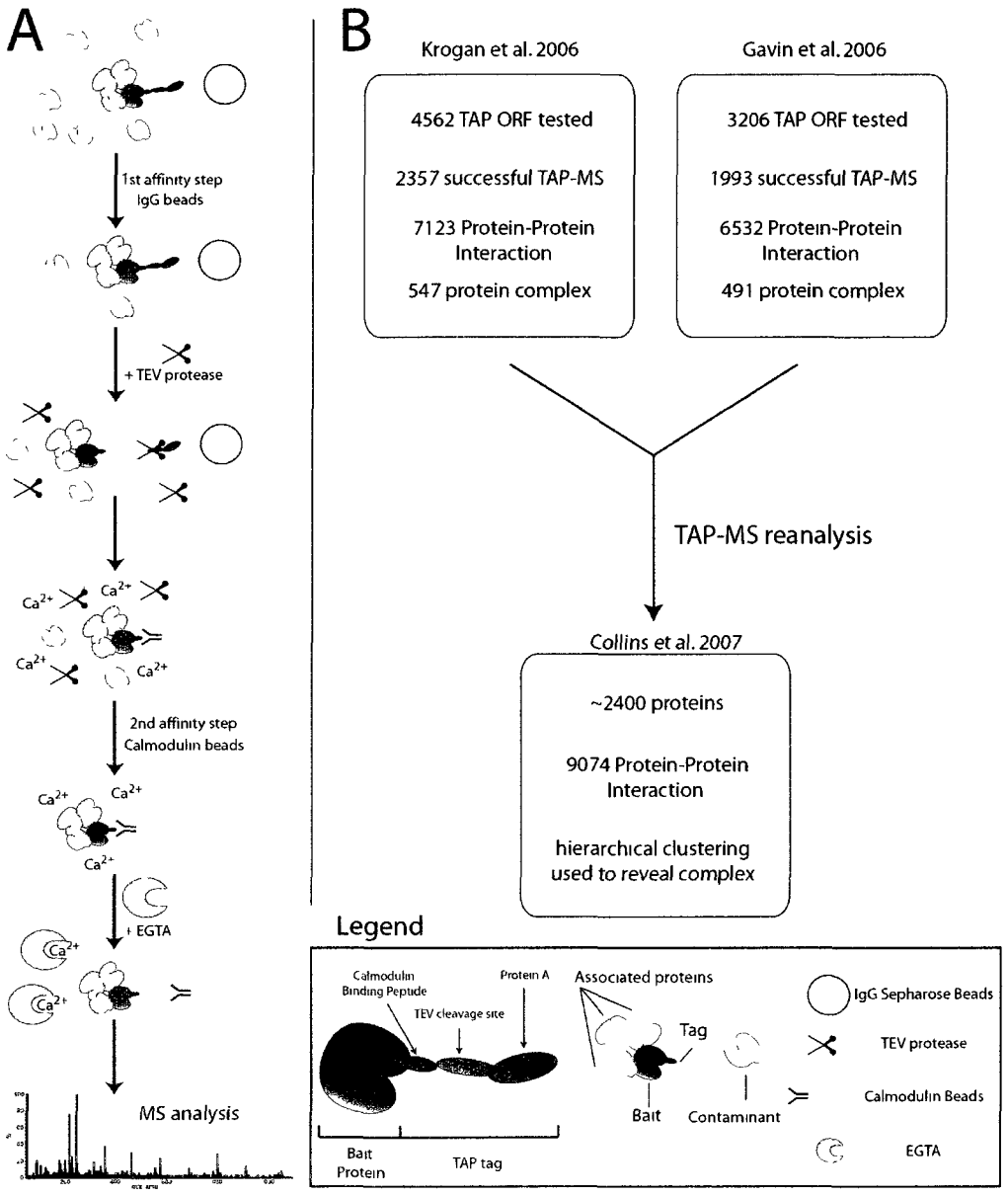


Figure 1-3: Principle of tandem affinity purification and its systematic application in budding yeast. (A) Schematic of tandem affinity purification as developed by Rigaut et al. (Rigaut et al., 1999). Whole cell extract from strain expressing a TAP tagged bait under the control of its native promoter is clarified and incubated with beads coated with rabbit immunoglobulin (IgG). Following extensive washing of the beads, TEV protease is added to cleave the TAP tagged fused to the bait thus releasing it along with its interaction partners. Calcium is added to the partially purified protein complex and subsequently calmodulin beads for a second purification step. Following more washing, a calcium chelator is added to elute the highly purified bait with its interaction partners ready for MS analysis. **(B)** The TAP-MS approach was systematically applied in budding yeast by two groups to characterize hundreds of protein complexes (Gavin et al., 2006; Krogan et al., 2006). Subsequently, the data from these two studies was unified through bioinformatic reanalysis (Collins et al., 2007a).

Systematic application of the TAP-MS procedure was performed by two independent group on a genome-wide scale in budding yeast (Gavin et al., 2006; Krogan et al., 2006) (Figure 1-3B). These studies together observed over 9,000 distinct pair-wise protein-protein interactions involving the majority of the yeast proteome. Through refinement of the raw association data using either a Markov clustering algorithm (Krogan et al., 2006) or socio-affinity and iterative clustering (Gavin et al., 2006), 547 and 491 protein complexes were generated, respectively. The use of two different approaches in the identification of protein complexes prevented an easy integration of these datasets. To circumvent this issue, reanalysis of both TAP-MS datasets was performed as a whole utilizing a purification enrichment scores (Collins et al., 2007a) (Figure 1-3B). Hierarchical clustering of the resulting high confidence protein-protein interaction dataset was performed to visualize protein complexes. This method proved sensitive at detecting known and novel protein complexes but also at identifying weak or low confidence protein interaction. In particular, 7,504 of the 12,122 protein-protein interaction identified by TAP-MS (Gavin et al., 2006; Krogan et al., 2006) representing ~61% of all protein-protein interaction reported were marked as being of low confidence (Collins et al., 2007a).

Despite the success of the TAP-MS approach, it is important to note that it possesses some inherent flaws. For one, the utilization of two affinity steps affects negatively the sample recovery, especially the calmodulin binding step which leads to losses of more than 50% of the bait (Rigaut et al., 1999). In budding yeast this can be offset in part by the ease of growing large cultures but for other systems this is a significant drawback. Another limitation of the TAP-MS approach is that it is often incapable of

detecting transient or weak protein interactions that are lost as part of the extensive purification process. Finally, TAP-MS was observed to be biased toward abundant cytosolic protein (Jensen and Bork, 2008).

Large-scale single step AP-MS studies have also been performed in budding yeast. In 2002, Ho et al. utilized an overexpression system based on the *GAL1* galactose inducible or on the tetracycline inducible *tet* promoters to study 725 FLAG tagged proteins (Ho et al., 2002). Baits used in this study were selected based on their functions (e.g. kinase) or involvement in biological pathway (e.g. DNA damage response). The use of the small hydrophilic octapeptide FLAG tag enabled faster, less stringent affinity purification to be performed resulting in a more encompassing detection of protein-protein interaction than TAP-MS. Following data curation, 3,617 pair-wise protein interactions between 493 baits and 1,578 preys were reported. In addition, numerous proteins with link to DNA damage response were submitted to AP-MS following treatment with DNA damaging agents such as hydroxyurea or methyl methane sulfonate. As such, dynamic changes in protein-protein interaction could be captured providing an additional layer of information on the DNA damage response network.

An intrinsic property in the use of a single affinity purification step is an increase in the amount of proteins co-purified with the bait of interest (Glatter et al., 2009). This affects both the interaction partners and proteins non-specifically associated with the bait increasing the complexity of the sample to analyse by MS. Thus more proteins must be identified by MS following a single affinity step than for TAP. It is interesting to note that

the 3D ion trap mass spectrometers and hybrid quadrupole time-of-flight mass spectrometers used prior to 2005 possessed a high duty cycle, the measure of the time needed to perform MS and MS/MS scans. As such, mass spectrometers were often overwhelmed with peptides resulting in high false negative levels for the detection of peptides (Tabb et al., 2005). Consequently, single step AP-MS was considered less effective than TAP-MS for protein-protein interaction mapping since it produces data with a higher background level while missing some known interaction partners (von Mering et al., 2002).

AP procedures utilizing at least two purification steps remained the norm until key development in the field of proteomic (Figure 1-4). The first is the introduction of significantly better mass spectrometer utilizing 2D linear ion trap (Schwartz et al., 2002) and subsequently "orbitrap" (Makarov et al., 2006) which reduces by more than fivefold the mass spectrometers duty cycle while improving their sensitivity and mass accuracy (Blackler et al., 2006). These novel mass spectrometers now enable proteomic researchers to "dig deep" in a proteome and characterize complex samples efficiently. Another significant development was the adoption of solid support for affinity reagents with deactivated surfaces reducing co-purification of non-specifically associated proteins. For example, magnetic beads which can be coated with antibodies or various antigens, have been beneficial to AP (Archambault et al., 2003; Page et al., 1999). Moreover, the magnetic beads further reduce background contaminants in AP experiments due to a more thorough bead washing enabled by the reduced bead volume. Finally experimental procedures were developed to properly segregate interaction partners from non-specifically associated proteins. For example, Selback et al. reported the combination of quantitative proteomic,

RNA interference and affinity purification to map protein-protein interaction (Selbach and Mann, 2006). This method called QUICK, for Protein interaction screening by quantitative immunoprecipitation combined with knockdown, enabled researchers to circumvent background contaminations and selectively identified protein-protein interactions. This method was an extension of the seminal report by Blagoev et al. which was the first to combine stable isotopic amino acids in cell culture (SILAC) with AP-MS (Blagoev et al., 2003). Briefly, SILAC functions by introducing isotopically labelled amino acid in a cell culture (Ong et al., 2002). These "heavy" amino acids are incorporated in proteins and upon MS analysis can be distinguished from unlabelled or "light" amino acid due to their increased mass (see (Ong et al., 2003) for an extensive discussion of SILAC). The use of quantitative proteomic in AP-MS experiments enabled Blagoev et al. to study how the interactome of the SH2 domain of Grb2 was rapidly modified following epidermal growth factor treatment (Blagoev et al., 2003). To do so, two cell populations are grown using the same conditions, but with light or heavy amino acids, and mixed prior to affinity purification. Then upon MS analysis it is possible to determine the provenance of each proteins based on their mass resulting in relative quantitation of the samples. This approach was further extended by the use of triple labelled SILAC to identify interaction partners specifically associated with two different protein phosphatase 1 isoforms over an untagged WT control (Trinkle-Mulcahy et al., 2006).

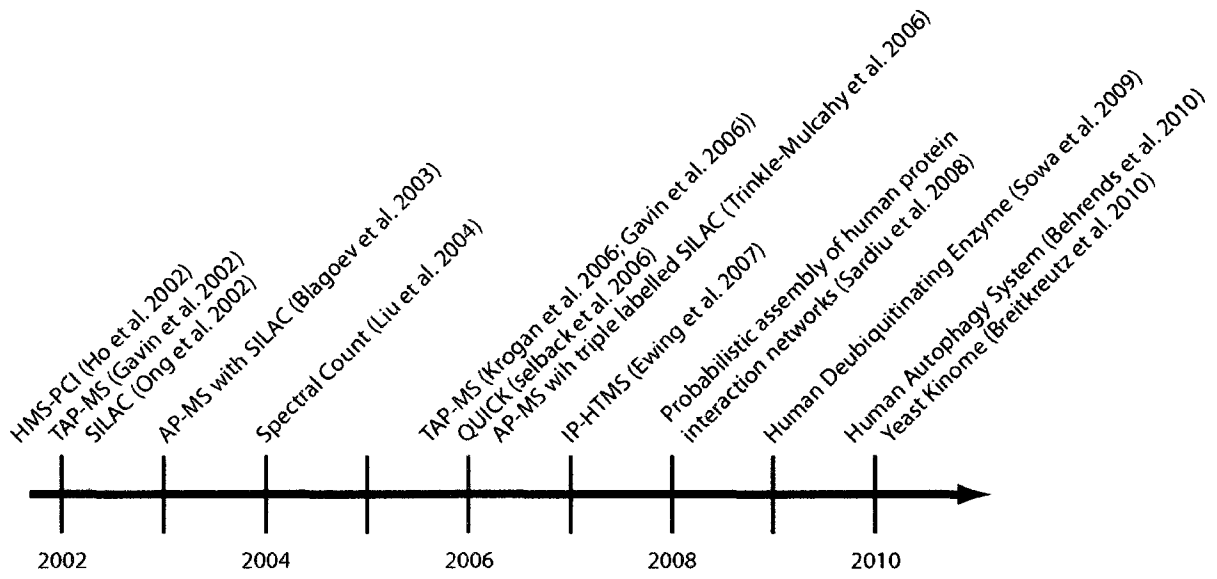


Figure 1-4: Key development and reports with regards to large-scale AP-MS. See text for additional details.

Label-free quantitative approaches can also be used effectively for AP-MS. One quantitative information readily available in shotgun proteomic studies is spectral count. Spectral count refers to the total number of peptide matching to a given protein sequence in one LC-MS/MS analysis and was determined to correlate with protein abundance (Liu et al., 2004). By comparing the spectral count for a given protein between two or more experiments, it is possible to infer the relative abundance of the target protein in the samples. Utilizing spectral count data in combination with vector algebra and statistical methods, a probabilistic approach to the study of protein-protein interaction was developed to analyze the data generated by 27 single step AP-MS experiments (Sardiu et al., 2008). Based uniquely on label-free quantitative information, it was possible to remove contaminant proteins (background) and to identify protein complexes. Chen et al. also shown that a simple hierarchical clustering of AP-MS data was sufficient to identify interaction partners (Chen et al., 2009). Endogenous affinity purification of the Miwi and Mili proteins was performed; the purified proteins analyzed by LC-MS/MS and the resulting data clustered to distinguish between contaminants and interaction partners. Their approach was both successful and simple representing an attractive option for AP-MS studies.

Above, I introduced the principle foundation of the proteomic field. In particular, I described how affinity purification when coupled to mass spectrometry can be used to systematically map protein-protein interaction on a large-scale. The significant developments described here changed the proteomic field from specialized analytical tools to a mainstream part of life sciences. In addition, systematic proteomics studies were made

possible by standardization of the reagents and tools utilized. Unfortunately, current proteomic tools are not adequate for the study of all aspects of the proteome, such as DNA associated proteins. My thesis work revolved around the development of a novel tool to characterize the interface between DNA and protein and its use in the large-scale study of chromatin associated proteins.

1.4 Chromatin - The Interface between Genes and Proteins

Proper maintenance and regulation of an organism's genetic material are essential for its survival. It is thus unsurprising that all organisms possess a wide array of redundant mechanisms to not only prevent damage and mutations to their genetic material but also to control its access. Seminal work performed over thirty years ago demonstrated that DNA is organized around a core protein complex, termed the nucleosome (Kornberg, 1974). Nucleosomes are octomeric protein complexes composed of two copies of the four different canonical histones (H2A, H2B, H3 and H4), the crystal structure of which has been determined at high resolution (Luger et al., 1997). DNA is wrapped around nucleosomes to not only package or compact its unruly length but also to regulate the access of other biomolecules to the genetic material. Chromatin, the combination of protein and DNA, also involves the functions of numerous non-histone proteins. For instance, the histone-dependent restriction of DNA access is opposed by the sequence specificity of transcription factors and other chromatin-associated proteins, such as nucleosome remodelers and

modifiers, resulting in a well orchestrated regulation of gene expression (Wang et al., 2007).

Early on it was recognized that another abundant histone protein, histone H1, participated in increasing the chromatin compaction, and formation of chromatin's most condensed form, the chromosomes. Furthermore chromatin researchers observed that a wide array of histone variants existed in addition to the core nucleosome members. These histone variants are of great interest since many possess specialized roles and functions (Boulard et al., 2007; Izzo et al., 2008; Loyola and Almouzni, 2007). Numerous proteins and protein complexes have been implicated in modifying the composition of core nucleosomes with histone variants at specific genomic loci, at centromeres for instance, or under specific cellular conditions, such as DNA damage. In addition to the exchange of core histone for their variants, chromatin is also controlled by an intricate number of histone post-translational modifications (PTM) (Taverna et al., 2007a). A wide array of enzymes including lysine acetyltransferases (KAT), lysine methyltransferases (KMT), kinases, E3 ubiquitin ligases and many more are known to modify one or many residues of histone proteins. This plethora of modifications act as signals, or anchors, for chromatin binding actors and this phenomenon is termed the histone code (Strahl and Allis, 2000).

Histones are small basic proteins that are present in all eukaryotes that can be grouped as either core histones (H2A, H2B, H3 and H4) or linker histones (H1). Histones were discovered more than a hundred years ago by Albrecht Kossel as an acid soluble class of proteins associated with nucleic acid; this is some of the work for which he received the

Nobel Prize of Physiology and Medicine in 1910. The fact that histones are soluble in acidic solutions has been exploited for a long time to enable their purification (see reference (Shechter et al., 2007) for an overview of the acid extraction). Early work often used electrophoresis methods to study the acid extracted histones but in recent years mass spectrometry has become the method of choice to analyze histones and their post-translational modifications (reviewed in reference (Garcia et al., 2007)). For instance, a number of groups have embarked on the cataloguing of histones present in various organisms using mass spectrometry. Bonenfant et al. performed such a study to better characterize the histone H2A and H2B isoforms present in Jurkat cells (Bonenfant et al., 2006). By using both top-down (intact proteins) and bottom-up (enzymatically digested proteins) approaches, they were able to identify nine histone H2A and eleven histone H2B variants. In addition, numerous post-translational modifications, such as phosphorylation and acetylation, were detected on some of these variants. Another similar study was performed by Thomas et al. in which they used top-down mass spectrometry to study the histone H3 isoforms from HeLa cells (Thomas et al., 2006). The authors were able to easily distinguish three histone H3 isoforms containing a wide number of post-translational modifications. It was also observed that acid extracted histone H3.1 from colchicine treated cells was a heterogeneous population containing zero, one or two phosphate moieties on serine 10 and 28. From these top-down studies, and other work, it becomes quite apparent that a single histone molecule can contain a diversity of post- translational modifications at once. This was further reinforced by the work of two independent groups which used mass spectrometry to study histone protein-PTM relationship. Jiang et al. analyzed acid extracted

histone H3 from budding yeast following their enzymatic digestion with the endoproteinase Glu-C (Jiang et al., 2007). This enzyme digests a protein into long peptides since it selectively cuts peptide bonds at the C-terminal of glutamic acid. Glu-C digestion of histone H3 produces a large peptide of its heavily modified N-terminal tail containing the first 50 amino acids. Analysis of this peptide demonstrated that certain modifications are linked, i.e. present simultaneously on one peptide. In particular, the authors observed that Set1-dependent methylation of H3 lysine four (H3K4) occurred concurrently with acetylation of lysine 9, 14 and 23. Taverna et al. performed a similar study in the ciliated protozoan *Tetrahymena thermophilia* (Taverna et al., 2007b). Acid extracted histone H3 was resolved on a PolyCAT A column, digested by Glu-C and analyzed by a mass spectrometer capable of performing electron transfer dissociation fragmentation and proton transfer charge reduction. The use of this modern mass spectrometer enabled the first fifty amino acids of histone H3 to be fully sequenced in a single scan. Using this approach, the diversity of PTMs on histone H3 was discerned and a clear connection between certain methylation and acetylation modifications established. These highly significant technical achievements overshadow the fact that it remains extremely difficult to analyze histones and their PTMs. For instance, many of the histone H2B isoforms present in human cells identified by Bonenfant et al. only differ from one another by a single or a few amino acids (Bonenfant et al., 2006). This means that distinguishing these isoforms in shotgun proteomic experiments is extremely challenging since most peptide identified will match equally well with multiple histone proteins. Furthermore, the detection of some large PTM on histones, such as ubiquitination, or labile PTM, such as phosphorylation, by mass spectrometry requires

forethought and experimental planning. As such, proteomic studies of chromatin still require innovation in both the instruments and methodologies in order to fully grasp the complexities and subtleties of histone biology.

Histone deposition on DNA is also heavily regulated by proteins called histone chaperones. While nucleosomes will assemble in the absence of histone chaperone *in vitro*, the resulting chromatin is aggregated and insoluble at physiological ionic strength (Kornberg, 1974). Furthermore, histones are not "free" but rather associated with histone chaperones in cell extracts of yeast and other organisms (Eitoku et al., 2008). Histone chaperones are thus critical to chromatin function. Histone chaperones are defined as *"factors that associates with histones and stimulates a reaction that involves histone transfer without being part of the final product"* (Probst et al., 2009) (Figure 1-5A). Interestingly, histone chaperones exhibit a large degree of variation in their structure but also in their affinity for histone proteins (reviewed in (Hansen et al., 2010)). For instance Nap1 has a high affinity for histone H2A/H2B dimers whereas Asf1, Rtt106 and Hir1 associate with histone H3/H4 tetramers. These chaperones were studied as part of my thesis work and will be discussed in details in chapter 3.

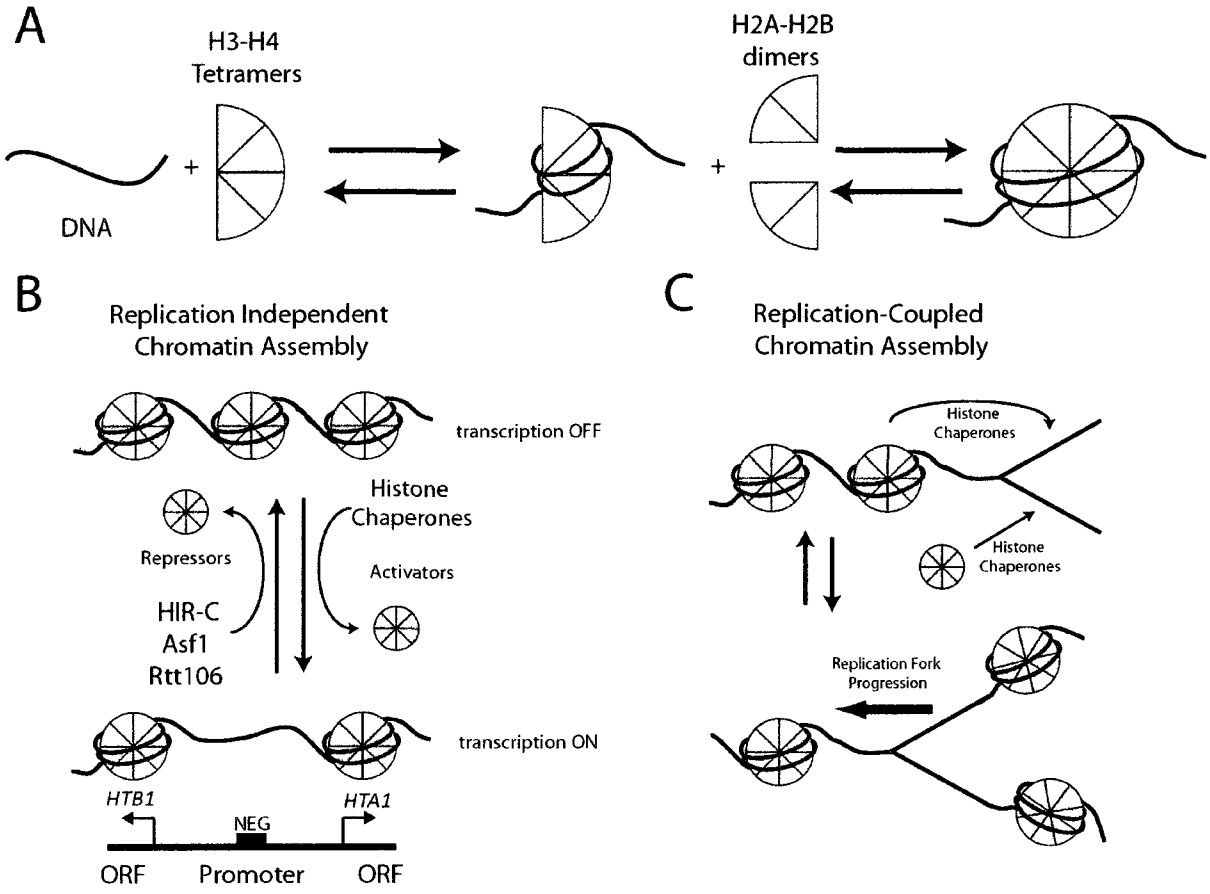


Figure 1-5: Roles of histone chaperones in nucleosome assembly and disassembly. (A) General principle of nucleosome assembly. Histone H3/H4 tetramers are deposited onto DNA with the action of histone chaperones. Subsequently, two histone H2A/H2B dimers will be deposited to form complete nucleosome. **(B)** Nucleosomes are refractory to gene transcription and must be displaced for transcription to occur. For instance, the histone chaperones Asf1, HIR and Rtt106 are repressors of the *HTA1-HTB1* promoter by promoting nucleosome deposition at this promoter. In general prior to transcription, histone chaperones will create nucleosome free region at promoter which will permit the recruitment of the transcription machinery. Following transcription chromatin will be reassembled at the promoter to enable its proper regulation. **(C)** As DNA is replicated, chromatin must be disassembled ahead of the replication fork and reassemble following it by the action of histone chaperone. This can be a conservative process where nucleosome are recycled or formed from newly synthesized histones (Xu et al., 2010).

Nucleosomes are refractory to gene transcription and need to be removed or displaced to enabled transcription (Knezetic and Luse, 1986; Lorch et al., 1987). Numerous histone chaperones are involved in this process (see Figure 1-5B for schematic). A well studied model of histone chaperone based transcription regulation in *S. cerevisiae* is the histone genes themselves. In budding yeast, the two different isomers of histone H2A (HTA1, HTA2) and H2B (HTB1, HTB2) are transcribed from a pair of divergent promoters (*HTA1-HTB1*; *HTA2-HTB2*) whereas histone H3 and H4 are transcribed from another pair of promoter (*HHT1-HHF1*; *HHT2-HHF2*). Interestingly, a central regulatory region in the promoter region of *HTA1-HTB1*, *HHT1-HHF1* and *HHT2-HHF2* termed NEG is sufficient to control the expression of these histone genes (Osley et al., 1986). Histone genes are repressed throughout the cell-cycle with the exception of S-phase where histone proteins numbers are doubled to properly assemble newly synthesized DNA into nucleosomes (Gunjan et al., 2005). Previous work showed that two histone chaperones, the nucleosome assembly factor Asf1 as well as the HIR complex, are involved in the repression of the *HTA1-HTB1* promoter (Green et al., 2005). Recently, a third histone chaperone, Rtt106, was reported to participate in the repression of the *HTA1-HTB1* promoter (Fillingham et al., 2009). This was achieved by combining genomic screens, biochemical studies and some proteomics experiments, which I performed (see Appendix 1-1 for details regarding my publications not covered directly in this thesis). By enabling proper nucleosome assembly at the NEG, these histone chaperones participate in the control of histone genes transcription. Chapter 3 will expand on this notion with a detailed study of the physical association between histone chaperones involved in the regulation of the *HTA1-HTB1* promoter.

The activity of numerous histone chaperones is also coupled to DNA replication since nucleosomes must be disassembled in front of the replication fork and reassembled after it (Figure 1-5C). Two significant challenges must be overcome as DNA is replicated. For one, the information present in the nucleosome (such as histone PTMs) must be preserved and propagated to newly synthesize histone proteins. Furthermore, damage to exposed (or nucleosome-free) DNA must be prevented. One histone chaperone involved in this process is the chromatin assembly factor 1 complex (CAF-1) which is recruited directly at the replication fork where it participates in nucleosome assembly and disassembly (Smith and Stillman, 1989). Asf1 is also involved in replication coupled chromatin assembly and directly interacts with CAF-1 (Mello et al., 2002) as is Rtt106 (Li et al., 2008). These chaperones can function by either recycling current nucleosomes, incorporate newly synthesized ones or assemble nucleosomes composed of old and new histones. The exact mechanism by which CAF-1, Asf1 and Rtt106 functions at the replication forks is still under debate (Probst et al., 2009).

To truly understand the dynamic and complex nature of chromatin, a detailed understanding of the combinatorial role of histone and of their chaperone along with the chromatin-associated proteins they recruit is required. The field of chromatin research has rapidly progressed in recent years due to the development of novel technologies. Advances in molecular biology, genomic sciences, structural biology and proteomics enabled the field to grow to its current state. The role of proteomic technologies in the study of chromatin as well as emerging methods to study chromatin-associated proteins will be discussed below.

1.5 Proteomic Tools to Study Chromatin Associated Protein Networks

Mass spectrometry has had a deep impact on the field of chromatin research by enabling researchers to better observe the diversity in both the histone composition and their modifications. But mass spectrometry driven proteomic research also enabled a better understanding of the composition of chromatin itself. An early report by Shiio et al. attempted to better characterize the protein content of chromatin with and without the expression of the oncogene Myc (Shiio et al., 2003). This was achieved by first isolating nuclei from P493-6 human B lymphocytes and subsequently isolating the chromatin fraction prior to ICAT labelling and mass spectrometric analysis. This study resulted in the identification of 64 nuclear proteins and their relative quantitation upon Myc induction. While this study was small in scope, it introduced the notion that proteomics is not only valuable for the identification of chromatin bound proteins but can also be exploited to study the dynamic nature of chromatin protein networks.

A major shift in the perception of protein–protein interactions has started to emerge in recent literature. In particular, the view of protein complexes has been extended to increasingly include a macromolecular, or network, view of the interconnections between protein complexes. This is in part derived from the emergence of systems biology and has been fuelled by technical improvements in the analysis of these connections whether they are between protein, DNA, RNA or other types of biomolecules. For instance, a recent bioinformatic study (Wang et al., 2009) re-analyzing the Saccharomyces cerevisiae TAP protein interaction datasets (Gavin et al., 2006; Krogan et al., 2006) produced ComplexNet, a network of interactions between proteins and protein complexes. This

effort exposed novel associations, such as a physical link between a poorly characterized protein complex composed of Yer071c and Yir003w, and the yeast actin-capping heterodimer.

Despite these recent advances, defining the interaction networks of chromatin-associated proteins using conventional methodologies remains challenging. This problem is exemplified by the various protein complexes involved in transcription. For instance, RNA polymerase II (RNAPII) has been extensively studied in many organisms (Cho et al., 1998; Maldonado et al., 1996). Yet despite all of these efforts, novel RNAPII-associated proteins were only discovered once technical improvements were made to alleviate the challenge posed by the presence of DNA (Aygun et al., 2008; Jeronimo et al., 2007). A common view in the field of chromatin proteomic is that the presence of DNA hinders the effective purification of protein complexes, thus DNA should be excluded from the sample under study. This was first recognized many years ago in a seminal study by Lai and Herr which demonstrated that the common laboratory reagent ethidium bromide (EtBr) could be used to disrupt DNA–protein interactions (Lai and Herr, 1992). The authors effectively demonstrated that their affinity purification of the transcription factor Oct-2 from 293 cells suffered from non-specific association of some proteins with the DNA that was co-purified. Removal of DNA by chemical treatment with EtBr, or with enzymatic digestion of DNA by micrococcal nuclease (MNase), was sufficient to abolish these non-specific associations. This study was critical since it described a fast and efficient way to determine whether a protein–protein interaction was dependent on the presence of DNA. But more importantly, it also showed that some DNA–protein associations could be enriched in the laboratory.

The importance of this fact is only starting to be recognized in the literature (Rusk, 2009). The practice of removing DNA prior to affinity purification remains common to this day in many laboratories. The reasons for this are multiple and include a reduced viscosity of the samples, improved solubility of large complexes and streamlined sample preparation. A good example of this strategy was reported by Foltz et al. for the study of CENP-A, a centromeric histone H3 variant, in human cells (Foltz et al., 2006). The authors used MNase treatment to drastically reduce DNA size in their cellular extracts prior to two step purification of CENP-A and histone H3. The resulting mass spectrometric analysis of protein networks enabled the identification of numerous new CENP-A interaction partners which were shown to possess critical roles in the maintenance of chromosome stability (Foltz et al., 2006).

A similar procedure was used by Du et al. to study H2A.X, a histone variant critical for efficient DNA damage response (Du et al., 2006). The authors used DNase I treatment to completely remove DNA from their cell extract prior to purification of H2A.X-FLAG and subsequent analysis by mass spectrometry. This type of procedure enabled H2A.X interaction partners to be specifically observed under DNA damage conditions and not under control conditions. Benzonase, an endonuclease digesting both DNA and RNA, is another reagent that is commonly used to degrade chromatin and thus preventing DNA-protein association from being observed (Chen and Gingras, 2007). Reagents affecting protein-DNA interaction are often used for validating previously observed protein-protein interaction. For instance, a report studying the transcription factor Oct-4 observed an interaction with the oestrogen receptor beta which was validated to be a direct interaction,

and not to be mediated through chromatin, since it was insensitive to both benzonase and EtBr (van den Berg et al., 2008). While these studies are successful at purifying protein complexes off chromatin, they forgo information in doing so, particularly about the chromatin environment surrounding the baits. Thus, researchers attempting to obtain a more holistic view of chromatin must use different protocols in which the DNA remains fully or partially intact.

Early attempts at performing functional proteomics studies of chromatin-associated proteins often used protocols not designed especially for this unique class of proteins; hence, these studies suffered from both sensitivity and specificity issues. A good example of this problem is the use of the well established chromatin immunoprecipitation (ChIP) method which permits the identification of a protein's DNA binding site (Massie and Mills, 2008). ChIP is typically composed of a few key steps: (1) stabilization of protein–DNA interaction through the use of chemical crosslinkers, such as formaldehyde; (2) DNA shearing to improve the solubility of the protein–DNA complexes; (3) affinity purification of the complex of interest using antibodies; (4) amplification and analysis of the DNA of interest by some means (PCR, Chip, etc.). At first glance the general idea of ChIP appears directly applicable to proteomics studies; however, one major problem needs to be addressed, the impossibility of amplifying protein samples prior to their analysis. Thus, direct transfer of the ChIP protocol to proteomic studies produces less than optimal results, mainly due to low sample recovery and the heterogeneity of the sample (Ricke and Bielinsky, 2005). In short, researchers attempting to study chromatin networks need to purify large amount of material, of high purity, at a reasonable cost. Proteomic approaches

attempting to do so will be discussed in details below along with their shortcomings (Table 1-1; Figure 1-6).

Déjardin and Kingston presented a new method termed proteomics of isolated chromatin segments (or PICh) for the analysis of proteins associated with specific chromatin loci (Dejardin and Kingston, 2009). The PICh method relies on nucleic acid probes that recognize specific genomic loci which are then enriched with their associated proteins (Fig. 1-6B). The procedure begins by fixing approximately 3 x 10^9 cells with formaldehyde which stabilizes both protein-protein and protein-DNA interactions. The cells were then lyzed and the chromatin solubilized by sonication. To specifically purify the genomic loci of interest, a 25 base pair probe made of locked nucleic acid (which possesses a higher melting temperature than a regular nucleic acid) linked to a desthiobiotin moiety was used. The probe was efficiently hybridized with the chromatin samples under stringent detergent conditions and then subsequently purified using streptavidin beads and eluted with excess biotin. The purified proteins were resolved on a SDS-PAGE gel and identified by mass spectrometry. This new technique was first applied for the purification of proteins associated with telomeres. Telomeres were selected since they contain repetitive sequences which significantly reduced the amount of material needed per experiment. The authors used a probe directed at telomere and a probe with the same nucleic acid composition but in a randomized order as a control. Three cell lines were studied, two of which had telomerase activity, the enzyme responsible for maintaining telomere length, and one cell line defective in telomerase activity which served as a comparative control. The authors purified approximately 200 proteins associated with telomere chromatin, but

not with the scrambled probe, in the different cell lines and approximately half these hits were shared between cell lines. 33 proteins observed by mass spectrometry following PICh purifications, such as Apollo and Ku70, were previously reported to be present at telomere, supporting the specificity of the method. Colocalization of putative telomere proteins with the telomeric protein Rap1 by indirect immunofluorescence along with ChIP experiments were used to validate the PICh datasets. For most novel proteins purified by PICh with telomeres, the secondary validations supported their telomeric localization and clearly demonstrated the strength of the new method. One drawback of the PICh procedure is its low sensitivity which translates in needing really high amount of starting material needed per experiment (Table 1-1). By targeting a specific DNA sequence which is present at one or a few copies per cell, it becomes extremely difficult to purify sufficient associated proteins for mass spectrometry analysis. While PICh represents a significant step forward in the study of chromatin-associated protein complexes it remains to be seen how widely applicable the method will become.

Table 1-1: Characteristics and performances of methods to study chromatin bound proteins.

Method	Publication	Soluble Protein-Protein Interaction	Chromatin Bound Protein-Protein Interaction	Sensitivity	Purification Specificity
AP MS	1997	Yes	No	Variable	Variable
mChIP	2009	Yes	Yes	High	High
PiCh	2009	No	Yes	Very Low	Low
Minichromosome Purification	2009	No	Yes	Low	Low
SILAC-Based DNA Interaction Screen	2009	No	Yes	Low	Low

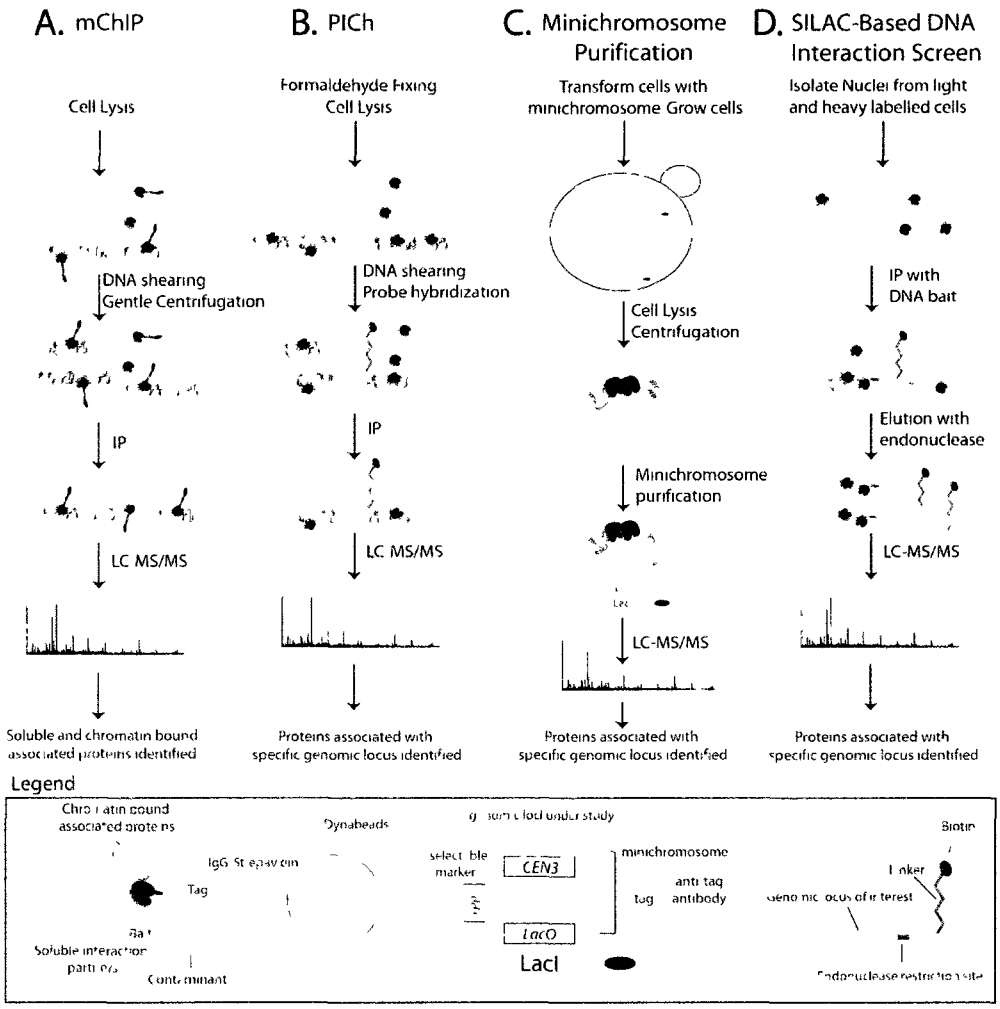


Figure 1-6: Proteomic schemes for the study of chromatin bound proteins. (A) The modified chromatin immunopurification (mChIP) method developed during my graduate studies enables the purification of chromatin associated protein networks. The mChIP method is described in detailed in chapter 2 and was utilized in a large-scale in chapter 3. **(B)** The PICh method uses a DNA probe to selectively purify proteins associated with a genomic locus. **(C)** Minichromosomes containing a genomic locus of interest can be isolated and the proteins associated with it identified by LC-MS/MS. **(D)** The SILAC-based DNA interaction screen uses quantitative proteomics and in vitro reconstituted protein–DNA interaction to identify proteins associated with a genomic locus.

The most straightforward solution to the problem created by the low abundance of a genomic locus in proteomic studies is to artificially increase its abundance. One example of this is the report by Akiyoshi et al. who utilized minichromosomes containing centromeric DNA from the third chromosome of budding yeast (*CEN3*) to study the proteins associated with this locus (Akiyoshi et al., 2009). These minichromosomes have been shown to be functional and more importantly to assemble complete kinetochore (Bloom and Joglekar, 2010; Clarke and Carbon, 1980; Ivanov and Nasmyth, 2005). In addition to the *CEN3* locus, the minichromosome contained a *TRP1* selectable marker, an origin of replication and two copies of lactose operators (*LacO*). This minichromosome or another minichromosome with a mutated *CEN3* region incapable of forming kinetochore were transformed into host cells constitutively expressing high levels of FLAG tagged LacI. LacI has been shown to bind with great affinity to LacO sequences (Lin and Riggs, 1975) enabling the minichromosome to be purified away from chromatin (Figure 1-6C). These cells were grown in light or heavy labelled SILAC media, the minichromosome purified individually from an equal amount of material, the purified protein mixed and quantitative LC-MS/MS performed. In this manner, kinetochore proteins were enriched with WT *CEN3* over mutant *CEN3* minichromosome such as Fin1, a protein phosphatase 1 regulatory subunit not previously known to associate with kinetochore (Akiyoshi et al., 2009). The use of the minichromosome purification strategy simplified the analysis of kinetochore proteins by multiplying by over 30 times the amount of *CEN3* loci present in a single cell (Akiyoshi et al., 2009). While this approach can successfully purify proteins associated with a single

genomic locus in yeast, it requires the development of an extensive array of reagents which renders its application to numerous genomic loci impractical.

A different approach was taken by Schultz-Norton et al. to detect chromatin-associated protein complexes with the oestrogen receptor alpha (ERα) in an *in vitro* assay (Schultz-Norton et al., 2008). The assay was designed to study chromatin bound protein complexes by reconstituting them *in vitro* prior to mass spectrometry identification, thus bypassing some of the technical issues previously discussed. Their assay was conducted as follows: first HeLa nuclear cell extract and recombinant ERα were prepared. Then, DNA oligos containing the Xenopus laevis A2 oestrogen responsive element sequence and its flanking elements were left at room temperature to anneal. The double stranded DNA was then incubated with the purified ERα and subsequently with HeLa cell nuclear extract. The resulting protein complexes associated with chromatin were then isolated from the unbound material using an agarose gel. The gel section corresponding to the ERα containing DNA bound complexes was excised, its protein content concentrated and analyzed by mass spectrometry. Follow-up work on some of these newly discovered proteins associated with ERα revealed that they impacted the transcriptional function of ERα. This new procedure reported by Schultz-Norton et al. represents an attractive approach to study chromatin bound protein complexes but might be difficult to standardize since knowledge of DNA binding sites as well as previously purified bait protein is required. Finally, the assay is performed on nucleosome-free DNA which does not mimic chromatin perfectly.

The use of DNA probes for the enrichment of protein complexes was also used by Rubio et al. to study the CCTC-binding factor CTCF in Jurkat cells (Rubio et al., 2008). The authors used 163 base pair DNA probes corresponding to either WT or mutant c-myc insulator sequences coupled with biotin. Insulator sequences participate in proper gene expression by preventing inappropriate enhancers' action as well as by restricting heterochromatin spreading (Komura et al., 2007). The probes were immobilized on magnetic beads coated with streptavidin enabling efficient immunopurification to be performed once the probes were mixed with nuclear extracts from Jurkat cells. By Western blotting, the authors showed a strong enrichment of CCTC-binding factor CTCF with the WT probe over the mutant one, reproducing previously shown data and validating their enrichment method. The next step was to analyze the proteins associated with their probes by mass spectrometry. To better distinguish real hits from background contamination, the ICAT method was used to quantify the proteins associated with the probes. Putative interaction partners were easily identified since they were enriched with the WT probes over the mutant c-myc insulator. This method showed a significant enrichment of the cohesion subunit Scc3/SA1 with the WT probe. Subsequent validation experiments showed that the CCTC-binding factor CTCF and Scc3/SA1 physical distribution on chromosome often overlap. This represented a novel role for CCTC-binding factor CTCF in cohesion which may have clinical implications.

This type of approach was shown by Mittler et al. to be compatible with SILAC, a quantitative proteomics method (Mittler et al., 2009). The method used by the authors is similar to the examples previously described but with a few key differences in their sample

preparation and probe design (Fig. 1-6D). The probes used in this study were designed to contain 40 base pairs encompassing a binding site for the transcription factor AP2 and its surrounding sequences, a restriction site for an endonuclease and a biotin moiety. A control probe containing a two point mutation in the AP2 binding site was also prepared to act as a negative control in the immunoprecipitation experiments. The nuclear extracts used in this paper were prepared from HeLa-S3 cells grown in suspension from media containing either light lysine (lysine-d0) or heavy $^{2}H_{4}$-lysine (lysine-d4). The light extracts (unlabeled) were used in the control immunoprecipitation and the heavy ones with the wild-type probe. Following the elution of the protein complexes by the addition of an endonuclease, the purified proteins from immunopurification were combined, run on a SDS-PAGE gel and analyzed by mass spectrometry. This protocol resulted in the identification of more than 250 proteins, a small fraction of which were specifically enriched with the wild-type probe. Among these proteins, TFAP2A was found to be enriched with the wild-type probe over the mutated probe by more than 6 fold which demonstrates the specificity of the purification. In a subsequent experiment, the authors tested whether their new method could specifically enrich proteins associated with methylated DNA. To do so, they synthesized two new probes containing a CpG island found ahead of the MAT2 loci with and without methylation. Using the experimental protocol described earlier, the authors were successful at efficiently enriching proteins associated with the methylated probe over the unmodified one.

The work described by Mittler et al. and Rubio et al. clearly demonstrates the feasibility of performing efficient proteomics studies of DNA-protein interactions (Mittler et

al., 2009; Rubio et al., 2008). The use of biotinylated DNA probes, a relatively inexpensive reagent, was able to purify proteins in sufficient quantity for mass spectrometric analysis. In both cases though, the methods suffered from significant background contaminations but the authors were still able to detect few specific interactions through the use of quantitative proteomic methods. These reports provide a basis for more elaborate work which could study single-nucleotide polymorphism and other genomic variations from a proteomic perspective. While this kind of work is bound to positively affect the chromatin field, it does pose the unusual constraint of designing experiments around a genomic locus rather than a protein. Thus to better understand the function of a given chromatin binding protein, a different protein-centered approach is needed. The development and large-scale utilization of such a technique is the focus on my thesis and will be discussed extensively in chapter 2 and 3.

1.6 Summary and Overall Significance

In the last decade, proteomics established itself as a critical technique for life science research. Above, I have described how technical progress in proteomics enabled both qualitative and quantitative protein analysis to become routine in numerous laboratories. In particular, the identification of protein-protein interaction for soluble baits is now commonly employed and as a result, thousands of protein-protein interactions have been reported in different organisms. As such subunits of abundant protein complexes such as the proteasome and spliceosome have been defined in details. Focus has thus shifted

away from genome-wide studies toward targeted "interactome mapping" such as the recent yeast kinome (Breitkreutz et al., 2010), human deubiquitinating enzymes (Sowa et al., 2009) and human autophagy system (Behrends et al., 2010).

In this thesis, I describe the development and application of an affinity purification method designed to facilitate the characterization of DNA binding baits. This method, called modified chromatin immunopurification or mChIP, enables chromatin bound protein to be effectively purified and to identify their direct and indirect (i.e. through chromatin) interaction partners ((Lambert et al., 2009) and chapter 2). In a proof-of-concept study, I applied the mChIP approach to the study of two histones and three non-histones chromatin bound protein (Lambert et al., 2009). Subsequently, I utilized the mChIP method to identify an association between the Rtt106 histone chaperone and the HIR complex (Fillingham et al., 2009). mChIP is thus applicable to all type of DNA binding proteins tested thus far.

Chapter 3 of my thesis describes a large-scale study in which over 100 chromatin associated baits were mChIP and the co-purified network of associated proteins identified by mass spectrometry. The resulting mChIP-MS dataset, when compared to present literature data, showed a significant improvement in the characterization of transcription factors and other low abundance DNA binding proteins. The mChIP procedure was further employed to dissect the physical interaction between histone chaperones involved in the regulation of the *HTA1-HTB1* promoter. In particular, the role of acetylation of lysine 56 on histone H3 with regards to the Rtt106, Asf1 and HIR was defined. The work presented in this thesis as well as other collaborations performed during my graduate studies (appendix

1-1) demonstrate the value of innovative proteomic approaches in the characterization of biological systems. In addition, the improved capacity in the identification of protein-protein interaction generated by the use of mChIP demonstrates the value of utilizing methodologies adapted to different bait classes.

1.7 Bibliography:

Aguilar, P.S., Frohlich, F., Rehman, M., Shales, M., Ulitsky, I., Olivera-Couto, A., Braberg, H., Shamir, R., Walter, P., Mann, M., *et al.* (2010). A plasma-membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. Nat Struct Mol Biol.

Akiyoshi, B., Nelson, C.R., Ranish, J.A., and Biggins, S. (2009). Quantitative proteomic analysis of purified yeast kinetochores identifies a PP1 regulatory subunit. Genes Dev *23*, 2887-2899.

Andersen, J.S., and Mann, M. (2000). Functional genomics by mass spectrometry. FEBS Lett *480*, 25-31.

Anderson, S. (1981). Shotgun DNA sequencing using cloned DNase I-generated fragments. Nucleic Acids Res *9*, 3015-3027.

Archambault, V., Li, C.X., Tackett, A.J., Wasch, R., Chait, B.T., Rout, M.P., and Cross, F.R. (2003). Genetic and biochemical evaluation of the importance of Cdc6 in regulating mitotic exit. Mol Biol Cell *14*, 4592-4604.

Aygun, O., Svejstrup, J., and Liu, Y. (2008). A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. Proc Natl Acad Sci U S A *105*, 8580-8584.

Baetz, K., Measday, V., and Andrews, B. (2006). Revealing hidden relationships among yeast genes involved in chromosome segregation using systematic synthetic lethal and synthetic dosage lethal screens. Cell Cycle *5*, 592-595.

Behrends, C., Sowa, M.E., Gygi, S.P., and Harper, J.W. (2010). Network organization of the human autophagy system. Nature *466*, 68-76.

Belle, A., Tanay, A., Bitincka, L., Shamir, R., and O'Shea, E.K. (2006). Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A *103*, 13004-13009.

Berthelet, S., Usher, J., Shulist, K., Hamza, A., Maltez, N., Johnston, A., Fong, Y., Harris, L.J., and Baetz, K. (2010). Functional Genomics Analysis of the Saccharomyces cerevisiae Iron Responsive Transcription Factor Aft1 Reveals Iron-independent Functions. Genetics.

Blackler, A.R., Klammer, A.A., MacCoss, M.J., and Wu, C.C. (2006). Quantitative comparison of proteomic data quality between a 2D and 3D quadrupole ion trap. Anal Chem *78*, 1337-1344.

Blagoev, B., Kratchmarova, I., Ong, S.E., Nielsen, M., Foster, L.J., and Mann, M. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol *21*, 315-318.

Bloom, K., and Joglekar, A. (2010). Towards building a chromosome segregation machine. Nature *463*, 446-456.

Bonenfant, D., Coulot, M., Towbin, H., Schindler, P., and van Oostrum, J. (2006). Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. Mol Cell Proteomics *5*, 541-552.

Boulard, M., Bouvet, P., Kundu, T.K., and Dimitrov, S. (2007). Histone variant nucleosomes: structure, function and implication in disease. Subcell Biochem *41*, 71-89.

Breitkreutz, A., Choi, H., Sharom, J.R., Boucher, L., Neduva, V., Larsen, B., Lin, Z.Y., Breitkreutz, B.J., Stark, C., Liu, G., *et al.* (2010). A global protein kinase and phosphatase interaction network in yeast. Science *328*, 1043-1046.

Brizzard, B.L., Chubet, R.G., and Vizard, D.L. (1994). Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. Biotechniques *16*, 730-735.

Chen, C., Jin, J., James, D.A., Adams-Cioaba, M.A., Park, J.G., Guo, Y., Tenaglia, E., Xu, C., Gish, G., Min, J., and Pawson, T. (2009). Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. Proc Natl Acad Sci U S A *106*, 20336-20341.

Chen, G.I., and Gingras, A.C. (2007). Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. Methods *42*, 298-305.

Cho, H., Orphanides, G., Sun, X., Yang, X.J., Ogryzko, V., Lees, E., Nakatani, Y., and Reinberg, D. (1998). A human RNA polymerase II complex containing factors that modify chromatin structure. Mol Cell Biol *18*, 5355-5363.

Christie, K.R., Hong, E.L., and Cherry, J.M. (2009). Functional annotations for the Saccharomyces cerevisiae genome: the knowns and the known unknowns. Trends Microbiol *17*, 286-294.

Clarke, L., and Carbon, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature *287*, 504-509.

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

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., *et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Cooper, M.P., Machwe, A., Orren, D.K., Brosh, R.M., Ramsden, D., and Bohr, V.A. (2000). Ku complex interacts with and stimulates the Werner protein. Genes Dev *14*, 907-912.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., *et al.* (2010). The genetic landscape of a cell. Science *327*, 425-431.

de Godoy, L.M., Olsen, J.V., Cox, J., Nielsen, M.L., Hubner, N.C., Frohlich, F., Walther, T.C., and Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature *455*, 1251-1254.

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Deterding, L.J., Moseley, M.A., Tomer, K.B., and Jorgenson, J.W. (1991). Nanoscale separations combined with tandem mass spectrometry. J Chromatogr *554*, 73-82.

Du, Y.C., Gu, S., Zhou, J., Wang, T., Cai, H., Macinnes, M.A., Bradbury, E.M., and Chen, X. (2006). The dynamic alterations of H2AX complex during DNA repair detected by a proteomic approach reveal the critical roles of Ca(2+)/calmodulin in the ionizing radiation-induced cell cycle arrest. Mol Cell Proteomics *5*, 1033-1044.

Eitoku, M., Sato, L., Senda, T., and Horikoshi, M. (2008). Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. Cell Mol Life Sci *65*, 414-444.

Eng, J.K., Mccormack, A.L., and Yates, J.R. (1994). An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. J Am Soc Mass Spectr *5*, 976-989.

Fazekas de St Groth, S., Webster, R.G., and Datyner, A. (1963). Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. Biochim Biophys Acta *71*, 377-391.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science *246*, 64-71.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S., *et al.* (2009). Functional organization of the S. cerevisiae phosphorylation network. Cell *136*, 952-963.

Fields, S. (2009). Interactive learning: lessons from two hybrids over two decades. Proteomics *9*, 5209-5213.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Fillingham, J., Kainth, P., Lambert, J.P., van Bakel, H., Tsui, K., Pena-Castillo, L., Nislow, C., Figeys, D., Hughes, T.R., Greenblatt, J., and Andrews, B.J. (2009). Two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. Mol Cell *35*, 340-351.

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Garcia, B.A., Shabanowitz, J., and Hunt, D.F. (2007). Characterization of histones and their post-translational modifications by mass spectrometry. Curr Opin Chem Biol *11*, 66-73.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., *et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Glatter, T., Wepf, A., Aebersold, R., and Gstaiger, M. (2009). An integrated workflow for charting the human interaction proteome: insights into the PP2A system. Mol Syst Biol *5*, 237.

Green, E.M., Antczak, A.J., Bailey, A.O., Franco, A.A., Wu, K.J., Yates, J.R., 3rd, and Kaufman, P.D. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Curr Biol *15*, 2044-2049.

Gunjan, A., Paik, J., and Verreault, A. (2005). Regulation of histone synthesis and nucleosome assembly. Biochimie *87*, 625-635.

Hansen, J.C., Nyborg, J.K., Luger, K., and Stargell, L.A. (2010). Histone chaperones, histone acetylation, and the fluidity of the chromogenome. J Cell Physiol *224*, 289-299.

Haynes, P.A., Fripp, N., and Aebersold, R. (1998). Identification of gel-separated proteins by liquid chromatography-electrospray tandem mass spectrometry: comparison of methods and their limitations. Electrophoresis *19*, 939-945.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180-183.

Howson, R., Huh, W.K., Ghaemmaghami, S., Falvo, J.V., Bower, K., Belle, A., Dephoure, N., Wykoff, D.D., Weissman, J.S., and O'Shea, E.K. (2005). Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. Comp Funct Genomics *6*, 2-16.

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

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., *et al.* (2007). Multiple high-throughput analyses monitor the response of E. coli to perturbations. Science *316*, 593-597.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc Natl Acad Sci U S A *98*, 4569-4574.

Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S., and Sakaki, Y. (2000). 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 *97*, 1143-1147.

Ivanov, D., and Nasmyth, K. (2005). A topological interaction between cohesin rings and a circular minichromosome. Cell *122*, 849-860.

Izzo, A., Kamieniarz, K., and Schneider, R. (2008). The histone H1 family: specific members, specific functions? Biol Chem *389*, 333-343.

Jensen, L.J., and Bork, P. (2008). Biochemistry. Not comparable, but complementary. Science *322*, 56-57.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., *et al.* (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol Cell *27*, 262-274.

Jiang, L., Smith, J.N., Anderson, S.L., Ma, P., Mizzen, C.A., and Kelleher, N.L. (2007). Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. J Biol Chem *282*, 27923-27934.

Kitano, H. (2002). Systems biology: a brief overview. Science *295*, 1662-1664.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95-104.

Komura, J., Ikehata, H., and Ono, T. (2007). Chromatin fine structure of the c-MYC insulator element/DNase I-hypersensitive site I is not preserved during mitosis. Proc Natl Acad Sci U S A *104*, 15741-15746.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., *et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Kroll, E.S., Hyland, K.M., Hieter, P., and Li, J.J. (1996). Establishing genetic interactions by a synthetic dosage lethality phenotype. Genetics *143*, 95-102.

Kumar, A., and Snyder, M. (2002). Protein complexes take the bait. Nature *415*, 123-124.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lai, J.S., and Herr, W. (1992). Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations. Proc Natl Acad Sci U S A *89*, 6958-6962.

Lambert, J.P., Mitchell, L., Rudner, A., Baetz, K., and Figeys, D. (2009). A novel proteomics approach for the discovery of chromatin-associated protein networks. Mol Cell Proteomics *8*, 870-882.

Lamond, A.I., and Mann, M. (1997). Cell biology and the genome projects a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. Trends Cell Biol *7*, 139-142.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. Cell *134*, 244-255.

Lin, S., and Riggs, A.D. (1975). The general affinity of lac repressor for E. coli DNA: implications for gene regulation in procaryotes and eucaryotes. Cell *4*, 107-111.

Liu, H., Sadygov, R.G., and Yates, J.R., 3rd (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem *76*, 4193-4201.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203-210.

Loyola, A., and Almouzni, G. (2007). Marking histone H3 variants: how, when and why? Trends Biochem Sci *32*, 425-433.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006). Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal Chem *78*, 2113-2120.

Maldonado, E., Shiekhattar, R., Sheldon, M., Cho, H., Drapkin, R., Rickert, P., Lees, E., Anderson, C.W., Linn, S., and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature *381*, 86-89.

Mann, M., and Wilm, M. (1994). Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem *66*, 4390-4399.

Massie, C.E., and Mills, I.G. (2008). ChIPping away at gene regulation. EMBO Rep *9*, 337-343.

Measday, V., Baetz, K., Guzzo, J., Yuen, K., Kwok, T., Sheikh, B., Ding, H., Ueta, R., Hoac, T., Cheng, B., *et al.* (2005). Systematic yeast synthetic lethal and synthetic dosage lethal screens identify genes required for chromosome segregation. Proc Natl Acad Sci U S A *102*, 13956-13961.

Mello, J.A., Sillje, H.H., Roche, D.M., Kirschner, D.B., Nigg, E.A., and Almouzni, G. (2002). Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. EMBO Rep *3*, 329-334.

Merril, C.R., Goldman, D., Sedman, S.A., and Ebert, M.H. (1981). Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. Science *211*, 1437-1438.

Mittler, G., Butter, F., and Mann, M. (2009). A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. Genome Res *19*, 284-293.

Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M.B., Roberts, T.M., Brown, G.W., Varela, E., Hediger, F., Gasser, S.M., and Krogan, N.J. (2008). Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science *322*, 597-602.

Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., Luhrmann, R., and Mann, M. (1997). Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A *94*, 385-390.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. J Biol Chem *250*, 4007-4021.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Ong, S.E., Foster, L.J., and Mann, M. (2003). Mass spectrometric-based approaches in quantitative proteomics. Methods *29*, 124-130.

Osley, M.A., Gould, J., Kim, S., Kane, M.Y., and Hereford, L. (1986). Identification of sequences in a yeast histone promoter involved in periodic transcription. Cell *45*, 537-544.

Page, M.J., Amess, B., Townsend, R.R., Parekh, R., Herath, A., Brusten, L., Zvelebil, M.J., Stein, R.C., Waterfield, M.D., Davies, S.C., and O'Hare, M.J. (1999). Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. Proc Natl Acad Sci U S A *96*, 12589-12594.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551-3567.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol *10*, 192-206.

Rabilloud, T. (2000). Detecting proteins separated by 2-D gel electrophoresis. Anal Chem *72*, 48A-55A.

Ricke, R.M., and Bielinsky, A.K. (2005). Easy detection of chromatin binding proteins by the Histone Association Assay. Biol Proced Online *7*, 60-69.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol *17*, 1030-1032.

Ross-Macdonald, P., Coelho, P.S., Roemer, T., Agarwal, S., Kumar, A., Jansen, R., Cheung, K.H., Sheehan, A., Symoniatis, D., Umansky, L., *et al.* (1999). Large-scale analysis of the yeast genome by transposon tagging and gene disruption. Nature *402*, 413-418.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol *148*, 635-651.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314.

Rusk, N. (2009). Reverse ChIP. Nature Methods *6*.

Sardiu, M.E., Cai, Y., Jin, J., Swanson, S.K., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2008). Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc Natl Acad Sci U S A *105*, 1454-1459.

Sauer, U., Heinemann, M., and Zamboni, N. (2007). Genetics. Getting closer to the whole picture. Science *316*, 550-551.

Schultz-Norton, J.R., Ziegler, Y.S., Likhite, V.S., Yates, J.R., and Nardulli, A.M. (2008). Isolation of novel coregulatory protein networks associated with DNA-bound estrogen receptor alpha. BMC Mol Biol *9*, 97.

Schwartz, J.C., Senko, M.W., and Syka, J.E. (2002). A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom *13*, 659-669.

Selbach, M., and Mann, M. (2006). Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nat Methods *3*, 981-983.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nat Protoc *2*, 1445-1457.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem *68*, 850-858.

Shiio, Y., Eisenman, R.N., Yi, E.C., Donohoe, S., Goodlett, D.R., and Aebersold, R. (2003). Quantitative proteomic analysis of chromatin-associated factors. J Am Soc Mass Spectrom *14*, 696-703.

Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31-40.

Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. Cell *58*, 15-25.

Snyder, M., and Gallagher, J.E. (2009). Systems biology from a yeast omics perspective. FEBS Lett *583*, 3895-3899.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S.G., Cyert, M., Hughes, T.R., *et al.* (2006). Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell *21*, 319-330.

Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the human deubiquitinating enzyme interaction landscape. Cell *138*, 389-403.

Steen, H., and Mann, M. (2004). The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol *5*, 699-711.

Steinberg, T.H., Haugland, R.P., and Singer, V.L. (1996). Applications of SYPRO orange and SYPRO red protein gel stains. Anal Biochem *239*, 238-245.

Stevens, S.W., and Abelson, J. (1999). Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proc Natl Acad Sci U S A *96*, 7226-7231.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Tabb, D.L., Thompson, M.R., Khalsa-Moyers, G., VerBerkmoes, N.C., and McDonald, W.H. (2005). MS2Grouper: group assessment and synthetic replacement of duplicate proteomic tandem mass spectra. J Am Soc Mass Spectrom *16*, 1250-1261.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007a). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabanowitz, J., Chait, B.T., Hunt, D.F., and Allis, C.D. (2007b). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A *104*, 2086-2091.

Thomas, C.E., Kelleher, N.L., and Mizzen, C.A. (2006). Mass spectrometric characterization of human histone H3: a bird's eye view. J Proteome Res *5*, 240-247.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H., *et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M., *et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Trinkle-Mulcahy, L., Andersen, J., Lam, Y.W., Moorhead, G., Mann, M., and Lamond, A.I. (2006). Repo-Man recruits PP1 gamma to chromatin and is essential for cell viability. J Cell Biol *172*, 679-692.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., *et al.* (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature *403*, 623-627.

van den Berg, D.L., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol Cell Biol *28*, 5986-5995.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature *417*, 399-403.

Wang, G.G., Allis, C.D., and Chi, P. (2007). Chromatin remodeling and cancer, Part II: ATP-dependent chromatin remodeling. Trends Mol Med *13*, 373-380.

Wang, H., Kakaradov, B., Collins, S.R., Karotki, L., Fiedler, D., Shales, M., Shokat, K.M., Walther, T.C., Krogan, N.J., and Koller, D. (2009). A complex-based reconstruction of the Saccharomyces cerevisiae interactome. Mol Cell Proteomics *8*, 1361-1381.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., and Williams, K.L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev *13*, 19-50.

Wilm, M., and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. Anal Chem *68*, 1-8.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature *379*, 466-469.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., *et al.* (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901-906.

Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010). Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. Science *328*, 94-98.

Yates, J.R., 3rd, McCormack, A.L., Schieltz, D., Carmack, E., and Link, A. (1997). Direct analysis of protein mixtures by tandem mass spectrometry. J Protein Chem *16*, 495-497.

Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., *et al.* (2008). High-quality binary protein interaction map of the yeast interactome network. Science *322*, 104-110.

Chapter 2

A Novel Proteomics Approach for the Discovery of Chromatin-associated Protein Networks

The work presented in this chapter is published as:

Jean-Philippe Lambert, Leslie Mitchell, Adam Rudner, Kristin Baetz and Daniel Figeys. *A Novel Proteomics Approach for the Discovery of Chromatin-associated Protein Networks.* Molecular & Cellular Proteomics, 8 (4), 2009, 870-882.

Permission to reprint this work was obtained from *The American Society for Biochemistry and Molecular Biology*.

Authors contributions:

JPL developed the method, performed the purifications and MS analyses, analyzed the data, wrote the manuscript and produced figure 2-1 to 2-4 and 2-6.

LM generated figure 2-5 from JPL data and edited the manuscript.

AR participated in the method development.

KB edited the manuscript and provided reagents.

DF directed the project and writing of the manuscript.

Abstract:

Protein-protein interaction mapping has progressed rapidly in recent years, enabling the completion of several high throughput studies. However, knowledge of physical interactions is limited for numerous classes of proteins, such as chromatin-bound proteins, because of their poor solubility when bound to DNA. To address this problem, I have developed a novel method, termed modified chromatin immunopurification (mChIP), which allows for the efficient purification of protein-DNA macromolecules, enabling subsequent protein identification by mass spectrometry. mChIP consists of a single affinity purification step whereby chromatin-bound protein networks are isolated from mildly sonicated and gently clarified cellular extracts using magnetic beads coated with antibodies. I applied the mChIP method in Saccharomyces cerevisiae cells expressing endogenously tandem affinity purification (TAP)-tagged histone H2A or the histone variant Htz1 and successfully co-purified numerous chromatin-bound protein networks as well as DNA. I further challenged the mChIP procedure by purifying three chromatin-bound bait proteins that have proven difficult to purify by traditional methods: Lge1, Mcm5, and Yta7. The protein interaction networks of these three baits dramatically expanded our knowledge of their chromatin environments and illustrate that the innovative mChIP procedure enables an improved characterization of chromatin-associated proteins.

2.1 Introduction

Chromatin, the complex packaging of DNA with proteins, is in many ways the master control center for the cell. At the most fundamental level, DNA is wrapped around histone proteins in a structure referred to as the nucleosome (Van Holde et al., 1974). Histones H2A, H2B, H3, and H4 form the core of the nucleosome around which the DNA is wrapped (Luger et al., 1997). In addition to core histones variant forms, such as histone Htz1 in yeast or H2A.Z in mammalian cells, also contributes to the diverse range of biological processes regulated by chromatin (Boulard et al., 2007). Numerous post-translational modifications of histone proteins modulate DNA-protein interactions as well as regulate the intricate and temporal associations of other proteins and protein complexes, such as chromatin-remodelling or -modifying proteins and transcription factors, with chromatin (for a review, see Ref. (Strahl and Allis, 2000)). To truly understand the regulatory role that chromatin plays in the cell and disease states will require a detailed understanding of the intricate protein-protein interactions that occur on it.

The study of protein-protein interactions has grown by leaps and bounds in recent years. Members of the Figeys laboratory and others have successfully performed large scale studies of protein-protein interactions in both yeast and human cells by immunopurification and high throughput mass spectrometry (Brajenovic et al., 2004; Ewing et al., 2007; Gavin et al., 2006; Gavin et al., 2002; Krogan et al., 2006). These studies, along with numerous small scale studies, have provided the scientific community with a wealth of information regarding the elaborate protein networks that function as the building blocks

of life. Although technical developments resulting in improved throughput and sensitivity have been achieved (for reviews, see references (Ethier et al., 2006) and (Parrish et al., 2006)), hurdles remain in the discovery of protein-protein interactions of chromatin- bound proteins/complexes. Proteins associated with DNA have the dual natures of participating in conventional soluble protein-protein interactions as well as participating in larger DNA-protein macrocomplexes. In conventional immunopurification protocols, these macrocomplexes are often lost during the clarification steps as they pellet along with the DNA/chromatin and hence are not available for the downstream immunopurification step. Consequently protein interaction networks have been discerned for the soluble fraction of chromatin-binding proteins; however, these networks may not be representative of the chromatin-bound protein environment. A holistic view of cellular protein-protein interactions occurring on chromatin requires the development of new purification techniques.

Over the years, a number of proteomics methodologies have been developed to study chromatin-bound proteins. Acid extraction, a well established protocol for the study of histones, utilizes the low solubility of the histone proteins that are complexed with DNA to specifically precipitate them from the cellular matrix; histones are subsequently resolubilized for analysis (Shechter et al., 2007). This protocol has enabled the detailed study of histone composition and their associated post-translational modifications in a variety of organisms (Galasinski et al., 2003). However, the denaturing conditions inherent to this method are not amenable to the identification of protein-protein interactions. Recently a number of methods have been reported that facilitate the identification of

interaction partners of proteins bound to chromatin (for a review, see reference (Smith et al., 2007)). A common approach in these methods is to drastically reduce, or entirely remove, genomic DNA in the cellular extract by various mechanical or enzymatic means, thereby enhancing the solubility of the protein of interest. Such procedures have provided insight into the interaction partners of a number of histone variants and their interaction partners in both mammalian (Du et al., 2006; Foltz et al., 2006) and yeast cells (Mizuguchi et al., 2004). However, these methods drastically reduce the number of associated proteins purified with the bait of interests and hence provide an incomplete picture of the physical interactions for any given bait.

Here I report a novel method, termed modified chromatin immunopurification (mChIP), for the purification and analysis of chromatin-bound proteins and their associated protein interaction networks. This method is reminiscent of the chromatin immunoprecipitation procedures used to define DNA binding sites for transcription factors and other DNA-binding proteins (for a review, see reference (Collas and Dahl, 2008)). The mChIP protocol involves DNA shearing by sonication, gentle clarification, and affinity purification of protein-DNA complexes. This method allows the purification of chromatin-bound proteins and their associated DNA in quantities sufficient for analysis by mass spectrometry. To optimize the mChIP protocol I initially focused on the identification of associated proteins of histone H2A (Hta2) and its variant (Htz1) in the budding yeast Saccharomyces cerevisiae. The resulting Hta2 and Htz1 mChIP protein networks include an array of chromatin-bound proteins and serve to define the "chromatin background" of the new mChIP procedure. I subsequently used the new mChIP procedure to explore the

cellular function of Lge1, Mcm5, and Yta7, three known chromatin-binding proteins with differential cellular roles. Although traditional tandem affinity purification (TAP) with these baits revealed limited insight into these proteins, mChIP identified novel protein interactions providing unique biological insights into the function of these proteins and illustrating the potential worth of our procedure.

2.2 Experimental Procedures:

2.2.1 Yeast Strains, Plasmids, and Genetics Methods

All yeast strains used in this study are listed in Table 2-1. Growth media and strains were prepared following standard practices. Strains from the TAP collection were obtained from Open Biosystems (Huntsville, AL). Genomic deletions and epitope tag integrations made for this study were designed with PCR-amplified cassettes as described previously (Longtine et al., 1998; Puig et al., 2001) and confirmed by either PCR analysis or immunoblotting for tag expression.

Table 2-1: List of strains used in chapter 2.

Strain	Genotype*	Source or Reference
BY4741	MATa *ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1*	(Sikorski and Hieter, 1989)
YSC1178-7499158	BY4741 H2A.2-TAP::HIS	TAP collection
YSC1178-7502629	BY4741 HTZ1-TAP::HIS	TAP collection
YSC1178-7503047	BY4741 LGE1-TAP::HIS	TAP collection
YSC1178-7501832	BY4741 MCM5-TAP::HIS	TAP collection
YSC1178-7500768	BY4741 YTA7-TAP::HIS	TAP collection
YKB509	BY4741 *swr1Δ::kanMX*	This study
YKB963	BY4741 HTZ1-TAP::HIS *swr1Δ::KanMX*	This study

* An introduction to *Saccharomyces cerevisiae* nomenclature can be obtained from the *Saccharomyces* genome database at http://www.yeastgenome.org/help/yeastGeneNomenclature.shtml.

2.2.2 Modified Chromatin Immunopurification

One-step affinity immunopurification was performed using TAP-tagged proteins and M-270 epoxy Dynabeads (Invitrogen) coated with rabbit IgG according to the manufacturer's instructions. Briefly 700 ml of cultured yeast cells grown in yeast, peptone, and dextrose (YPD) medium to an A_{600} of ~1 were pelleted and washed with water. Cells were resuspended in lysis buffer (100 mM HEPES, pH 8.0, 20 mM magnesium acetate, 10% glycerol (v/v), 10 mM EGTA, 0.1 mM EDTA + fresh yeast protease inhibitors mixture (Sigma-Aldrich)), frozen in liquid nitrogen in small droplets, and lysed using a coffee grinder half-filled with dry ice for 1 min. The dry ice from the ground cells was allowed to evaporate, and the resulting whole cell extract was sonicated three times for 30 seconds with at least 1 minute on ice between each pulse. Nonidet P-40 was added to a final concentration of 0.4%, and the sample was mixed by hand for 30 seconds. The extract was gently clarified by centrifugation at 1800 relative centrifugal force for 10 minutes (4°C), and the supernatant was transferred to a fresh tube. Freshly prepared rabbit IgG (Sigma-Aldrich) coated Dynabeads were added (200µl/sample), and the samples were incubated with end-over-end rotation for 3 hours at 4°C. Using a Dynal MPC-S magnet (Invitrogen) the beads were collected on the side of the sample tubes, and the supernatant was discarded. The beads were washed three times by resuspension in 1 ml of wash buffer (100 mM HEPES, pH 7.4, 20 mM magnesium acetate, 10% glycerol (v/v), 10 mM EGTA, 0.1 mM EDTA, 0.5% Nonidet P-40) and subsequently transferred to a fresh tube following each wash. Finally the beads were resuspended in 1 ml of elution buffer (0.5 M NH_4OH, 0.5mM EDTA) and incubated

with end-over-end rotation for 20 min at room temperature. The protein eluates were transferred to a fresh tube and were evaporated to dryness using a SpeedVac.

The protein sample was resuspended in 1X loading buffer (50 mM Tris-HCl, pH 8, 2% SDS, 100mM DTT, 10% glycerol) and resolved on a NuPAGE 4–12% SDS-PAGE gel. For protein visualization, the gels were silver-stained. For Western blot analysis, the proteins were transferred onto nitrocellulose membrane, blocked in 5% nonfat milk in TBST (20mM Tris base, 150mM NaCl, 0.1% Tween 20), and probed with polyclonal rabbit anti-TAP antibodies (Open Biosystems) or with polyclonal rabbit anti-histone H3 antibodies (Abcam Inc.). When stated, standard immunopurification was performed without the use of sonication and with clarification of the yeast extracts at 14,000 rpm for 15 min (4°C) while maintaining the same buffer and bead conditions as for the modified ChIP method outlined above.

2.2.3 Micrococcal Nuclease S7 and DNase I Digestion prior to mChIP

Following cell lysis and sonication, $MgCl_2$ and $CaCl_2$ were adjusted to a final concentration of 10 and 3 mM, respectively. 500 units/ml micrococcal nuclease S7 (BioShop Canada Inc.) or 15µg/ml DNase I (Sigma) were added, and samples were incubated for 20 min at room temperature. The reminder of the sample preparation was performed as described above.

2.2.4 PCR Analysis of Immunopurified DNA by mChIP

Aliquots of the final eluates (DNA sample) or genomic DNA prepared by phenol/chloroform extraction of cell extracts (input) were resuspended in 100µl of 10mM Tris, EDTA, pH 8.0 and treated with 40µg of proteinase K for 2hours at 37°C. DNA fragments were then purified using the GFX PCR DNA and gel band purification kit (GE Healthcare) according to the supplier's protocols. DNA was eluted with 50 µl of water, and 1 µl was used for each reaction. PCR was performed in 25-µl batches using the iQ SYBR Green Supermix (Bio-Rad) following the supplier's instructions. Immunoprecipitated DNA was amplified by PCR with the following primers pair: GAL1 F, 5'-GAA-GAGTCTCTCGCCAATAAGAAACAGG; GAL1 R, 5'-GAACATTCGTAAAGTTTATCGCAAG (Krogan et al., 2003). PCR products were resolved on a 1.5% agarose gel and stained with ethidium bromide.

2.2.5 Mass Spectrometry Analysis

Gel bands were excised, reduced, alkylated, and digested as described previously (Abu-Farha et al., 2008) based on the original protocol of Wilm et al. (Wilm et al., 1996). Peptide solutions were dried in a SpeedVac and stored at -20°C until mass spectrometric analysis. LC-MS/MS was performed by dissolving the peptide samples in 5% formic acid and loading them into a 200-µm x 5-cm precolumn packed in house with 5-µm YMC ODS-A C_{18} beads (Waters) using a micro Agilent 1100 HPLC system (Agilent Technologies). The peptides were desalted on line with 95% water, 5% acetonitrile, 0.1% formic acid (v/v) for

10 min at 10 µl/min. Then the flow rate was split before the precolumn to produce a flow of rate of ~200 nl/min at the column. Following their elution from the precolumn, the peptides were directed to a 75-µm x 5-cm analytical column packed with 5-µm YMC ODS-A C_{18} beads. The peptides were eluted using a 1-hour gradient (5–80% acetonitrile with 0.1% formic acid) into an LTQ linear ion trap mass spectrometer (Thermo-Electron). MS/MS spectra were acquired in a data-dependant acquisition mode that automatically selected and fragmented the five most intense peaks from each MS spectrum generated.

Peak lists were generated from the MS/MS .raw file using Mascot Distiller 2.0.0.0 (Matrix Science) to produce a .mgf file with default parameters with the exception that for each MS/MS individual peak lists were generated assuming a 2+ and a 3+ charge. All .mgf files from one sample were merged into a single file and then analyzed and matched to the 6298 *S. cerevisiae* protein sequences in the National Center for Biotechnology Information (NCBI) database (released April 2007) using the Mascot 2.1.04 database search engine (Matrix Science) with trypsin as digestion enzyme, carbamidomethylation of cysteine as a fixed modification, and methionine oxidation as a variable modification. Peptide and MS/MS mass tolerances were set at ±2 and ±0.8 Da, respectively, with one miscleavage allowed and the significance threshold set to 0.01 ($p < 0.01$). Finally an ion score cutoff of 30 was chosen to produce a false-positive rate of less than 1% in the MS data (Elias et al., 2005). A protein hit required at least two "bold red peptides," i.e. the most logical assignment of the peptide in the database selected, to be reported. Furthermore when peptides matched to more than one database entry, only the highest scoring protein was considered.

2.3 Results

2.3.1 mChIP Facilitates the Purification of Protein Networks Bound to DNA

The study of proteins associated with chromatin is difficult. Well established protocols, such as TAP (Puig et al., 2001; Rigaut et al., 1999), failed to identify proteins shown previously to associate with the *S. cerevisiae* histone Hta2 and its variant, Htz1 (data not shown and reference (Gavin et al., 2006) and (Krogan et al., 2006)). In my hands, I found that most of the proteins expected to be associated with chromatin were lost during the sample preparation, particularly in the centrifugation steps. I therefore postulated that the poor solubility of large protein-DNA complexes, further aggravated by the necessary centrifugation steps, was the cause of this problem. To address this issue, I have developed a novel method based on the principle of ChIP, an established method for the purification of DNA associated with specific proteins (for a review, see reference (Collas and Dahl, 2008)). The mChIP method ensures that chromatin-bound protein complexes remain in solution by shearing DNA using sonication prior to the immunopurification (Figure 2-1). This step breaks DNA into fragments that are sufficiently small (~2000 bp) to remain soluble following mild centrifugation. Not surprisingly, I found that by limiting the clarification of the samples before performing the IP more contaminant proteins were found bound to the IgG-agarose beads (data not shown). I resolved this problem by using magnetic beads that dramatically reduce the binding of nonspecific proteins to the affinity matrix (see "2.2 Methods" for details).

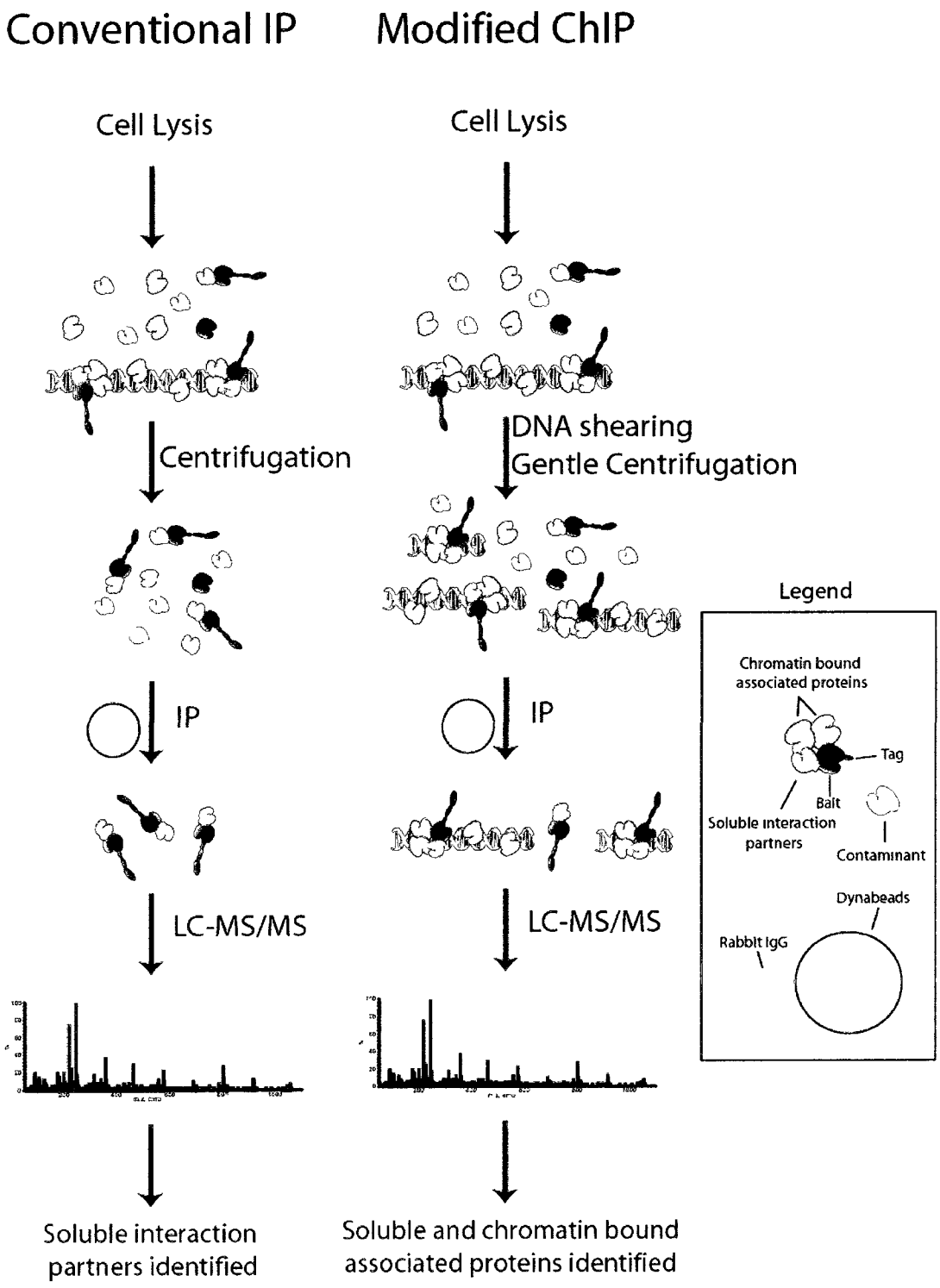


Figure 2-1: Experimental protocol for immunopurification of chromatin-bound protein complexes using the mChIP approach as compared with conventional IP methods.

The mChIP method was used to identify chromatin-bound protein networks from yeast cells expressing the endogenously TAP-tagged histone Hta2 (Hta2-TAP) or Htz1 (Htz1-TAP) (Fig. 2-2A). To ensure that the novel mChIP protocol was successfully purifying genuine protein-DNA complexes and not experimental artifacts, I also performed mChIP from cells expressing Htz1-TAP mChIP in an *swr1Δ* mutant background. Swr1 is a component of the SWR1 complex that is required for the incorporation of Htz1 into nucleosomes (Kobor et al., 2004; Krogan et al., 2004; Mizuguchi et al., 2004), and *swr1Δ* cells display a dramatic reduction of Htz1 incorporation in chromatin (Kobor et al., 2004; Krogan et al., 2004). In addition, a traditional immunopurification method was used to purify Htz1-TAP (Figure 2-2A, IP; see "2.2 Methods" for details). Although both protocols resulted in the purification of roughly equal amounts of Hta2-TAP or Htz1-TAP as assessed by Western blot (Figure 2-2B), the mChIP of both Hta2-TAP and Htz1-TAP resulted in the purification of an intricate mixture of proteins that were not present in purifications from untagged control cells or Htz1- TAP *swr1Δ* cells as wells as a conventional Htz1-TAP IP (Figure 2-2A; see Appendix 2-1 for a complete list of identified proteins). Because of the use of gradient gels to resolve the mChIP proteins, small proteins, including core histones, are poorly silver-stained but observed in large amounts by Coomassie (Figure 2-2A), Western blot (Figure 2-2C), and mass spectrometry analysis (Appendix 2-1). Histone H3 was only co-purified in the Hta2-TAP and Htz1-TAP using mChIP suggesting that this method is purifying intact nucleosomes and possibly chromatin (Figure 2-2C). Hta2 and Htz1 have been shown to localize to the *GAL1* promoter (Krogan et al., 2003), and I demonstrated that the *GAL1* promoter region can be amplified from our Hta2-TAP and Htz1-TAP mChIPs eluates but not

from the Htz1-TAP *swr1Δ* mChIP or Htz1-TAP traditional IP (Figure 2-2D). These results demonstrate that the mChIP protocol can purify chromatin from sonicated whole cell extract (Figure 2-2E).

Mass spectrometry analysis also indicates that the mChIP method can successfully identify chromatin-bound protein networks (see Appendix 2-1 for a complete list of identified proteins). Hta2-TAP and Htz1-TAP mChIPs along with the untagged control mChIP purification were performed in triplicate. Ribosomal proteins and other common contaminant proteins as well as proteins observed only in a single mChIP experiment were removed from the final high confidence data sets (Appendix 2-2). The high confidence data set of Hta2-TAP and Htz1-TAP includes several well characterized chromatin proteins such as the core histones, the topoisomerase Top2 (Berger, 1998), RNA polymerase subunits, and replication factors A and C (Cullmann et al., 1995; Longhese et al., 1994) among others. Numerous chromatin-remodelling complexes also co-purified with both Hta2-TAP and Htz1-TAP including the imitation-switch (ISWI) chromatin-remodelling complex (Mellor and Morillon, 2004). In contrast, Swr1 and Rvb1, components exclusively found in the SWR1 complex (Kobor et al., 2004; Krogan et al., 2004; Mizuguchi et al., 2004), only co-purified with Htz1-TAP, whereas Rvb2, as SWR1 complex component that also functions in additional chromatin-associated complexes, co-purified also with Hta2-TAP. Htz1-TAP also co-purified the SWR1 complex proteins Arp4, Arp6, Swc4, Swc6, and Vps71 in one of three replicates (Appendix 2-1). As expected the absence of Swr1 caused a marked decrease in the number of proteins co-purified by Htz1p-TAP; in fact, the protein interaction profile greatly resembled that of Htz1-TAP generated using traditional IP methods (Fig. 2-2A). In

both cases, the Htz1 interaction map included Nap1 and Kap114, two proteins shown previously to interact with Htz1 in the soluble (cytosolic) fraction of the cell (Mosammaparast et al., 2002). The remarkable differences in associated protein profiles between the mChIP and traditional IP support the use of the mChIP procedure in generating protein interaction networks of chromatin-bound proteins. This can be further demonstrated by comparing the 98 protein interactions observed for Hta2-TAP by mChIP against the 42 protein interactions for Hta2-TAP IP-MS purifications curated in the BioGRID (Stark et al., 2006) (Fig. 2-2F). I observed an overlap of 13 protein interactions between traditional IP and mChIP. Furthermore I found that the mChIP-derived protein interactions for Hta2-TAP have a greater proportion of nuclear localization than those found in the BioGRID only, indicating an improved accessibility to chromatin protein networks in mChIP purifications compared with classical methodologies (Fig. 2-2F).

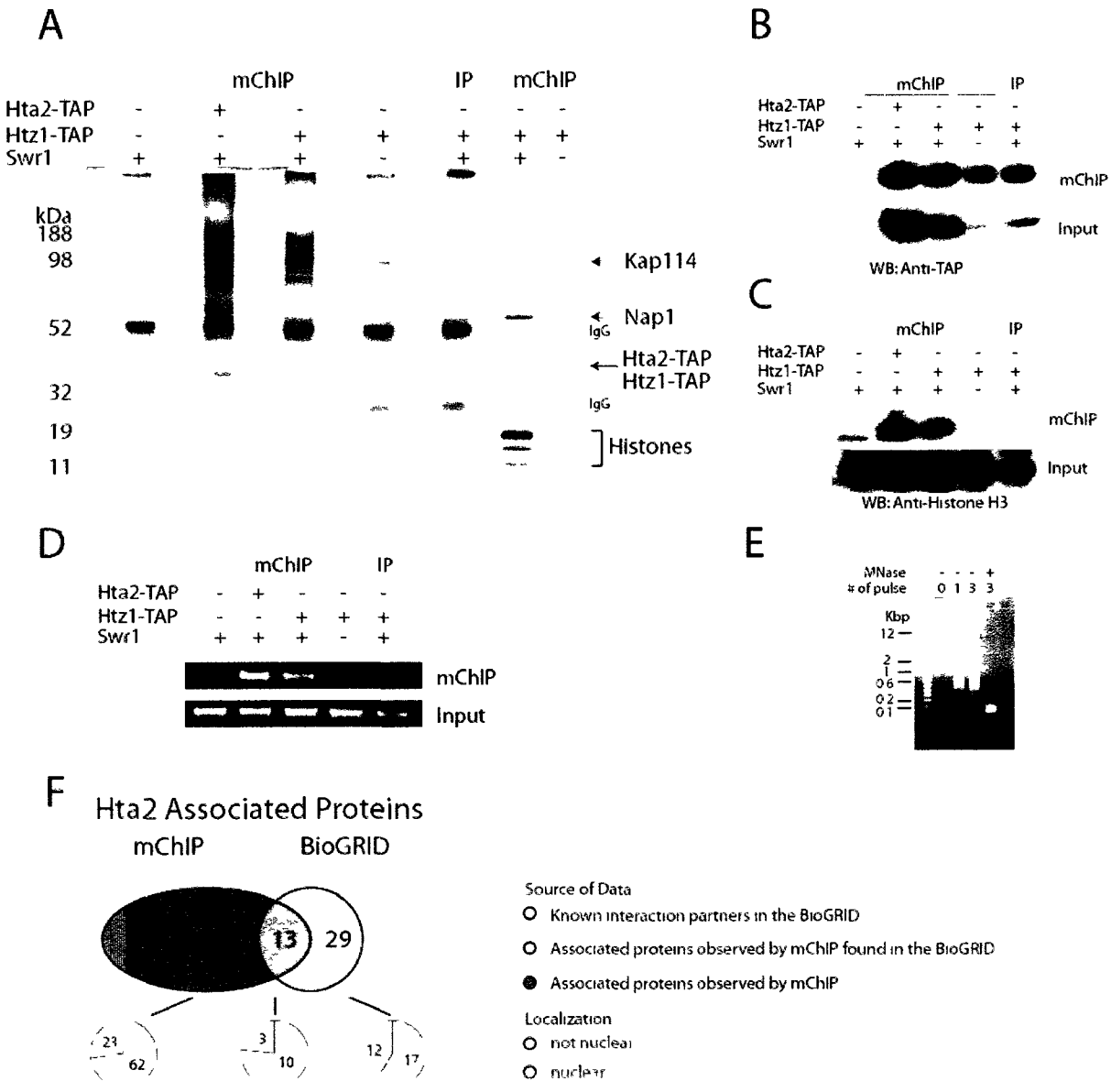


Figure 2-2: Purification of chromatin-bound protein associated with Hta2 and Htz1 by mChIP. (A) silver-stained SDS-PAGE gel of protein network associated with H2A-TAP (YSC1178-7499158), Htz1-TAP (YSC1178-7502629), and Htz1-TAP *swr1Δ* (YKB963) purified by mChIP compared with an untagged control (YPH499) and with a traditional Htz1-TAP (YSC1178-7502629) purification (IP). Additional mChIP from Htz1-TAP and Htz1-TAP *swr1Δ* strains were stained with coomassie blue to better visualize the histone proteins. An aliquot of the protein samples purified was also transferred on nitrocellulose membrane for Western blot (WB) analysis against TAP-tagged **(B)** or histone H3 proteins **(C)**. **(D)**, PCR analysis for GAL1 locus (+1039 to +1331) following mChIP showing the presence of chromatin co-purifying with the histones H2A and Htz1p. **(E)**, DNA isolated before and after DNA shearing by sonication and micrococcal nuclease S7 (MNase) treatment resolved on a 1.5% agarose gel stained with ethidium bromide. The sample prepared for mChIP usually contains chromatin fragments of 500–1500 base pairs. Results shown are representative of three experiments. **(F)**, comparison of histone H2A-associated proteins observed by mChIP and those previously reported in IP-MS experiments with H2A as a bait (literature data from BioGRID) and of their nuclear localization (Huh et al., 2003).

2.3.2 mChIP Protein Networks Are Affected by DNA Size

I was interested in determining how the size of DNA fragments in the cellular extracts affects the proteins co-purified by mChIP. As the mChIP method purifies large chromatin pieces (500-2000 base pairs; Figure 2-2E), it was expected that some co-purified proteins would not interact directly with the bait protein, but rather the interactions would be mediated through DNA. This was tested by adjusting the size of DNA in the cellular extracts with one of two enzymes: micrococcal nuclease S7 (MNase), which cleaves DNA between nucleosomes, or DNase I, which digests most DNA in solution given enough time (Du et al., 2006). This makes for a system where the DNA in an extract can be large (sonication), medium (MNase), or small/absent (DNase I) (Figure 2-3B). Htz1-TAP and untagged control mChIPs were performed from extracts treated by sonication, MNase, or DNase I, and the purified proteins were resolved on a 4-12% NuPAGE gel and silver-stained (Figure 2-3A). I found that as the DNA size was reduced in the cellular extract (Figure 2-3B) the amount of co-purifying proteins was concomitantly reduced as assessed by silver stain (Figure 2-3A). Notably DNA fragment size did not significantly affect the purification efficiency of Htz1-TAP and the associated core histones, in particular histone H3 (Figure 2-3C and D), reflecting the fact that the mChIP method can efficiently purify large protein-DNA macromolecules. These data support the view that the mChIP method isolates both proteins that directly or indirectly interact with the bait and proteins that are not necessarily physically interacting with the bait but are bound to the DNA/chromatin that is co-purified by the bait. A specific example of this phenomenon is the association of Top2 and Nap1 with Htz1-TAP by mChIP. As seen in Figure 2-3A, the histone chaperone Nap1 is

strongly associated with Htz1-TAP independently of the DNA shearing method used reflecting their strong direct interaction. On the other hand, Top2, a DNA topoisomerase, is dependent on the presence of DNA for its associations with Htz1-TAP and is lost upon DNase I treatment. Thus manipulating the size of the DNA fragments can be used as a tool to further dissect the protein networks associated with a bait of interest. But for the purpose of this work, all the subsequent experiments in this chapter were done using sonication as the DNA shearing method because it provided us with the largest number of associated proteins upon mChIP.

The mChIP procedure functions efficiently for histone baits (Figure 2-2) because these proteins are tightly associated with DNA. The question of whether other classes of chromatin-associated proteins would be as easily amenable to mChIP purifications remained open. To answer this, three non-histone baits that are difficult to purify by the traditional IP method were selected for mChIP purification. The baits chosen were Lge1, Mcm5, and Yta7, and each was purified in triplicate following the mChIP protocol developed previously.

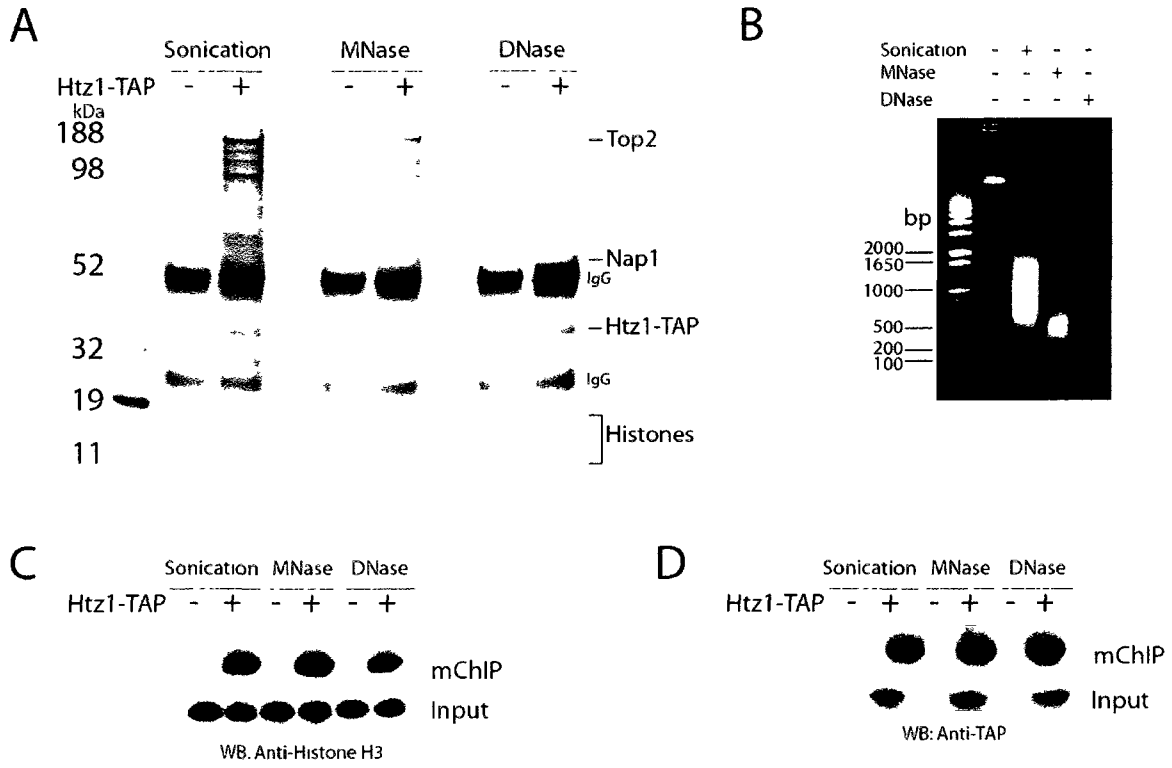


Figure 2-3: The size of chromatin fragments in whole cell extract determines the magnitude of the protein networks co-purifying with Htz1-TAP by mChIP. The mChIP procedure was performed from the Htz1-TAP strain (YSC1178-7502629) or untagged control strain (YPH499) previously treated with sonication, MNase or DNase I (see "2.2 Methods" for details). **(A)** 90% of the protein eluted for each sample was resolved on a NuPAGE 4–12% SDS-PAGE gel and silver-stained. **(B)** phenol/chloroform extractions from the whole cell extracts used in *A* were resolved on a 1% agarose gel, and the chromatin fragments were visualized by ethidium bromide to define their size. 2.5% of each input and mChIP was analyzed by WB using antibodies against histone H3 **(C)** or the TAP tag **(D)**.

2.3.3 Efficient Purification of the Poorly Characterized Lge1 and of Its Associated Proteins by mChIP

Lge1, which stands for large cells 1, is a protein of unknown function whose deletion results in the formation of abnormally large cells, a rare phenotype in S. cerevisiae (Jorgensen et al., 2002). Using traditional TAP and MS techniques, Lge1 was observed previously to interact with only two proteins: Bre1 (Collins et al., 2007; Ho et al., 2002; Krogan et al., 2006), an E3 ubiquitin ligase required for monoubiquitination of histone H2B (Hwang et al., 2003), and Isw1 (Krogan et al., 2006), a component of the chromatin-remodelling complex Isw. Interestingly histone H2B monoubiquitination at lysine 123 is required for the methylation of histone H3 at lysine 4 (Dover et al., 2002; Wood et al., 2003). The exact role of Lge1 in this "cross-talk" is still undefined (Lee et al., 2007; Wood et al., 2005). Lge1 was also shown to participate in the induction of *PDR3* transcript independently of its association with Bre1p and of its role in histone H2B ubiquitination (Zhang et al., 2005). *lge1Δ* cells display genetic interactions with more than 369 genes indicating that its function impacts a myriad of cellular functions. These puzzling aspects of Lge1 biology and its apparent association with chromatin make it a prime candidate for mChIP analysis in hopes of further refining its cellular functions. Although traditional methods only identified Isw1 (Krogan et al., 2006), mChIP of Lge1-TAP copurified an unexpectedly large number of associated proteins (Fig. 2-4A). I reproducibly identified 52 proteins co-purifying with Lge1-TAP in two of three purifications, 40 of which were unique to Lge1-TAP and not identified in the Hta2-TAP mChIPs that serve as a chromatin background control (Fig. 2-4B and see Appendix 2-1 and 2-2). Bre1 was observed as the top

hit in each of the Lge1 mChIP experiments, confirming that Lge1 and Bre1 strongly interact. In addition, a plethora of new associated proteins were found. Surprisingly Kog1, for kontroller of growth, was the second MS hit in two of the three Lge1 mChIP experiments and thus strongly associated with Lge1. Kog1 is a member of the TORC1 complex, a protein complex containing the Tor1 kinase, and is involved in control of cell growth in response to nutrient level (Barbet et al., 1996). Two more subunits of the TORC1 complex (Tco89 and Tor1) were also co-purified reproducibly with Lge1 by mChIP, and the last member of TORC1, Lst8, was identified in one experiment. The association between Lge1 and the TORC1 complex raises the possibility that Lge1 could affect cell size via interplay with the action of the TORC1 complex. Furthermore a significant amount of Ssd1 was co-purified in the three Lge1 mChIPs, solidifying the connection between Lge1 and the TORC1 complex since previous report linked TORC1 to Ssd1 (Reinke et al., 2004). Another subset of Lge1-associated proteins (Msh2, Msh3, Msh6, and Rad52) was found to be involved in DNA repair. These associations are consistent with previous reports showing the sensitivity of *lge1Δ* cells to UV radiation (Hanway et al., 2002). Finally spindle pole body (SPB) components make up a large fraction of proteins co-purifying with Lge1 by mChIP. This includes core SPB components such as Spc29, Spc110, Spc97, Spc98p, Spc42, Cdc31, Nud1, and Cnm67 but also associated kinases Kin3 and Kin4 and the mitotic exit network protein Bub2 (for a review see Reference (Jaspersen and Winey, 2004)). Both the number of SPB components and their high scores in our MS analysis suggest that a significant fraction of the cellular Lge1 may be present at the SPB and may contribute to spindle dynamics.

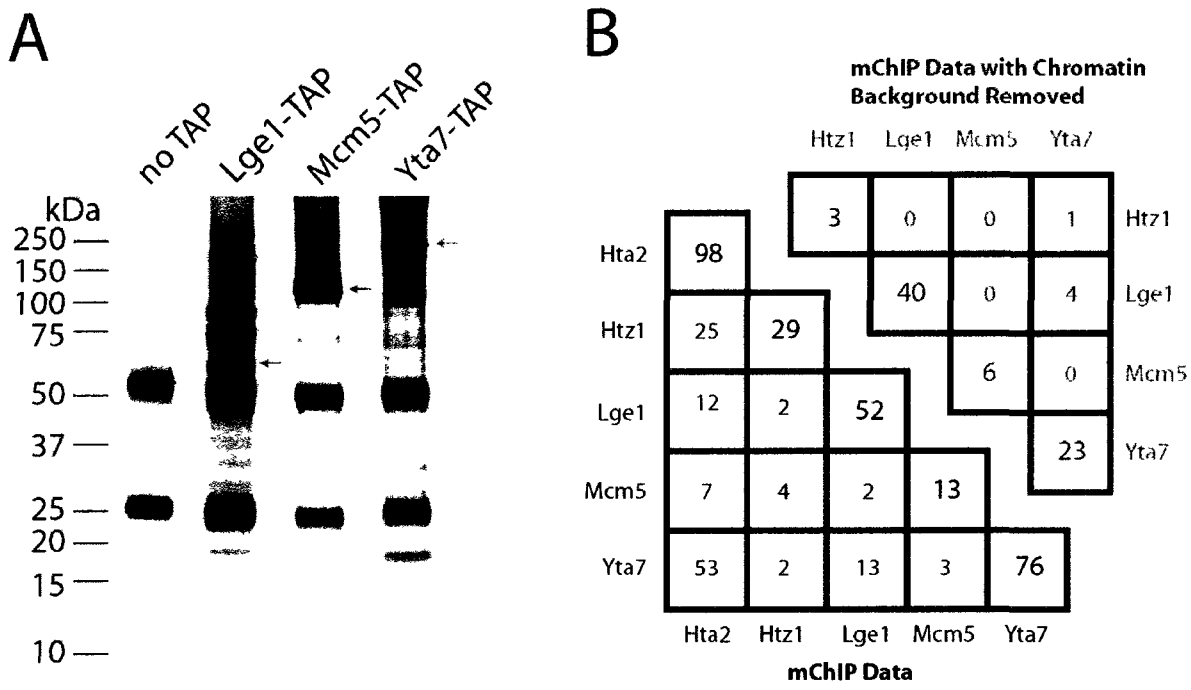


Figure 2-4: mChIP efficiently purifies chromatin-associated protein networks when non-histone proteins are used as baits. (A) Silver-stained 4–12% NuPAGE gels of the protein networks associated with Lge1-TAP (YSC1178-7503047), Mcm5-TAP (YSC1178-7501832), and Yta7-TAP (YSC1178-7500768). mChIP purifications performed according to the "Materials and Methods" are shown as compared with an untagged control (YPH499). The red arrows indicate the TAP-tagged bait protein. Results shown are representative of three experiments. **(B)** Summary of the overlap between the associated proteins observed in the high confidence data set (supplemental Table 2) and each mChIP bait (bottom left triangle) and with the chromatin background or Hta2-TAP high confidence data set removed (top right triangle).

2.3.4 Purification of the Minichromosome Maintenance (MCM) Helicase and of Its Associated Interaction Partners by mChIP

The MCM complex is composed of six highly conserved components termed Mcm2 through Mcm7 (for a review, see Reference (Forsburg, 2008)). The MCM complex participates in DNA replication and elongation, and each subunit is essential for cell survival in *S. cerevisiae*. To effectively function, the MCM complex associates with a number of distinct proteins such as the DNA replication initiation factor Cdc45 as well as other chromatin-associated proteins such as Ctf4 at specific cell cycle stages and chromatin locations (Aparicio et al., 1997; Gambus et al., 2006; Sheu and Stillman, 2006). To date, successful affinity purification of the MCM complex has required a large amount of cells (up to 10 liters) (Gambus et al., 2006), and high throughput studies failed to purify the highly abundant MCM complex (Gavin et al., 2006; Ho et al., 2002; Krogan et al., 2006). I hypothesized that the tight interactions of the MCM complex with chromatin might be the cause of the low efficiency in previous immunopurifications and therefore tested the mChIP method with a component of the complex, Mcm5-TAP. The three Mcm5 mChIP experiments were successful at reproducibly co-purifying 13 proteins, six of which were only identified in the Mcm5-TAP (Figure 2-4B). High levels of the six core MCM components (Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, and Mcm7) were found in the three Mcm5-TAP mChIP experiments despite the relatively small quantity of cells used for the purifications (700 mL cultures). Mcm3 and Mcm6 were identified in only the Mcm5-TAP, whereas Mcm2, Mcm4, Mcm5, and Mcm7 were identified in both Mcm5-TAP and Hta2-TAP mChIPs. Furthermore additional non-core MCM complex components were also co-purified by

mChIP, Tof1 and Ctf4. Sap185 was identified as a novel interaction partner for Mcm5. Sap185 is a phosphatase shown previously to interact with Sit4, another phosphatase, and is involved in the G1 to S cell cycle progression (Luke et al., 1996). The co-purification of Sap185 with Mcm5-TAP by mChIP raises the possibility that Sap185 may interact with one or more MCM core subunits and dephosphorylate them in a cell cycle-dependent manner.

2.3.5 Determination of Yta7 Interaction Partners on Chromatin by mChIP

One protein that was identified in both Hta2-TAP and Htz1-TAP mChIP experiments was Yta7. Yta7 contains three domains: two AAA ATPase domains of unknown function and one bromodomain, a domain known to bind acetylated histones (Mujtaba et al., 2007). Yta7 had been observed previously only at chromatin boundary regions between transcriptionally active and silent chromatin (Tackett et al., 2005). Recently Yta7 was also reported to bind to the promoter regions of histone genes and to regulate their transcriptional expression (Gradolatto et al., 2008). Therefore, I was surprised to observe Yta7 as one of the top preys associated with both Hta2-TAP and Htz1-TAP, hinting to a broad distribution for this protein on chromatin and a potentially global regulation of chromatin structure for Yta7 in *S. cerevisiae*. To further investigate, the mChIP method was applied to identify the chromatin-associated protein networks of Yta7 (Figure 2-4A). Yta7-TAP was successfully purified using the mChIP method, and 76 proteins were reproducibly identified (Figure 2-4). The overlap between Yta7-TAP and Hta2-TAP protein interaction profiles was large at 53 proteins, suggesting a wide distribution of Yta7 along chromosomes

(Figure 2-4B). However, 23 proteins were specific to Yta7-TAP (Figure 2-4B). Previously reported interaction partners of Yta7 such as the core histones, Htz1, Top2, Spt16, Sas3, and Ylr455w (Tackett et al., 2005) were observed upon mass spectrometric analysis (see Appendix 2-1 for a complete list of interaction partners). In addition our mChIP method identified numerous other proteins including Pob3, a subunit of the Facilitates Chromatin Transcription (FACT) complex known to interact with Spt16 (Wittmeyer and Formosa, 1997); Chd1, a nucleosome-remodelling factor containing a chromodomain (Tran et al., 2000); Nto1, a subunit of the NuA3 acetyltransferase complex (for a review, see Reference (Lafon et al., 2007)), and others. This suggests that despite the broad distribution of Yta7 it is likely found in specific chromatin microenvironments.

2.4 Discussion

In this study, I present a novel proteomics method, termed mChIP, for the purification of proteins that interact with chromatin. My method increases the solubility of protein-DNA complexes by introducing a DNA shearing step and subsequently minimizing centrifugation in the sample preparation procedure. Together these steps ensure the maintenance of protein-DNA macromolecules in the soluble fraction of the whole cell extract and hence make them available for purification in sufficient quantities for mass spectrometric analysis. Using Hta2-TAP and Htz1-TAP as tests baits I applied the mChIP method to build a chromatin-bound protein network in yeast. To further demonstrate the

power of this method, I applied it to three non-histone proteins and demonstrated that the mChIP method is specific and sensitive for these functionally distinct baits.

2.4.1 mChIP Method Expands Chromatin-bound Protein Networks

I initially tested the mChIP method for its ability to purify chromatin-bound protein networks associated with the canonical histone H2A (Hta2) and its variant (Htz1) in *S. cerevisiae*. I observed that our newly developed method minimized the co-purification of nonspecific interacting proteins (Appendix 2-1) while providing a wealth of new information concerning protein interactions of the histone baits (Figure 2-2F and Appendix 2-1). Additionally I determined that contrary to classical TAP (Rigaut et al., 1999) the mChIP method efficiently purifies both chromatin-bound and cytosolic interaction partners of Htz1-TAP. This was accomplished by purifying Htz1-TAP in cells lacking SWR1. Swr1 is the enzyme responsible for the insertion of Htz1 into nucleosome, and its deletion greatly reduces Htz1 interaction with chromatin (Kobor et al., 2004; Krogan et al., 2004; Mizuguchi et al., 2004). Notably SWR1 deletion causes a significant reduction in the cellular level of Htz1 as compared with wild type cells (Keogh et al., 2006). However, because of the efficiency of the mChIP method, Htz1-TAP and its interaction partners were still immunopurified in *swr1Δ* cells in sufficient quantity to allow for MS analysis. I observed that in the absence of Swr1 the chromatin-bound interaction partners of Htz1 were lost, whereas the interaction with the two histone chaperones Nap1 and Kap114, present in the cytoplasm and nucleus, were maintained (Figure 2-2A, Appendix 2-2). This supports my

view that the DNA and chromatin-bound proteins co-purified with the baits are not artefacts produced in the sample preparation but represent genuine interactions. Moreover, I demonstrated that by adjusting the DNA fragment size in the cellular extracts through the use of enzymes I can modulate the extent of interaction partners that co-purify with Htz1-TAP bait (Figure 2-3). This technique may be used to precisely define chromatin-bound protein environments on chromatin that to date have not been extensively characterized.

An interesting side product observed in the development of the mChIP is the ability to efficiently co-purify DNA associated with the bait of interest with this new method (Figure 2-2D). This was exploited by my collaborators to localize the Msn4 transcription factor to the *HSP12* gene under stress conditions by mChIP (Mitchell et al., 2008). Chromatin immunoprecipitation that circumvents the use of formaldehyde, often named native ChIP have been described previously (O'Neill and Turner, 2003), but these protocols are lengthy and restricted to histone proteins. In contrast the mChIP method produced good results for the isolation of DNA associated with both histone (Figure 2-2D) and non-histone proteins (Mitchell et al., 2008) while remaining a relatively fast and convenient protocol.

mChIP experiments performed using histone baits provide an efficient means to define the chromatin protein back- ground, or base line, in *S. cerevisiae* with which other mChIP purifications may be compared (Figure 2-4B). In addition, it enabled me to define the depth of the current literature regarding chromatin-based protein interactions. When

comparing the Hta2-TAP mChIP data with previously reported Hta2-TAP IP-MS experiments, it became apparent that the mChIP procedure significantly expanded the number of associated proteins detected (Figure 2-2F). Furthermore Hta2-TAP interaction partners obtained through mChIP had a higher proportion of nuclear localization (Figure 2-2F) reflecting an increase in sensitivity and specificity of the mChIP method because of the special care taken to maintain the integrity of macromolecules during the purification process (Figure 2-1).

2.4.2 Efficient mChIP Purifications of Non-histone Chromatin Baits

The new protocol I developed was optimized using histones as the archetype of chromatin-associated baits. These baits by their very nature are strongly bound to DNA and constitute the foundation of chromatin. A crucial test of the mChIP protocol was to efficiently purify non-histone baits associated with chromatin. Yta7, Mcm5, and Lge1 were successfully purified by mChIP without the need of any bait- specific optimization (Figure 2-4A). Furthermore the protein networks identified for each baits are significantly different, demonstrating the specificity of the mChIP procedure (Figures 2-4B and 2-5). For example, Mcm5 is known to be present at replication forks, which are very specialized chromatin environments that are devoid of most common chromatin protein (Groth et al., 2007). This is reflected by the mChIP data for Mcm5 that contain a small fraction of the previously defined chromatin background by histone mChIP (Appendix 2-2). On the other hand, Yta7 has been shown recently to be bound at the promoter sequence of a large number of genes

(Gradolatto et al., 2008). This is in agreement with the mChIP data obtained for Yta7 that contain a large proportion of the proteins associated with histone baits by mChIP, reflecting its broad chromatin localization. Thus, mChIP data can be use to provide an indirect readout of the distribution of a bait protein on chromatin.

When the mChIP data for the five bait proteins was compared with protein-protein interaction data contained in the BioGRID, I observed that the overlap was particularly low for Mcm5 and Lge1 (Figure 2-6). This is consistent with the view that traditional TAPs are not effective for numerous chromatin-associated proteins and thus poorly represented in the literature (Jensen and Bork, 2008). Mcm5 provide a clear example of this situation. A total of 16 distinct interaction partners have been curated in the BioGRID with Mcm5 (Appendix 2-2). Interestingly only three of these interaction partners are known to be present in the nucleus (Appendix 2-2). In contrast, 13 interaction partners were found for Mcm5 in at least two of three mChIP purifications (Appendix 2-2) with 10 of these found in the nucleus (~75%) representing a clear enrichment for this localization (p-value <0.005). This effectively demonstrates the improvement in the coverage of the Mcm5 nuclear interactome by mChIP as compared with classical methodologies.

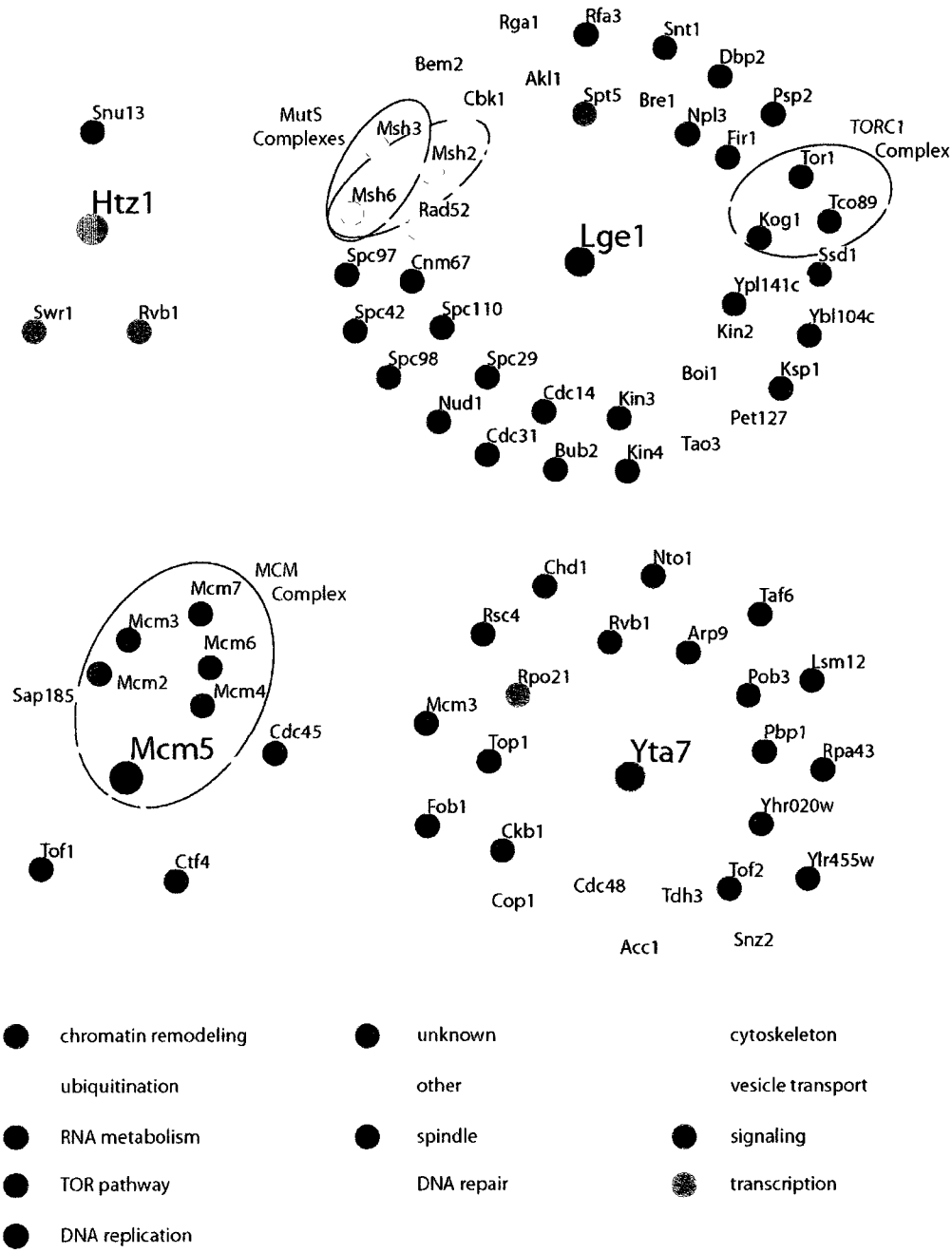


Figure 2-5: Distinct protein networks co-purify by mChIP with histone Htz1- TAP, Lge1-TAP, Mcm5-TAP, and Yta7-TAP (baits are shown in bold text). Associated preys displayed are those reproducibly identified (minimum of two of three mChIPs) that were not identified in Hta2-TAP mChIP (Appendix 2-2). Osprey was used to display proteins identified in one or more mChIP purifications for each bait protein. *Gray edges* represent physical interaction observed by mChIP, and *node color* corresponds to gene ontology annotations as listed in the *Saccharomyces* Genome Database. Some known protein complexes are denoted by a *black circle*.

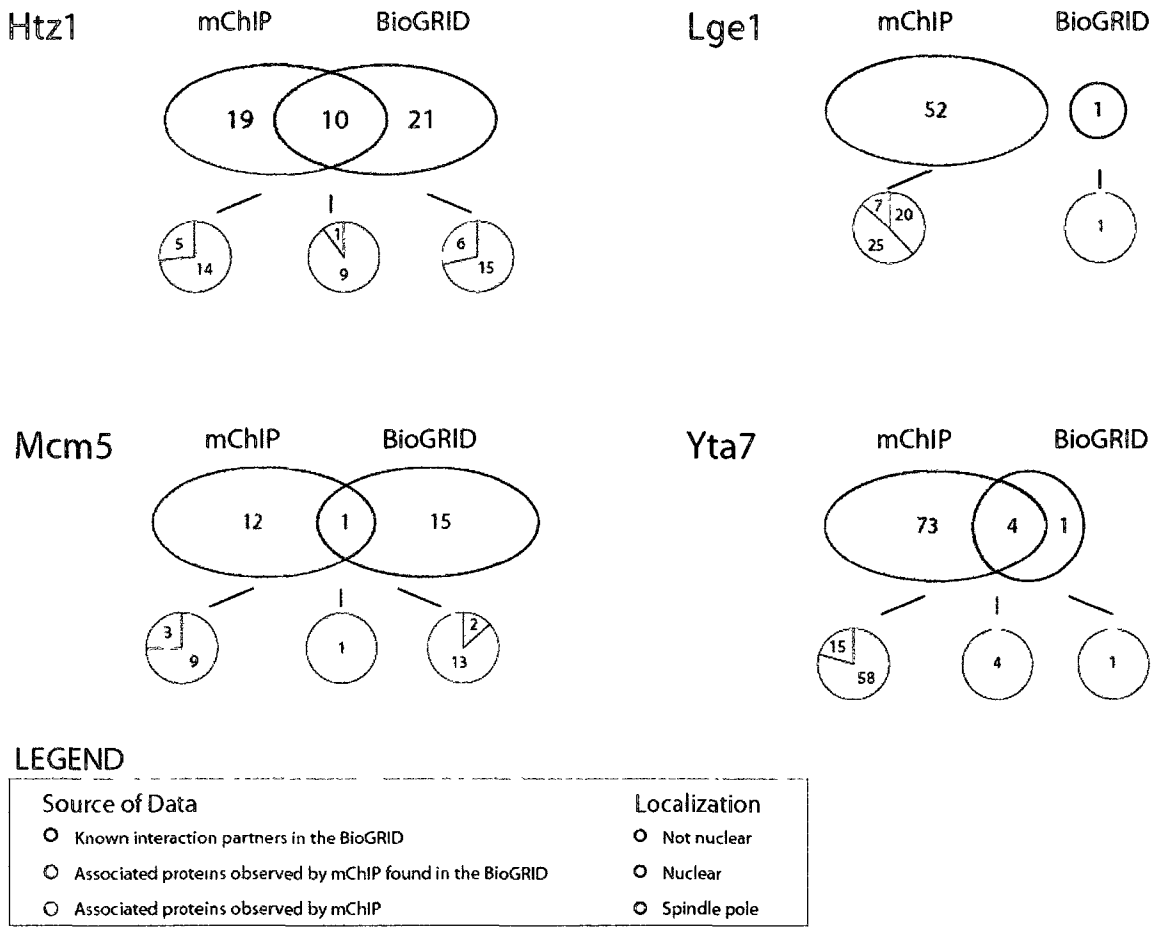


Figure 2-6: mChIP improve the depth and nuclear coverage of protein networks associated with chromatin-bound proteins. Shown are Venn diagrams of unique and shared associated proteins found in two of three mChIP purifications and traditional TAPs as curated in the BioGRID (Appendix 2-2). The localization of each associated proteins was assessed by using the yeast green fluorescent protein fusion localization database (Huh et al., 2003).

2.4.3 Gain of Novel Protein-Protein Interaction Information through mChIP Purifications

The mChIP method was also shown to be successful at purifying large protein networks associated with poorly characterized non-histone proteins (Fig. 2-4). The most striking example of this is the success of the mChIP procedure at purifying Lge1. To my knowledge, the only reported protein-protein interactions for Lge1 is with Bre1 (Collins et al., 2007; Ho et al., 2002; Krogan et al., 2006), Isw1 (Collins et al., 2007), and Ybr296c-A (Tarassov et al., 2008). Despite this apparently limited interactome of Lge1, I observed over 50 proteins co-purifying with Lge1 in at least two distinct mChIP purifications (Figure 2-4B), most of which were not part of the chromatin background defined by Hta2-TAP (Figure 2-5). In particular, various protein complexes, such as TORC1 and the SPB, were observed to associate with Lge1 in a reproducible manner by mChIP. This new layer of information now accessible through the use of the mChIP procedure can help explain some of the known phenotypes associated with the *LGE1* gene. For instance, the most drastic phenotype of *lge1Δ* cells is their size (Jorgensen et al., 2002). Lge1 mChIP revealed a significant enrichment for the TORC1 complex (Tor1, Kog1, Tco89, and Lst8). The TORC1 complex has been shown to regulate the cell cycle by integrating a number of cell signals (nutrients, hormones, and growth factors levels) (Loewith et al., 2002). The exact mechanism of the TORC1 complex regulation of the cell cycle remains nebulous, but it was observed in fission yeast (*Schizosaccharomyces pombe*) that TORC1 inhibition through rapamycin treatment results in a significant reduction of cell size (Petersen and Nurse, 2007). In *S. cerevisiae*, *lge1Δ* cells were observed to tolerate rapamycin treatment better than wild-type cells in a genome-wide study (Xie et al., 2005). This partial rescue to TORC1 inhibition by rapamycin

through the deletion of *lge1*, in addition to their association in mChIP purification, points to their participation in common cellular events. One such event could be in the control of cell cycle through the monoubiquitination of histone H2B at lysine 123. Hwang et al. (Hwang et al., 2003) showed that mutation of histone H2B at lysine 123 to an arginine (K123R), preventing its monoubiquitination, resulted in an increase in cell size. The authors speculated that histone H2B monoubiquitination could function as part of a "biochemical calculation" that could participate in determining cell size and impact cell cycle check, such as START initiation (Hwang et al., 2003). The resistance of *lge1Δ* cells to rapamycin is also shared by *bre1Δ* and *rad6Δ* cells (Xie et al., 2005); these proteins are directly involved in histone H2B Lys-123 monoubiquitination. In light of my Lge1 mChIP and of the previous literature regarding *LGE1* and the TORC1 complex, the TORC1 complex could participate in this biochemical calculation to efficiently regulate the cell cycle.

A few years ago, Zhang et al. (Zhang et al., 2005) demonstrated that Complex Proteins Associated with Set1 (COMPASS) was capable of dimethylating lysine 233 of Dam1, a protein localizing to the kinetochore that is required for proper chromosome segregation and spindle integrity (Courtwright and He, 2002). This process was shown to be independent of Bre1-Rad6 and of H2B Lys-123 monoubiquitination (Zhang et al., 2005), but the possibility of a cross-talk between COMPASS and Bre1-Rad6 on non-histone proteins was raised (Wood and Shilatifard, 2005). Although this possibility has not been demonstrated to be correct yet, it remains an attractive hypothesis in light of the large Lge1 interactome observed by mChIP that would locate Lge1 (and potentially Bre1) to the spindle pole body. Whether Bre1 possesses an enzymatic activity toward any spindle pole

body constituents is still unknown, but this process could be linked to proper cell cycle progression and thus be related to the known *lge1Δ* cell phenotype.

In summary, I have developed a novel method for the purification of protein networks associated with chromatin. The mChIP method was shown to be capable of purifying a large number of chromatin-bound protein complexes associated with the histones Hta2 and Htz1 in an efficient and specific manner. The mChIP protocol was also capable of establishing protein networks of chromatin-associated baits that are hard to purify by traditional AP-MS. Importantly, I demonstrated that this method is not limited to purifying proteins that are physically linked to the bait but can purify proteins that interact with the DNA or chromatin with which the bait interacts. This represents an advance in the identification of protein-protein interactions from the detection of direct interactions to a more global association between complexes or macromolecular environments. As I have shown in this chapter, the ability to efficiently and specifically purify protein networks on chromatin can provide a novel means to improve our understanding of protein functions. In the next chapter, I show a large-scale application of the mChIP procedure to explore the chromatin networks associated of poorly characterized DNA-binding proteins in S. cerevisiae and how this can improve our comprehension of their individual functions.

Bibliography:

Abu-Farha, M., Lambert, J.P., Al-Madhoun, A.S., Elisma, F., Skerjanc, I.S., and Figeys, D. (2008). The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase. Mol Cell Proteomics *7*, 560-572.

Aparicio, O.M., Weinstein, D.M., and Bell, S.P. (1997). Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell *91*, 59-69.

Barbet, N.C., Schneider, U., Helliwell, S.B., Stansfield, I., Tuite, M.F., and Hall, M.N. (1996). TOR controls translation initiation and early G1 progression in yeast. Mol Biol Cell *7*, 25-42.

Berger, J.M. (1998). Type II DNA topoisomerases. Curr Opin Struct Biol *8*, 26-32.

Boulard, M., Bouvet, P., Kundu, T.K., and Dimitrov, S. (2007). Histone variant nucleosomes: structure, function and implication in disease. Subcell Biochem *41*, 71-89.

Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T., and Drewes, G. (2004). Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. J Biol Chem *279*, 12804-12811.

Collas, P., and Dahl, J.A. (2008). Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. Front Biosci *13*, 929-943.

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

Courtwright, A.M., and He, X. (2002). Dam1 is the right one: phosphoregulation of kinetochore biorientation. Dev Cell *3*, 610-611.

Cullmann, G., Fien, K., Kobayashi, R., and Stillman, B. (1995). Characterization of the five replication factor C genes of Saccharomyces cerevisiae. Mol Cell Biol *15*, 4661-4671.

Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem *277*, 28368-28371.

Du, Y.C., Gu, S., Zhou, J., Wang, T., Cai, H., Macinnes, M.A., Bradbury, E.M., and Chen, X. (2006). The dynamic alterations of H2AX complex during DNA repair detected by a proteomic approach reveal the critical roles of Ca(2+)/calmodulin in the ionizing radiation-induced cell cycle arrest. Mol Cell Proteomics *5*, 1033-1044.

Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat Methods *2*, 667-675.

Ethier, M., Lambert, J.P., Vasilescu, J., and Figeys, D. (2006). Analysis of protein interaction networks using mass spectrometry compatible techniques. Anal Chim Acta *564*, 10-18.

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

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Forsburg, S.L. (2008). The MCM helicase: linking checkpoints to the replication fork. Biochem Soc Trans *36*, 114-119.

Galasinski, S.C., Resing, K.A., and Ahn, N.G. (2003). Protein mass analysis of histones. Methods *31*, 3-11.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol *8*, 358-366.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., *et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Gradolatto, A., Rogers, R.S., Lavender, H., Taverna, S.D., Allis, C.D., Aitchison, J.D., and Tackett, A.J. (2008). Saccharomyces cerevisiae Yta7 regulates histone gene expression. Genetics *179*, 291-304.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin challenges during DNA replication and repair. Cell *128*, 721-733.

Hanway, D., Chin, J.K., Xia, G., Oshiro, G., Winzeler, E.A., and Romesberg, F.E. (2002). Previously uncharacterized genes in the UV- and MMS-induced DNA damage response in yeast. Proc Natl Acad Sci U S A *99*, 10605-10610.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180-183.

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

Hwang, W.W., Venkatasubrahmanyam, S., Ianculescu, A.G., Tong, A., Boone, C., and Madhani, H.D. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. Mol Cell *11*, 261-266.

Jaspersen, S.L., and Winey, M. (2004). The budding yeast spindle pole body: structure, duplication, and function. Annu Rev Cell Dev Biol *20*, 1-28.

Jensen, L.J., and Bork, P. (2008). Biochemistry. Not comparable, but complementary. Science *322*, 56-57.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J., and Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. Science *297*, 395-400.

Keogh, M.C., Mennella, T.A., Sawa, C., Berthelet, S., Krogan, N.J., Wolek, A., Podolny, V., Carpenter, L.R., Greenblatt, J.F., Baetz, K., and Buratowski, S. (2006). The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. Genes Dev *20*, 660-665.

Kobor, M.S., Venkatasubrahmanyam, S., Meneghini, M.D., Gin, J.W., Jennings, J.L., Link, A.J., Madhani, H.D., and Rine, J. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol *2*, E131.

Krogan, N.J., Baetz, K., Keogh, M.C., Datta, N., Sawa, C., Kwok, T.C., Thompson, N.J., Davey, M.G., Pootoolal, J., Hughes, T.R., *et al.* (2004). Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. Proc Natl Acad Sci U S A *101*, 13513-13518.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., *et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V., *et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell *12*, 1565-1576.

Lafon, A., Chang, C.S., Scott, E.M., Jacobson, S.J., and Pillus, L. (2007). MYST opportunities for growth control: yeast genes illuminate human cancer gene functions. Oncogene *26*, 5373-5384.

Lee, J.S., Shukla, A., Schneider, J., Swanson, S.K., Washburn, M.P., Florens, L., Bhaumik, S.R., and Shilatifard, A. (2007). Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell *131*, 1084-1096.

Loewith, R., Jacinto, E., Wullschleger, S., Lorberg, A., Crespo, J.L., Bonenfant, D., Oppliger, W., Jenoe, P., and Hall, M.N. (2002). Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. Mol Cell *10*, 457-468.

Longhese, M.P., Plevani, P., and Lucchini, G. (1994). Replication factor A is required in vivo for DNA replication, repair, and recombination. Mol Cell Biol *14*, 7884-7890.

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luke, M.M., Della Seta, F., Di Como, C.J., Sugimoto, H., Kobayashi, R., and Arndt, K.T. (1996). The SAP, a new family of proteins, associate and function positively with the SIT4 phosphatase. Mol Cell Biol *16*, 2744-2755.

Mellor, J., and Morillon, A. (2004). ISWI complexes in Saccharomyces cerevisiae. Biochim Biophys Acta *1677*, 100-112.

Mitchell, L., Lambert, J.P., Gerdes, M., Al-Madhoun, A.S., Skerjanc, I.S., Figeys, D., and Baetz, K. (2008). Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 is essential for complex integrity. Mol Cell Biol *28*, 2244-2256.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mosammaparast, N., Ewart, C.S., and Pemberton, L.F. (2002). A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. EMBO J *21*, 6527-6538.

Mujtaba, S., Zeng, L., and Zhou, M.M. (2007). Structure and acetyl-lysine recognition of the bromodomain. Oncogene *26*, 5521-5527.

O'Neill, L.P., and Turner, B.M. (2003). Immunoprecipitation of native chromatin: NChIP. Methods *31*, 76-82.

Parrish, J.R., Gulyas, K.D., and Finley, R.L., Jr. (2006). Yeast two-hybrid contributions to interactome mapping. Curr Opin Biotechnol *17*, 387-393.

Petersen, J., and Nurse, P. (2007). TOR signalling regulates mitotic commitment through the stress MAP kinase pathway and the Polo and Cdc2 kinases. Nat Cell Biol *9*, 1263-1272.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218-229.

Reinke, A., Anderson, S., McCaffery, J.M., Yates, J., 3rd, Aronova, S., Chu, S., Fairclough, S., Iverson, C., Wedaman, K.P., and Powers, T. (2004). TOR complex 1 includes a novel component, Tco89p (YPL180w), and cooperates with Ssd1p to maintain cellular integrity in Saccharomyces cerevisiae. J Biol Chem *279*, 14752-14762.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol *17*, 1030-1032.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nat Protoc *2*, 1445-1457.

Sheu, Y.J., and Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. Mol Cell *24*, 101-113.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Smith, J.C., Lambert, J.P., Elisma, F., and Figeys, D. (2007). Proteomics in 2005/2006: developments, applications and challenges. Anal Chem *79*, 4325-4343.

Stark, C., Breitkreutz, B.J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Res *34*, D535-539.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M., Aitchison, J.D., Rout, M.P., and Chait, B.T. (2005). Proteomic and genomic characterization of chromatin complexes at a boundary. J Cell Biol *169*, 35-47.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S.W. (2008). An in vivo map of the yeast protein interactome. Science *320*, 1465-1470.

Tran, H.G., Steger, D.J., Iyer, V.R., and Johnson, A.D. (2000). The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. EMBO J *19*, 2323-2331.

Van Holde, K.E., Sahasrabuddhe, C.G., and Shaw, B.R. (1974). A model for particulate structure in chromatin. Nucleic Acids Res *1*, 1579-1586.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature *379*, 466-469.

Wittmeyer, J., and Formosa, T. (1997). The Saccharomyces cerevisiae DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein. Mol Cell Biol *17*, 4178-4190.

Wood, A., Krogan, N.J., Dover, J., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J.F., *et al.* (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell *11*, 267-274.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2005). The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. Mol Cell *20*, 589-599.

Wood, A., and Shilatifard, A. (2005). Guided by COMPASS on a journey through chromosome segregation. Nat Struct Mol Biol *12*, 839-840.

Xie, M.W., Jin, F., Hwang, H., Hwang, S., Anand, V., Duncan, M.C., and Huang, J. (2005). Insights into TOR function and rapamycin response: chemical genomic profiling by using a high-density cell array method. Proc Natl Acad Sci U S A *102*, 7215-7220.

Zhang, K., Lin, W., Latham, J.A., Riefler, G.M., Schumacher, J.M., Chan, C., Tatchell, K., Hawke, D.H., Kobayashi, R., and Dent, S.Y. (2005). The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. Cell *122*, 723-734.

Chapter 3

Defining the Budding Yeast Chromatin Associated Interactome

The work presented in this chapter was accepted for publication as:

Jean-Philippe Lambert, Jeffrey Fillingham, Mojgan Siahbazi, Jack Greenblatt, Kristin Baetz and Daniel Figeys. *Defining the Budding Yeast Chromatin Associated Interactome.* Molecular Systems Biology, accepted for publication on November 4th 2010.

Permission to reprint this work was obtained from the *Nature Publishing Group.*

Authors contributions:

JPL performed the purifications and MS analyses, analyzed the data, wrote the manuscript and produced all figures except 3-11D and 3-11E.

JF generated figure 3-11D and 3-11E, provided reagents and edited the manuscript in part under the supervision of JG.

MS generated reagents under the supervision of KB.

KB edited the manuscript and provided reagents.

DF directed the project and writing of the manuscript.

Abstract:

Recently, I reported a novel affinity purification method termed modified Chromatin immunopurification (mChIP) which permits selective enrichment of DNA-bound proteins along with their associated protein network. In this chapter, I report a large-scale study of the protein network of 102 chromatin related proteins from budding yeast that were analyzed by mChIP coupled to mass spectrometry. This effort resulted in the detection of 3576 high confidence protein associations with 724 distinct preys. Approximately 75% of the baits had significantly improved interaction coverage using mChIP compared to the classical affinity purification methodologies. Furthermore, many lower abundance transcription factors that previously failed using conventional affinity purification methodologies were successful with mChIP. mChIP was also used to perform a targeted study of Asf1 and its associated proteins to better unravel the physical interplay between Asf1 and two other histone chaperones, Rtt106 and the HIR complex.

3.1 Introduction:

The maintenance of cellular fitness requires living organisms to integrate multiple signals into coordinated outputs. Central to this process is the regulation of the expression of the genetic information encoded into DNA. As a result, there are numerous constraints imposed on gene expression. The access to DNA is restricted by the formation of nucleosomes in which DNA is wrapped around histone octamers to form chromatin wherein the volume of DNA is considerably reduced. As such, nucleosome positioning is critical and must be defined precisely, particularly during transcription (Workman, 2006). Furthermore, nucleosomes can be actively assembled/disassembled by histone chaperones and can be made to "slide" along DNA by the actions of chromatin remodelers. Moreover, the histone proteins are heavily regulated at the expression level and by extensive post-translational modifications (PTM) (Campos and Reinberg, 2009). Histone PTMs have also been shown to help recruit numerous chromatin-associated factors in accordance with the histone code (Strahl and Allis, 2000). Although our understanding of chromatin and its roles has improved, we still have limited knowledge of the chromatin-associated protein complexes and their interactions.

Progress in the chromatin field has been closely intertwined with technical improvements in both genomic and proteomic technologies. For instance, the chromatin immunoprecipitation (ChIP) protocol has been used for many years to define the binding sites of a protein on DNA (Kuo and Allis, 1999). Early uses of the ChIP protocol were coupled to PCR and restricted to the study of a few genomic loci at a time. Currently, it is possible to determine the location of a given protein throughout a whole genome due to

the development of better detection platforms, such as ChIP-chip (Ren et al., 2000) and ChIP-seq (Barski et al., 2007). The analysis of proteins associated with chromatin has also benefited from this technical progress. For instance, detailed analyses of histone isomers and their PTMs by MS have been conducted in numerous organisms (Bonenfant et al., 2006; Masumoto et al., 2005; Thomas et al., 2006). These analyses enabled researchers to identify novel modifications (Garcia et al., 2007) and to uncover cooperative actions among multiple histone modifications (Jiang et al., 2007; Taverna et al., 2007). This extra level of complexity was not apparent in studies that used traditional methods.

One area of chromatin research that still requires technical improvement is the identification and characterization of protein complexes associated with chromatin (Lambert et al., 2010). Affinity purification and mass spectrometry (AP-MS) has emerged as a powerful tool for characterizing protein-protein interactions and biological systems in general (Gingras et al., 2007; Gstaiger and Aebersold, 2009). To date, AP-MS has been successfully applied to the majority of model organisms including budding yeast (Rigaut et al., 1999), fission yeast (Cipak et al., 2009; Kim et al., 2009a), *Drosophila melanogaster* (Veraksa et al., 2005), *Caenorhabditis elegans* (Ooi et al., 2010), mouse (Bienvenu et al., 2010), mouse stem cells (Kim et al., 2009b) and human cells (Glatter et al., 2009). Furthermore, numerous large-scale studies have been performed both in budding yeast (Gavin et al., 2006; Ho et al., 2002; Krogan et al., 2006) and in human cells (Ewing et al., 2007) resulting in an improved characterization for thousands of gene products. As well, our understanding of many chromatin related processes, such as transcription and translation, has greatly benefited from AP-MS studies. Indeed, two recent applications

focused on the study of the transcription machinery and encompassed 77 baits involved in RNA polymerase II (RNAPII). These bait proteins were immunopurified by tandem affinity purification (TAP) and analyzed by MS (Jeronimo et al., 2007)(Cloutier et al., 2009), leading to the discovery of novel proteins associated with RNAPII. Despite the success of the previous studies, they all targeted protein complexes that were present in the soluble fraction of the cell and not the less soluble proteins which are associated with chromatin.

Two techniques were reported to enable the purification of protein complexes associated with particular genomic loci. The first approach relies on specific nucleic acid probes which are affixed to solid supports (i.e. beads). These nucleic acid sequences act as affinity probes and replace antibodies. The proteins associated with the nucleic acid probes can then be selectively purified and subsequently identified by mass spectrometry (Burckstummer et al., 2009; Dejardin and Kingston, 2009; Rubio et al., 2008; Schultz-Norton et al., 2008). The second approach uses mini-chromosomes that contain sequences of interest flanked with repetitive Lac operator sequences. The mini-chromosomes are then selectively purified from the bulk of chromatin using an immobilized Lac repressor (Akiyoshi et al., 2009; Unnikrishnan et al., 2010). These two approaches are well suited for studying specific genomic loci and their associated protein complexes. Unfortunately, these methods are limited in their applicability because they require many specialized tools (affinity probes), they focus only on distinct genomic loci and they require a large amount of materials. Thus, another approach is required for performing large-scale studies involving multiple baits.

As I described in Chapter 2, the modified Chromatin Immunopurification (or mChIP) approach efficiently purifies protein-DNA macromolecular complexes and enables their subsequent analysis by mass spectrometry (Lambert et al., 2009). The mChIP method consists of a single affinity purification step whereby chromatin-associated proteins are isolated from mildly sonicated and gently clarified cellular extracts using magnetic beads coated with antibodies (Lambert et al., 2009). mChIP was successfully applied to the study of both histones (Chapter 2; (Lambert et al., 2009)) and non-histone (Chapter 2; (Fillingham et al., 2009; Lambert et al., 2009)) chromatin-associated proteins. Furthermore, the mChIP method was shown to drastically increase the coverage of the interactome for chromatin-associated proteins that are difficult to purify, such as Lge1 and Yta7 (Chapter 2; (Lambert et al., 2009)). Moreover, mChIP can sensitively identify protein associations (which include both direct and indirect protein-protein interactions) present only at a few genomic loci (Fillingham et al., 2009).

In this chapter, I report the first large-scale mChIP characterization of the chromatin interactome in budding yeast. As part of this study, 102 baits known to bind DNA or with functional links to chromatin were successfully purified by mChIP. Mass spectrometry was used to identify the chromatin proteins associated with these baits. This mChIP study of the chromatin interactome resulted in the detection of 3576 high confidence protein associations with 724 distinct preys. To my knowledge, this is the first large-scale effort to map the chromatin-associated protein-protein interaction network.

3.2 Experimental Procedures:

3.2.1 Yeast Strains, Plasmids, and Genetics Methods:

All yeast strains and plasmids used in this study are listed in Table 3-1 and 3-2, respectively. Growth media and strains were prepared following standard practices. Strains from the TAP collection were obtained from Open Biosystems (Huntsville, AL). Genomic deletions and epitope tag integrations that were made for this study were designed with PCR-amplified cassettes as described previously (Longtine et al., 1998; Puig et al., 2001) and confirmed by either PCR analysis or immunoblotting for tag expression.

Table 3-1: List of strains used in chapter 3.

Strain	Description	Reference
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	*(Sikorski and Hieter, 1989)*
TAP strains	BY4741, BAIT-TAP::HIS3	*(Howson et al., 2005)*
YJP023	BY4741 *trp1Δ*::HIS3, CPR1-TAP::HIS3 +Ub-myc	this study
YJP024	BY4741 *trp1Δ*::HIS3, CPR1-TAP::HIS3 +Ub-myc (K48R G76A)	this study
JF08-22	BY4741, RTT106-TAP::HIS3 *asf1Δ*::NAT	*(Fillingham et al., 2009)*
YJP167	BY4741, HIR1-myc::KAN	this study
YJP168	BY4741, HIR1-myc::KAN *asf1Δ*::NAT	this study
YJP169	BY4741, RTT106-TAP::HIS3 Hir1-myc::KAN	this study
YJP171	BY4741, RTT106-TAP::HIS3 Hir1-myc::KAN *asf1Δ*::NAT	this study
YJP180	BY4741 *trp1Δ*::HIS3, RTT106-TAP::HIS3 *asf1Δ*::NAT, pRS314	this study
YJP181	BY4741 *trp1Δ*::HIS3, RTT106-TAP::HIS3 *asf1Δ*::NAT, pRS314 ASF1	this study
YJP182	BY4741 *trp1Δ*::HIS3, RTT106-TAP::HIS3 *asf1Δ*::NAT, pRS314 ASF1-V94R	this study
YJP183	BY4741 *trp1Δ*::HIS3, HIR1-TAP::HIS3 *asf1Δ*::NAT, pRS314	this study
YJP184	BY4741 *trp1Δ*::HIS3, HIR1-TAP::HIS3 *asf1Δ*::NAT, pRS314 ASF1	this study
YJP185	BY4741 *trp1Δ*::HIS3, HIR1-TAP::HIS3 *asf1Δ*::NAT, pRS314 ASF1-V94R	this study
YJP197	BY4741, ASF1-TAP::HIS3 *rtt106Δ*::NAT	this study
YJP198	BY4741, ASF1-TAP::HIS3 *hir1Δ*::KAN	this study
YJP225	BY4741, ASF1-TAP::HIS3 *rtt109Δ*::NAT	this study
YJP226	BY4741, ASF1-TAP::HIS3 *vps75Δ*::NAT	this study
YJP227	BY4741, ASF1-TAP::HIS3 *hat2Δ*::NAT	this study
MSY421	MATa, Δ(*hht1-hhf1*) Δ(*hht2-hhf2*) leu2-3, 112,ura3-62, trp1, his3, pMS329 [HHT1-HHF1, URA3, CEN]	*(Recht et al., 2006)*
YJP199	MSY421, ASF1-TAP::HIS3 + [pRS314-HHT2-HHF2]	this study
JR1	MSY421 + [pRS314-H3K56R, H4WT]	*(Recht et al., 2006)*
YJP200	JR1, ASF1-TAP::HIS3	this study
JF08-23	BY4741, Rtt106-TAP::HIS3 *hir1Δ*::KAN	*(Fillingham et al., 2009)*
JF08-24	BY4741, Rtt106-TAP::HIS3 *rtt109Δ*::NAT	*(Fillingham et al., 2009)*
JF08-237	MSY421, Rtt106-TAP::HIS3 + [pRS314-HHT2-HHF2]	this study
JF08-238	JR1, Rtt106-TAP::HIS3 + [pRS314-H3K56R, H4WT]	this study

Table 3-2: List of plasmid utilized in this chapter.

Plasmid	Description	Reference
PEy105	*pCUP1-UBI HIS-MYC TRP1 CEN*	*(Hochstrasser et al., 1991)*
pUB223	*pCUP1-UBI HIS-MYC-K48R G76A TRP1 CEN*	*(Willems et al., 1996)*
pRS314	*pRS314 TRP1 CEN*	*(English et al., 2006)*
pRS314 Asf1	*pRS314 ASF1-myc TRP1 CEN*	*(English et al., 2006)*
pRS314 Asf1 V94R	*pRS314 ASF1-V94R-myc TRP1 CEN*	*(English et al., 2006)*
pRS314-HHT2-HHF2	*pRS314-HHT2-HHF2 TRP1 CEN*	*(Recht et al., 2006)*
pRS314-H3K56R, H4WT	*pRS314-H3K56R, H4WT TRP1 CEN*	*(Recht et al., 2006)*

3.2.2 Modified Chromatin Immunopurification:

Modified Chromatin Immunopurification was performed as per Chapter 2 (published as (Lambert et al., 2009)) with minors modifications. Briefly, one-step affinity immunopurification was performed using TAP-tagged proteins and M-270 epoxy Dynabeads (Invitrogen) coated with rabbit IgG (Sigma-Aldrich) according to the manufacturer's instructions. Briefly, 700 ml of cultured yeast cells grown in yeast, peptone, and dextrose (YPD) medium to an OD_{600} of ~1 were pelleted and washed with water. Cells were resuspended in a lysis buffer (100 mM HEPES, pH 8.0, 20 mM magnesium acetate, 10% glycerol (v/v), 10 mM EGTA, 0.1 mM EDTA with fresh protease inhibitors mixture (Roche) and phosphatase inhibitors mixture (Roche)), frozen in liquid nitrogen in small droplets, and lysed using a coffee grinder half-filled with dry ice for 1 min. The dry ice from the ground cells was allowed to evaporate, and the resulting whole cell extract was sonicated three times for 30 seconds with at least 1 min on ice between each pulse. Nonidet P-40 was added to a final concentration of 0.4%, and the sample was mixed by hand for 30 seconds. The extract was gently clarified by centrifugation at 1800g for 10 min (4°C), and the supernatant was transferred to a fresh tube. In some cases, 75 units of Benzonase (Sigma-Aldrich) were added per mL of protein extract to completely remove DNA. Freshly prepared rabbit IgG coated Dynabeads were added (200 µL/sample), and the samples were incubated with end-over-end rotation for 3 hours at 4°C. Using a Dynal MPC-S magnet (Invitrogen) the beads were collected on the side of the sample tubes, and the supernatant was discarded. The beads were washed three times in fresh tubes by resuspension and transfer in 1 ml of ice-cold wash buffer (100 mM HEPES, pH 7.4, 20 mM

magnesium acetate, 10% glycerol (v/v), 10 mM EGTA, 0.1 mM EDTA, 0.5% Nonidet P-40). Finally, the beads were resuspended in 1 ml of elution buffer (0.5 M NH_4OH pH >11, 0.5 mM EDTA) and incubated with end-over-end rotation for 20 min at room temperature. The protein eluates were transferred to fresh tubes and were evaporated to dryness using a SpeedVac with no heat. The protein sample was resuspended in 1X loading buffer (50 mM Tris-HCl, pH 8, 2% SDS, 100 mM DTT, 10% glycerol) and resolved on a NuPAGE 4–12% SDS-PAGE gel with MES buffer unless mentioned otherwise. For protein visualization, the gels were silver-stained or stained with coomassie blue. For Western blot analysis, the proteins were transferred onto a nitrocellulose membrane, blocked in 5% nonfat milk in TBST (20 mM Tris base, 150 mM NaCl, 0.1% Tween 20), and then probed with anti-TAP (Open Biosystems), anti-H3 (Abcam), anti-H3K56Ac (Upstate), anti-actin (Abcam) or anti-myc antibodies (Roche).

3.2.3 Mass Spectrometry Analysis:

Gel bands were excised, reduced, alkylated, and digested as described previously in Chapter 2 with minor modifications. Briefly, peptide solutions were dried in a SpeedVac and stored at -20°C until the mass spectrometric analysis. LC-MS/MS was performed by dissolving the peptide samples in 5% formic acid and loading them into a 200µm x 5-cm precolumn packed in house with 5µm ReproSil-Pur C_{18}-AQ beads (Dr. Maisch HPLC GmbH) using a micro Agilent 1100 HPLC system (Agilent Technologies). The peptides were desalted on line with 95% water, 5% acetonitrile, 0.1% formic acid (v/v) for 10 min at 10 µl/min. The flow rate was then split before the precolumn to produce a flow rate of ~200 nl/min at the

column. Following their elution from the precolumn, the peptides were directed to a 75µm x 5cm analytical column packed with 5µm ReproSil-Pur C_{18}-AQ beads. The peptides were eluted using a one hour gradient (5–80% acetonitrile with 0.1% formic acid) into an LTQ linear ion trap mass spectrometer (Thermo-Electron). MS/MS spectra were acquired in a data-dependant acquisition mode that automatically selected and fragmented the five most intense peaks from each MS spectrum generated. Peak lists were generated from the MS/MS .raw file using Mascot Distiller 2.0.0.0 (Matrix Science) to produce a .mgf file with default parameters except that for each MS/MS individual peak lists were generated assuming a +2 and a +3 charge. All .mgf files from one sample were merged into a single file and then analyzed and matched to the 6298 *S. cerevisiae* protein sequences in the Saccharomyces Genome Database database (released April 2007) using the Mascot 2.1.04 database search engine (Matrix Science) with trypsin as the digestion enzyme, carbamidomethylation of cysteine as a fixed modification, and methionine oxidation as a variable modification. Peptide and MS/MS mass tolerances were set at +/- 3 and +/- 0.8 Da, respectively, with one miss-cleavage allowed and the significance threshold set to 0.01 ($p > 0.01$). Finally, an ion score cutoff of 30 was chosen to produce a false-positive rate of less than 1% in the MS data (Elias et al., 2005). A protein hit required at least two "bold red peptides," i.e. the most logical assignment of the peptide in the database selected. Furthermore, when peptides matched to more than one database entry, only the highest scoring protein was considered.

3.2.4 Raw mChIP-MS Data Curation:

Manual curation of raw protein association data generated by mChIP-MS was performed using a two step process. At first, a list of common background contaminant was generated from multiple mChIP-MS experiments from untagged yeast cells. These background proteins are highly abundant and involved in housekeeping roles such as metabolic processes and ribosomal biogenesis. This list was further supplemented by an exhaustive list of ribosomal proteins curated from the *Saccharomyces* Genome Database (www.yeastgenome.org) annotated as "structural constituent of ribosome" (GO:0003735) (Appendix 3-1). All ribosomal proteins were added to background contaminant since ribosomes are large macromolecule and as such, not all subunits were observed in the mChIP-MS untagged controls. These background proteins were removed from all raw mChIP-MS association data. Next, I applied another curation step designed to remove preys present at high frequency in the mChIP-MS association data but without relevance to chromatin biology. To do so, the number of times that a given prey was detected by mChIP-MS in the complete dataset, referred to as "mChIP abundance factor", was computed. Then, each prey which was observed in three or more mChIP-MS experiments were manually curated based on two additional criteria: molecular function and localization. Molecular functions that were targeted include protein folding, mRNA export, fatty acid biosynthesis, ribosome biogenesis and RNA processing as well as proteins located to the mitochondria and preribosome. The *Saccharomyces* Genome Database was used to

determine the molecular functions and localization of mChIP preys. In this way, 170 proteins were identified as not relevant to chromatin biology, labelled as non-specific mChIP binders and removed from the final mChIP-MS association data (Appendix 3-2).

3.2.5 Denaturing Cell Lysis for Western blot analysis:

To study the ubiquitination level of Cpr1, a quick denaturing cell lysis procedure previously described was used to minimize protein degradation (Kao and Osley, 2003). Briefly, 20mL of synthetic media lacking tryptophan (SC-TRP) was inoculated with the appropriate strains and grown to an OD_{600} of 0.15. $CuSO_4$ was then added to a final concentration of 250μM and the strains were grown for a further 3 hours at 30°C. Benomyl (10mg/mL in DMSO) or rapamycin (2nM in ethanol) were then added and the strains grown for extra 3 hours. The cells were then harvested by centrifugation at 4000 rpm for 5 minutes (4°C) and the supernatant discarded. 10mL of ice-cold trichloroacetic acid (TCA) (20% v/v) was used to resuspend the cells followed by centrifugation at 4000 rpm for 5 minutes (4°C). The pellet cells were then frozen on dry ice and stored at -80°C until ready to bed used. To lyse the cells, the cells pellets were resuspended in 200μL of 20% TCA and transferred to a 2mL bullet tube containing ~500μL of glass beads. Cells were vortexed six times for 30 seconds with one minute on ice between each pulse. The cell lysate were transferred to a fresh tube. The glass beads were washed with 500μL of 5% TCA and added to the cell lysate. Samples were left on ice for 20 minutes prior to centrifuging them for 15 minutes at 14000 rpm (4°C). The supernatant were then discarded and the protein pellet was resuspended in 300μL of 1X Laemmli sample buffer (60mM Tris, pH 6.8, 10% (v/v) SDS,

5% (v/v) fresh 2-mercaptoethanol, and 0.0025% (w/v) bromophenol blue) and 50µL of unbuffered 2M Tris to neutralize the pH of the lysate. The samples were then boiled at 95°C for 5 minutes and the samples further centrifuge at 14000 rpm for 5 minutes. The supernatants were collected and 5µL were used per Western blot analysis.

3.2.6 Chromatin Immunopurification:

Soluble chromatin was prepared from cells treated with formaldehyde and immunoprecipitated using standard procedures (Fillingham et al., 2009). Chromatin prepared from TAP-tagged strains was incubated with IgG-Sepharose (Amersham). Immunoprecipitated DNA was analyzed by semi-quantitative, multiplex PCR, always including an internal control (either a region of *ACT1* or a non-transcribed region of chromosome V, as indicated) for background. PCR reactions were separated on 6% PAGE, imaged, and in some cases quantified by using a Typhoon phosphoimager. For ChIP reactions that required quantification, the ratio of the experimental to the control signal for the precipitated DNA was divided by the ratio of the experimental to the control signal for the input DNA.

3.3 Results

3.3.1 Large-scale study of chromatin-associated proteins by mChIP-MS

My graduate work focused on better defining the interactome of DNA-binding proteins such as transcription factors. I identified 64 proteins (as of April 2010) manually curated in the *Saccharomyces* Genome Database (SGD; www.yeastgenome.org) to bind DNA (Appendix 3-3). Half of the DNA-binding proteins manually curated by SGD were transcription factors or transcriptional activators/repressors (32 out of 64). Interestingly, more than 70% of these transcription factors and transcriptional activators/repressors possess five or fewer known interaction partners previously observed by AP-MS in the BioGRID (Stark et al., 2006) (www.thebiogrid.com). Whereas, 15 out of the 64 DNA binding proteins had more than 20 protein-protein interactions reported by AP-MS (Appendix 3-3). These proteins were mostly histone proteins or members of large chromatin remodelling protein complexes which are present at high levels in the cell. As such, current AP-MS methods are viable for studying some types of DNA binding proteins (e.g. histones), but they provide little information about other classes (e.g. transcription factors).

The mChIP procedure was used to characterize the 32 known DNA binding proteins that have five or fewer reported interactions (Appendix 3-3) and other proteins with molecular functions relevant to chromatin biology. For instance, ten histone chaperones, ten lysine acetyl transferases (KAT), seven lysine methyl transferases (KMT), and seven nuclear proline isomerases were used as baits for mChIP (Appendix 3-4). The protein expression of endogenous C-terminally TAP-tagged baits (Howson et al., 2005) were first tested by Western blot. From the 130 strains tested, 110 showed the expression of a TAP-tagged bait protein at the correct molecular weight (Appendix 3-4). I subsequently used these 110 strains for large-scale mChIP purifications (Figure 3-1). The purified proteins

from each mChIP were resolved on 4-12% NuPAGE gels and the gels were silver-stained. Each lane corresponding to one mChIP experiment was sliced into 12 sections. The proteins present in the sections were in-gel digested with trypsin and were subsequently analyzed by mass spectrometry (see Materials and Methods for details). In total, 110 different TAP-tagged proteins were subjected to mChIP, and 102 of them were successfully analyzed by mass spectrometry (Figure 3-2 and Appendix 3-4).

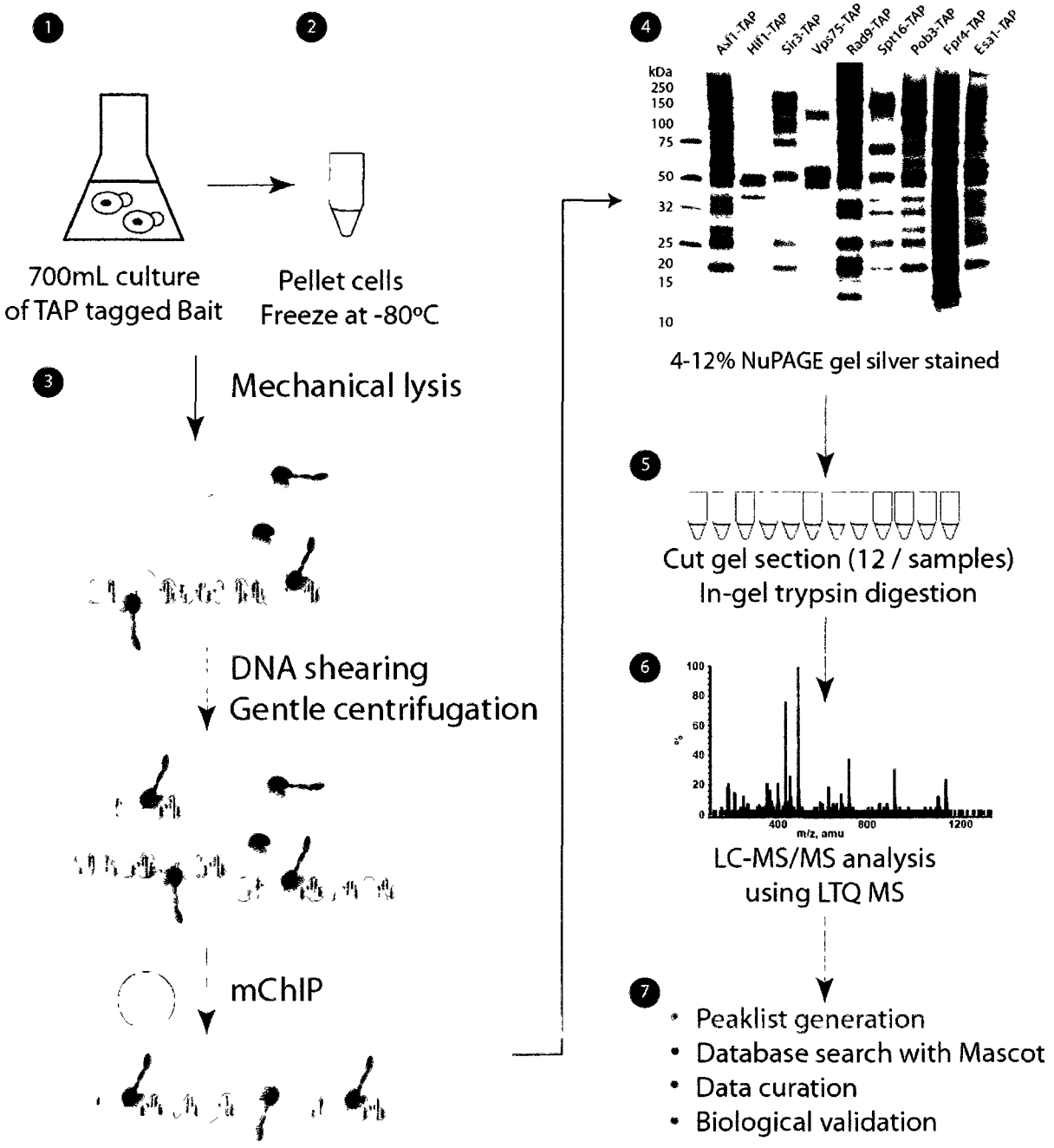


Figure 3-1: Experimental platform for the large-scale characterization of chromatin associated proteins by mChIP-MS in *Saccharomyces cerevisiae*. See *Methods* for complete details of the mChIP-MS pipeline.

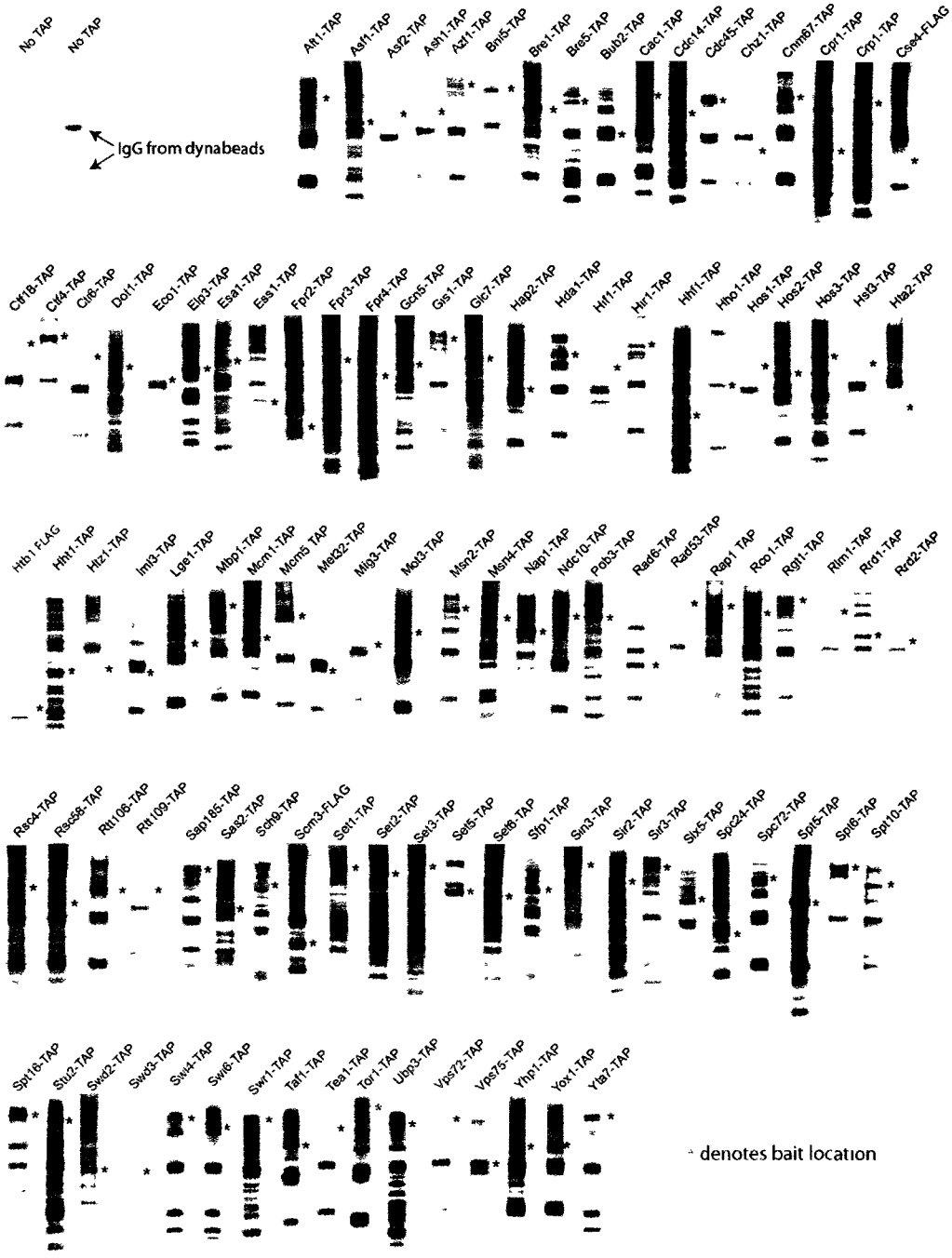


Figure 3-2: Image of all successful mChIP purification. 90% of the purified material by mChIP was resolved on 4-12% NuPAGE gradient gels with MES buffer and subsequently stained with silver or with Coomassie Brilliant Blue.

3.3.2 Curation and global analysis of mChIP-MS data

By design, the mChIP technique attempts to preserve protein-protein interactions by keeping the salt concentration in buffers and the sample centrifugation to a minimum (chapter 2; (Lambert et al., 2009)). Consequently, the mChIP of proteins globally associated with chromatin by ChIP (such as histones (Barski et al., 2007) or members of the RSC complex (Floer et al., 2010)) identified large numbers of associated proteins. Efficient data analysis is thus critical to fully appreciate the data generated by mChIP-MS. To refine the mChIP dataset, I first applied a step designed to remove common contaminants (Appendix 3-1). The list of common contaminants was compiled using control mChIP purifications and supplemented with ribosomal proteins (known common contaminants in AP-MS experiments (Gingras et al., 2007)) in the *Saccharomyces* genome database (www.yeastgenome.org). This first curation step resulted in a dataset containing 5723 protein associations among 102 unique baits and 896 distinct preys. The results were visualized as a heat map generated by hierarchical clustering of the dataset (Figure 3-4). Upon further examination of the heat map, it became clear that numerous prey proteins are detected at high frequencies in the mChIP results (vertical lines in Figure 3-4). While these high frequency preys were never observed in wild-type untagged controls, they do not appear relevant to chromatin biology and were removed from the final mChIP-MS dataset. To more systematically identify these non-specific mChIP preys, a mChIP abundance factor (the number of times a prey was identified in our mChIP screen) was determined for each prey. Examples of high abundance preys include Yra1 (54), Prp43 (50) and Vps1 (48), which have housekeeping roles not related to chromatin biology. Other

scoring algorithms have been reported, such as in (Ewing et al., 2007), to remove non-specific binders, but my dataset is not suitable for these algorithms. First, the mChIP does not identify only direct protein-protein interactions but also indirect protein associations mediated by chromatin. No previous scoring algorithms were designed to take this into account. Second, the baits studied by mChIP are functionally linked, and thus they often associate with the same preys. As such, some preys have a high mChIP abundance factor but, nonetheless, they need to be retained in the final mChIP dataset (e.g. histone chaperones co-purifying with histones). To circumvent these issues, a manual examination of the dataset was performed based on the prey's mChIP abundance factor, molecular function, and cellular localization (see methods for details). This led to the removal of 170 non-specific binders, resulting in a higher confidence mChIP dataset containing 724 prey proteins (Appendix 3-5). This refined dataset was used to generate a second heat map based on hierarchical clustering using the Pearson correlation (Figure 3-3A; Figure 3-4).

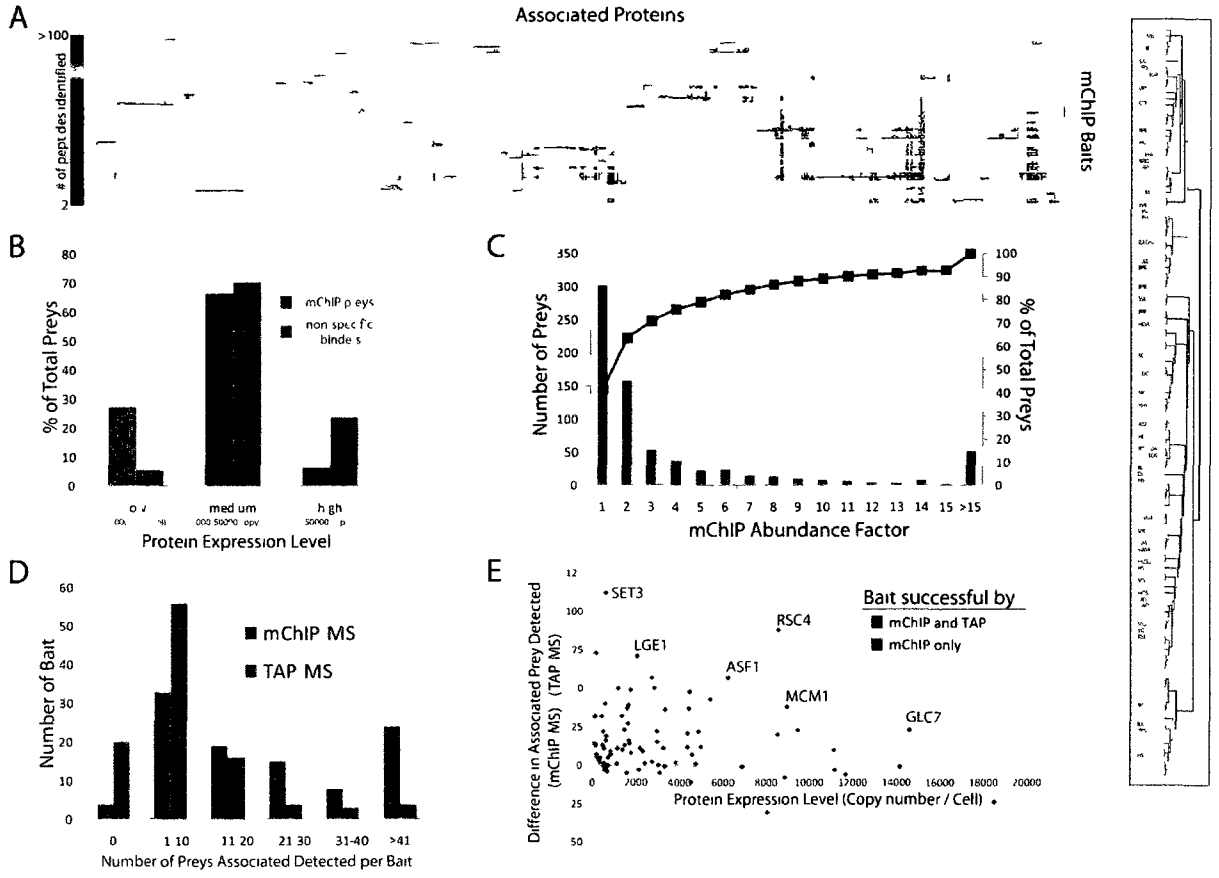


Figure 3-3: Characterization of the mChIP-MS interactome. (A) Heat map generated from the two dimensional hierarchical clustering (Pearson correlation) of 102 mChIP baits associated with 724 associated preys The color of the associated protein corresponds to the number of peptides detected by LC-MS/MS analysis following mChIP Clustering and heat map generations were performed using the Multiple Experiment Viewer (www tm4 org/mev html) **(B)** mChIP preys tend to be expressed at a lower level than non specific mChIP binders The expression levels of mChIP preys and of non-specific mChIP binders were obtained from reference (Ghaemmaghami et al , 2003) and binned in three categories (low, medium and high expression level) **(C)** Most preys are observed at low frequencies by mChIP-MS The mChIP abundance factor of each prey was determined, binned, and plotted in a bar graph (blue bars) The percentage of total preys as a function of mChIP abundance factor is also plotted (red line) **(D)** mChIP analysis increases the number of protein associations detected for the majority of the baits that were tested The data generated for 110 baits by mChIP-MS (blue bars) and TAP-MS (Gavin et al , 2006, Krogan et al , 2006) (red bars) was compiled, binned, and used to generate a bar graph **(E)** mChIP-MS improves the characterization of baits with low and high expression levels The difference in the number of protein associations between mChIP-MS and TAP-MS was determined and plotted against the bait expression levels obtained from reference (Ghaemmaghami et al , 2003)

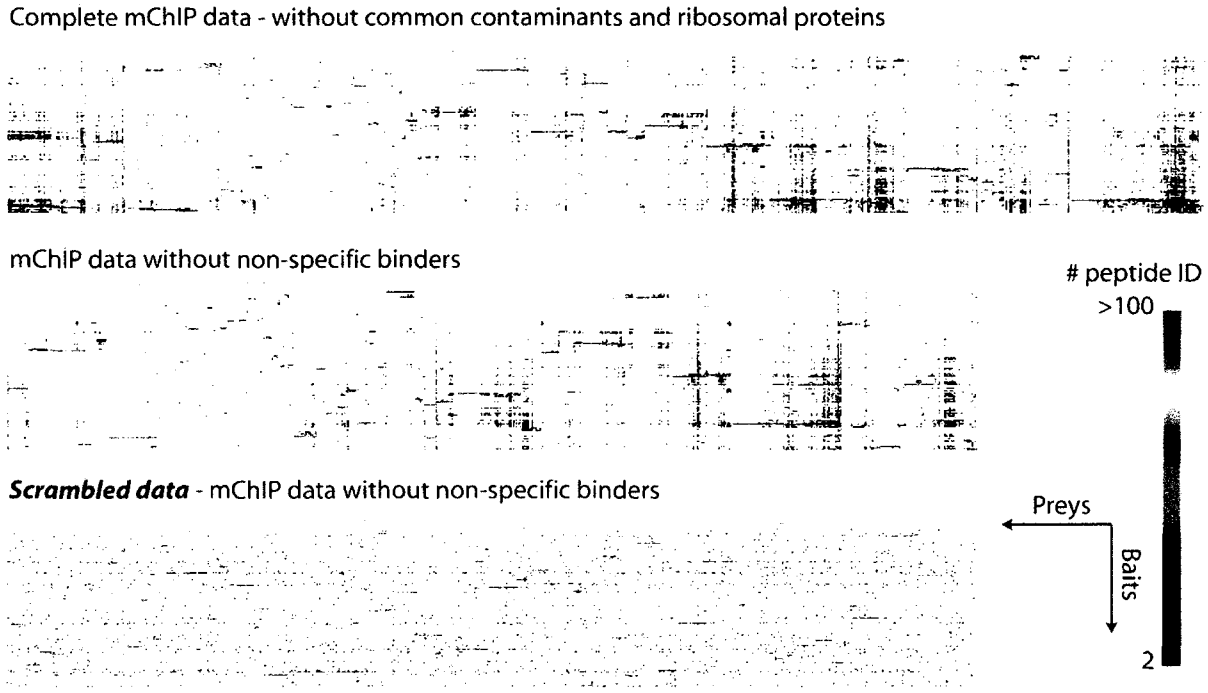


Figure 3-4: Heat map generated from mChIP data with various levels of curation. Heat maps generated from the two dimensional hierarchical clustering (Pearson correlation) of 102 mChIP baits associated with 896 preys (complete mChIP data), with 724 high-confidence preys (mChIP data without non-specific binders) or from scramble data of 724 preys. The color of the associated protein corresponds to the number of peptides detected by LC-MS/MS analysis following mChIP. Clustering and heat map generation were performed using the Multiple Experiment Viewer (www.tm4.org/mev.html).

Metrics to evaluate the quality of the final datasets were computed. A comparison of the list of non-specific binders with the dataset reveals that the non-specific binders are biased towards mid to high expression levels, whereas the mChIP preys are biased towards low to mid expression levels (Figure 3-3B). The higher expression levels of non-specific binders are consistent with the literature on AP-MS contaminants (Chen and Gingras, 2007). Furthermore, over 80% of the preys in the final dataset are each associated with less than 5 baits (Figure 3-3C). Additionally, the mChIP data was enriched for chromatin-related functions (such as chromosome segregation/division or transcriptional control), while the non-specific binders were not (Figure 3-5). Taken together, these metrics indicate that our manual removal of non-specific binders improved the overall quality of the mChIP dataset.

MIPS Functional Classification of mChIP Preys	p-values
modification by acetylation/deacetylation	<1e-14
DNA conformation modification	<1e-14
Transcription initiation	<1e-14
organization of chromosome structure	<1e-14
chromosome segregation/division	<1e-14
transcriptional control	<1e-14
general transcription activities	<1e-14
mitotic cell cycle and cell cycle control	<1e-14
DNA binding	7.616e-14
nuclear membrane	3.116e-13

MIPS Functional Classification of non-specific mChIP Binders	p-values
tRNA synthesis	<1e-14
rRNA synthesis	<1e-14
ribosome biogenesis	<1e-14
rRNA processing	<1e-14
DNA binding	<1e-14
RNA binding	<1e-14
rRNA modification	3.497e-14
extension/ polymerization activity	1.566e-08

Figure 3-5: Top Ten MIPS Functional Classification of mChIP preys and non-specific binders. FunSpec (Robinson et al., 2002) (http://funspec.med.utoronto.ca/) was used to identify the MIPS functional classification enriched in the mChIP data and non-specific mChIP binders. Only eight MIPS functional classifications were enriched in the non-specific binders.

Next, the mChIP data were compared with previously reported genome-wide TAP-MS data (Gavin et al., 2006; Krogan et al., 2006). For over 75% of the baits studied by mChIP-MS, more prey proteins were detected compared to TAP-MS. Furthermore, 18% of the baits that were successfully analyzed by mChIP-MS had previously failed by TAP-MS (Gavin et al., 2006; Krogan et al., 2006) (Figure 3-3D). Interestingly, there was no correlation between the increase in the number of associated proteins detected by mChIP and the bait expression level (Figure 3-3E). This finding suggests that the increase in the number of protein associations detected by mChIP-MS compared to those detected by TAP-MS is not mainly due to a more sensitive mass spectrometer, but rather to the purification technique itself. Overall, the budding yeast chromatin-associated interactome that is now accessible by mChIP-MS is an environment not previously investigated and worth further study.

3.3.3 mChIP improves the characterization of transcription factors

High abundance chromatin-associated proteins, such as histones and their chromatin associated protein networks, were successfully characterized by mChIP-MS. The results are consistent with the wealth of protein interaction data currently available in the literature for high abundance baits (Appendix 3-3 and Chapter 2). In this study, emphasis was placed on lower abundance targets such as transcription factors. For instance, the results from the mChIP-MS analysis of the Hap2 transcription factor (a member of the CCAAT-binding complex) was compared to traditional TAP-MS. mChIP-MS of Hap2-TAP

revealed over 80 associated proteins including Hap3 and Hap5 which are known to form a heterotrimer with Hap2 (Figure 3-6A). Previous efforts using the tandem affinity purification approach also revealed Hap2-TAP to be associated with Hap3 and Hap5 (Gavin et al., 2006; Krogan et al., 2006). Interestingly, the overexpression of Hap2-FLAG using a galactose-inducible construct followed by a one-step AP-MS analysis also produced an extensive interactome (Ho et al., 2002). A significant fraction of these associated proteins (~57%) did not possess functions related to chromatin because they are localized outside the nucleus (Figure 3-6B, Appendix 3-6). In contrast, Hap2-TAP mChIP uncovered associations with six transcription factors (Ste12, Dal81, Gln3, Stp1, Stp2 and Yap5) and with chromatin remodelling complexes (RSC, SAGA, etc). The association of Hap2 (a global regulator of carbohydrate metabolism) with Dal81 and Gln3 (two transcription regulators of nitrogen utilization pathways) suggests a broader role for Hap2 than previously reported. It should be noted that Hap2 levels might be negatively affected by glucose (Pinkham and Guarente, 1985). It is possible that some of the differences observed in the study by Ho et al. (Ho et al., 2002) are due to the carbon source used for cell growth. Still, my mChIP data suggest Hap2, Dal81, and Gln3 transcription factors may mediate crosstalk between the nitrogen utilization and non fermentable sugar utilization pathways.

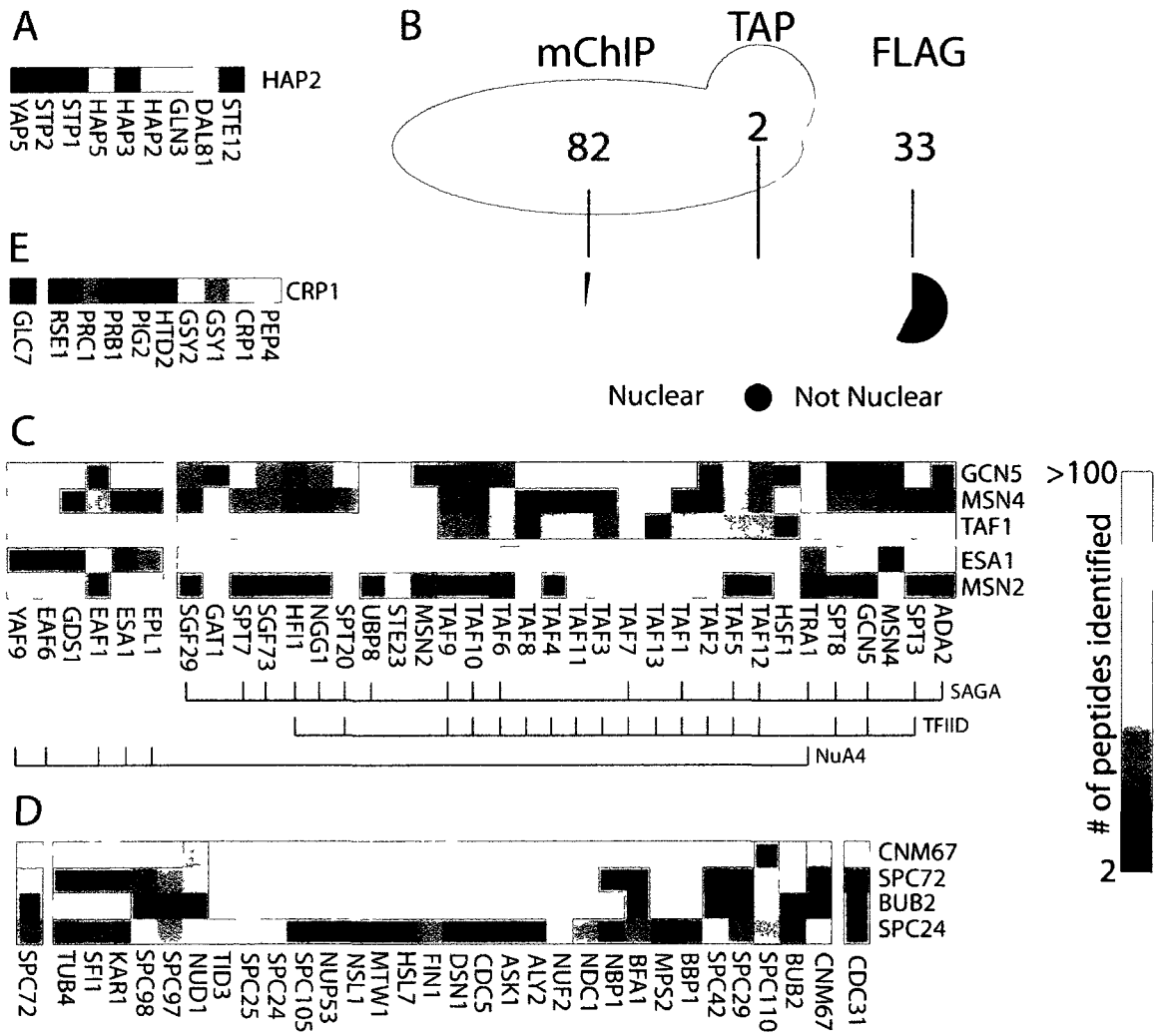


Figure 3-6: The chromatin associated interactome is a diverse landscape. (A) The Hap2 cluster contains the Hap2/Hap3/Hap5 heterotrimer and other transcriptional coactivators involved in regulating nitrogen usage. **(B)** Analysis of the Hap2 interactome by mChIP-MS, TAP-MS, and AP-MS. The overlap between the proteins associated with Hap2 by mChIP-MS, TAP-MS, and AP-MS (Supplementary Table S6) is shown. The cellular localization of Hap2 interaction partners was determined from reference (Huh et al., 2003) and displayed in pie charts. **(C)** A cluster of stress activated transcriptional activators containing the Msn2/Msn4 transcription factors and their transcriptional coactivators. SAGA, TFIID, and NuA4 protein complexes subunits are shown. **(D)** Spindle pole body cluster based on mChIP of Cnm67, Spc72, Bub2, and Spc24. **(E)** Crp1 cluster includes numerous genes involved in glycogen metabolism.

Another example of transcription factors studied by mChIP is the Msn2 and Msn4 proteins. In collaboration with the laboratory of Dr. Kristin Baetz, the transcription factor Msn4 was shown to associate with the NuA4 lysine acetyltransferase complex using traditional AP-MS (Mitchell et al., 2008). I further characterized this interaction by mChIP. First, mChIP-MS of Esa1-TAP (the catalytic subunit of the NuA4 complex) was performed and, as expected, it resulted in the co-purification of Msn4 (Figure 3-6C). Second, reciprocal mChIP of Msn4-TAP and the related Msn2-TAP resulted in the co-purification of NuA4 subunits (Figure 3-6C). Moreover, members of both the SAGA and TFIID complexes were associated with Msn2 and Msn4, which suggests that numerous transcriptional coactivators participate in Msn2 and Msn4 functions (Figure 3-6C). Based on the spectral count data, described in (Liu et al., 2004), it appears that Msn2 preferably associates with protein complexes that contain the Gcn5 rather than Esa1 KAT (Figure 3-6C). Conversely, Msn4 does not show this bias in its association with transcriptional coactivators (Figure 3-6C). Interestingly, the heat shock transcription factor Hsf1 was also detected in the Gcn5 KAT mChIP but absent from the Esa1 purification (Figure 3-6C). Association between heat shock transcription factors and KATs thus appear common in budding yeast.

mChIP-MS analyses of Msn2 and Msn4 also identified proteins uniquely associated with each of the transcription factors. For instance, Gds1, a protein of unknown function was associated with Msn4 and Esa1 with 5 and 6 peptide, respectively, but not with Msn2 (Figure 3-6C). On the other hand, Ste23 was the top MS hit in the Msn2 mChIP, but it was not detected with Msn4. Ste23 is a metalloprotease which is an ortholog of the mammalian insulin-degrading enzyme (Alper et al., 2009). Ste23 was also shown to catalyze the

cleavage of a peptide sequence corresponding to pro-α-factor *in vitro* (Alper et al., 2009). Furthermore, an additional link between Ste23 and Msn2 lies in the presence of a stress response element (STRE) upstream from the *STE23* gene (Treger et al., 1998a). STRE are often bound by the Msn2 and Msn4 transcription factors, and STRE-controlled genes are induced following heat shock (Treger et al., 1998b). Heat shock proteins, many of which possess STRE, are required for proper α-factor processing (Meacham et al., 1999). Based on this data, I postulate that Ste23 plays a role in proper stress responses in budding yeast.

3.3.4 mChIP facilitates the characterization of transcription factors that regulate the cell cycle

Coordinated gene expression is essential for maintaining cellular fitness (Zhou et al., 2009). In budding yeast, numerous transcription factors are critically involved in regulating the expression of multiple genes within precise parts of the cell-cycle (Wittenberg and Reed, 2005). In *S. cerevisiae*, the cell-cycle transition from G_1 to S begins with START, a coordinate transcriptional program resulting in the timed expression of hundreds of genes. Two protein complexes essential for this process are the MBF and SBF transcription factors composed of Swi4-Mbp1 and Swi4-Swi6, respectively (Moll et al., 1992). Previous AP-MS studies of MBF and SBF revealed interaction partners such as Whi5, Nrm1, and Msa1 with known roles in cell-cycle regulation. mChIP-MS analyses of Swi4-TAP, Swi6-TAP, and Mbp1-TAP successfully identified known interaction partners (such as Stb1) for both MBF and SBF which had not been previously identified by AP-MS methods (Figure 3-7).

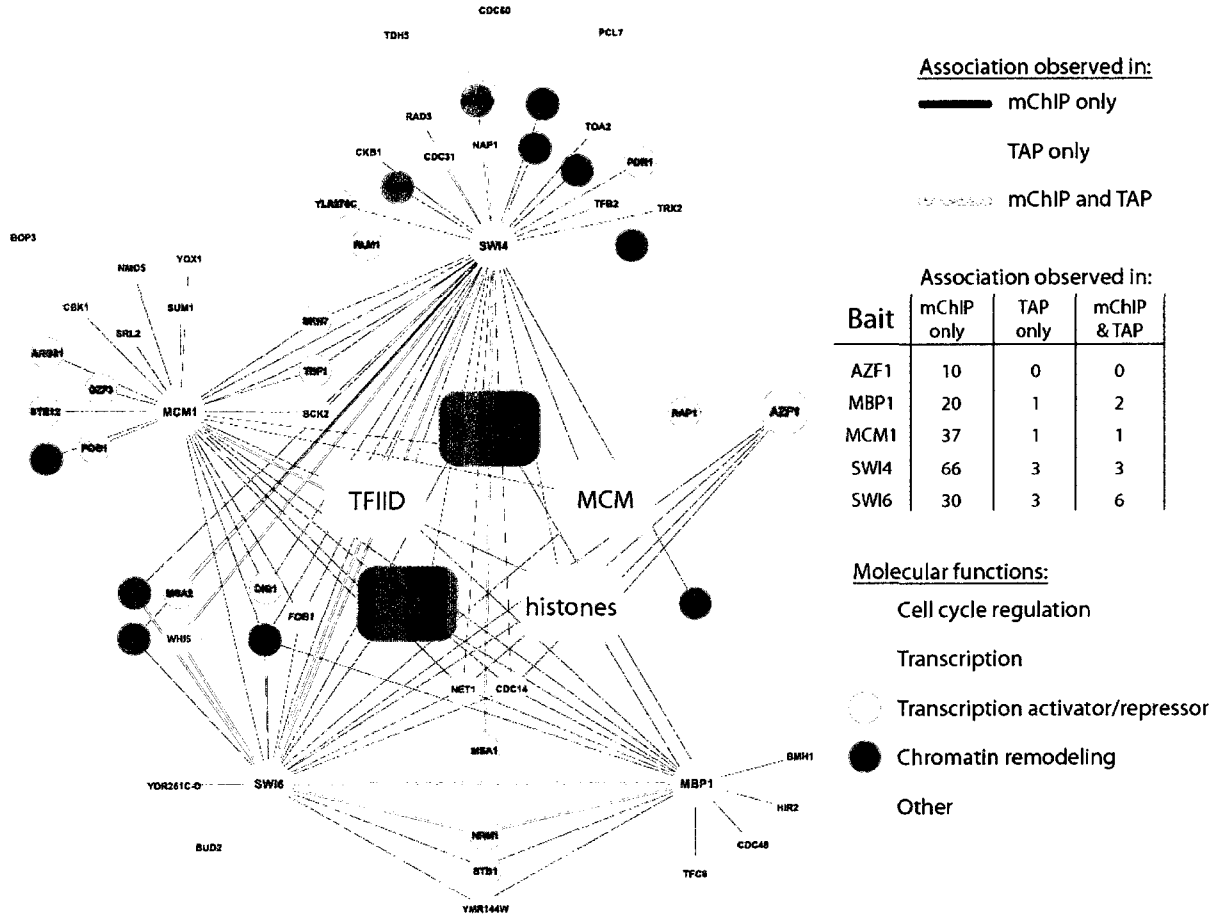


Figure 3-7: Five cell-cycle regulators are observed to be part of a dense network of chromatin associated proteins. Cytoscape (Shannon et al., 2003) was used to visualize the network of preys associated with five transcription factors (Azf1, Mbp1, Mcm1, Swi4, and Swi6) by TAP-MS (Gavin et al., 2006; Krogan et al., 2006), mChIP-MS (this study), or both. The large nodes represent the bait used for AP-MS experiments; whereas, the small nodes are the associated preys. In addition, some well characterized large complexes were manually collapsed into rectangles to simplify the network. The edge color corresponds to the source of the association, and the node color corresponds to protein molecular function curated from the *Saccharomyces* genome database.

Interestingly, the networks for the transcription factors Azf1 and Mcm1 showed an interconnection with the Swi4, Swi6 and Mbp1 networks (Figure 3-7). In fact, associations between Azf1-TAP and Swi6, as well as associations between Mcm1-TAP and Mbp1, Swi4, and Swi6 were detected by mChIP-MS (Figure 3-7). Mcm1 is a transcription factor that participates in the regulation of multiple genes depending on its associated proteins (Ferrezuelo et al., 2009). For instance, when Mcm1 interacts with Ste12, it participates in regulating the mating-specific genes (Errede and Ammerer, 1989), whereas association with Yox1 or Yhp1 leads to the regulation of genes expressed in the M to G_1 transition (Pramila et al., 2002). The mChIP data for Mcm1-TAP shows a wide array of associated proteins involved in properly regulating the cell cycle (e.g. Sum1) as well as transcriptional activators such as Gzf3 and Pog1 (Figure 3-7). Furthermore, Mcm1-TAP was found to associate with Bck2 and Ste12 by mChIP-MS. Bck2, which is known to activate numerous cell-cycle regulated genes (Ferrezuelo et al., 2009), was previously shown to be affected in strains lacking *ste12* or *mcm1,* thus indicating a common function (Ferrezuelo et al., 2009). The fact that Mcm1-TAP co-purified with both Ste12 and Bck2 by mChIP-MS supports a direct interplay between these transcription factors at specific promoters. Overall, I successfully purified several networks of transcription factors involved in cell-cycle regulation using the novel mChIP approach.

3.3.5 mChIP uncovers novel roles for the peptidyl proline isomerase CPR1

As part of my proteomic screen, many nuclear peptidyl proline isomerises, enzymes that catalyze conformational changes of proline residue (Lu et al., 2007), were studied. Seven nuclear peptidyl proline isomerases, including Cpr1, a known member of the Set3 complex (Pijnappel et al., 2001), were successfully analyzed by mChIP-MS. In particular, mChIP-MS of Cpr1-TAP revealed a large number of associated proteins including members of the Set3 complex (Figure 3-8A). In addition, all members of the TORC1 complex and some members of TORC2 were found. Moreover, numerous components of the spindle pole body, as well as proteins with spindle-related functions, were found with Cpr1. These finding suggest that Cpr1 possesses wider functions than previously thought, especially with regards to regulating cellular growth (Figure 3-8A). Surprisingly, the E3 ubiquitin ligase Bre1 and its interaction partner Lge1 were found to be associated with Cpr1 (Figure 3-8A). This association raises the possibility that Cpr1 is ubiquitinated by Bre1, which is supported by the presence of higher molecular weight bands in a Western blot for the TAP tag of the mChIP material, albeit at a low level (Figure 3-8B).

To study this possibility with greater sensitivity, Cpr1-TAP strains containing a plasmid encoding myc-tagged ubiquitin under the control of the copper inducible *CUP1* promoter were prepared. Following induction with $CuSO_4$, myc-tagged ubiquitin was expressed at high levels to facilitate the detection of ubiquitinated proteins. Using this strategy, Cpr1-TAP was observed to be ubiquitinated in a mid-log phase culture. Furthermore, the extent of Cpr1 ubiquitination was increased following a treatment with

rapamycin (a TORC1 inhibitor) or benomyl (a microtubule destabilizing agent) (Figure 3-8C), whereas, global ubiquitination levels were not increased (Figure 3-9). These higher molecular weight bands were abolished when a mutant ubiquitin K48R G76A protein that is incapable of forming poly-ubiquitin chains was expressed (Figure 3-8D). Therefore, these higher molecular weight bands were confirmed to be poly-ubiquitinated forms of Cpr1. Moreover, Cpr1 ubiquitination appears to be modulated in response to the two drug treatments, which suggests roles for Cpr1 in nutrient sensing and cell-cycle regulation. The role of Bre1 and Lge1 in Cpr1 ubiquitination is unknown but currently under study in the Figeys laboratory.

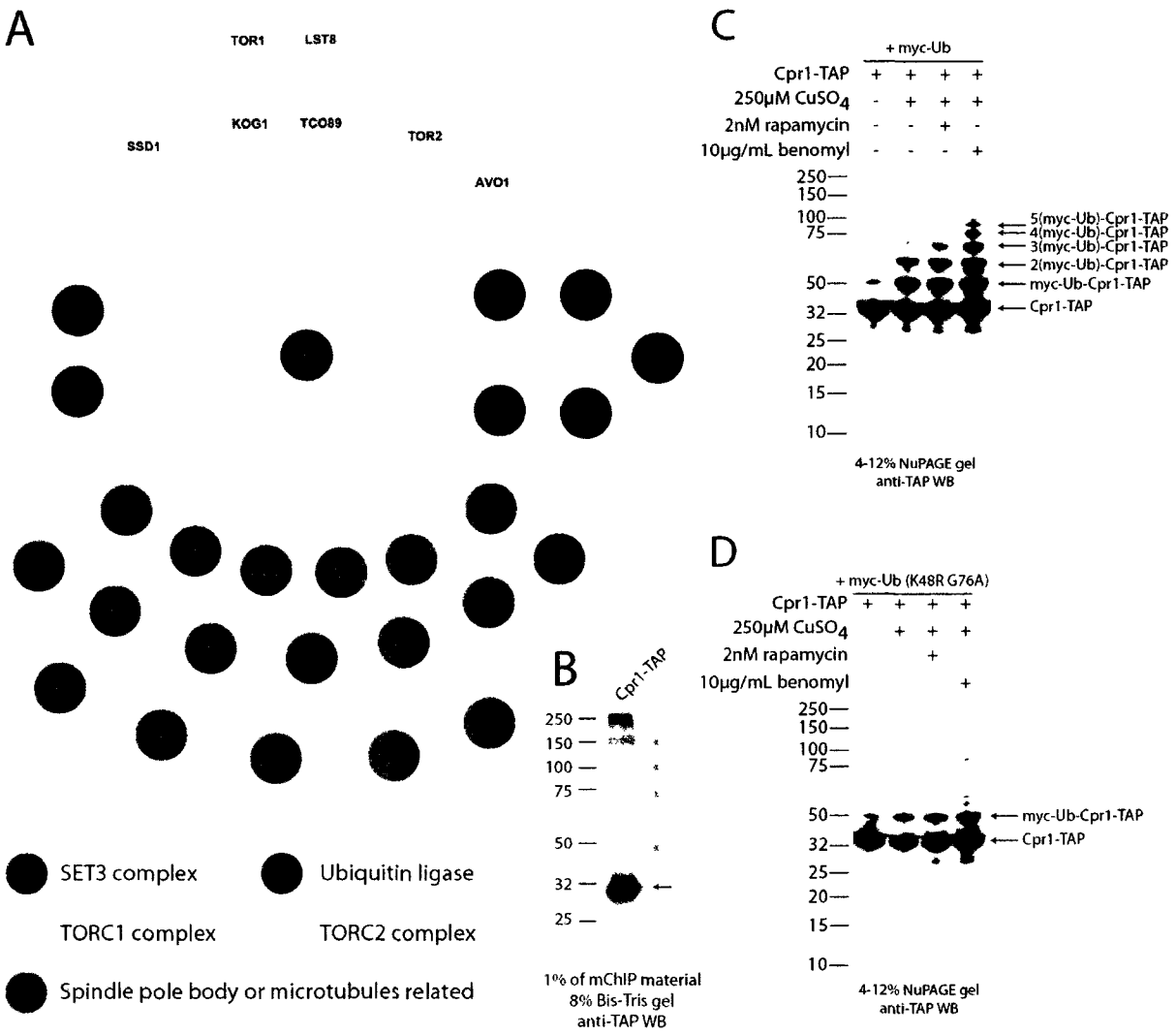


Figure 3-8: Cpr1 is a protein hub associated with nutrient sensing, cell division, and silencing. (A) Proteins associated with Cpr1 were visualized using Cytoscape and arranged based on their known molecular functions manually. **(B)** Approximately 1% of the Cpr1-TAP mChIP purified sample was resolved on an 8% Bis-Tris gel, transferred onto nitrocellulose, and immunoblotted against the TAP tagged. The arrow indicates the location of Cpr1-TAP and the asterisks mark potential ubiquitinated Cpr1-TAP. **(C)** Cpr1 is ubiquitinated in response to cellular stress. TCA extracted WCL from the indicated samples were resolved on a 4-12% NuPAGE gradient gel with MES buffer, transferred onto nitrocellulose, and immunoblotted for the TAP tag. The higher molecular bands observed for Cpr1 correspond to mono- and poly-ubiquitinated Cpr1 as marked and are not detected in a strain which cannot form polyubiquitin chains **(D)**.

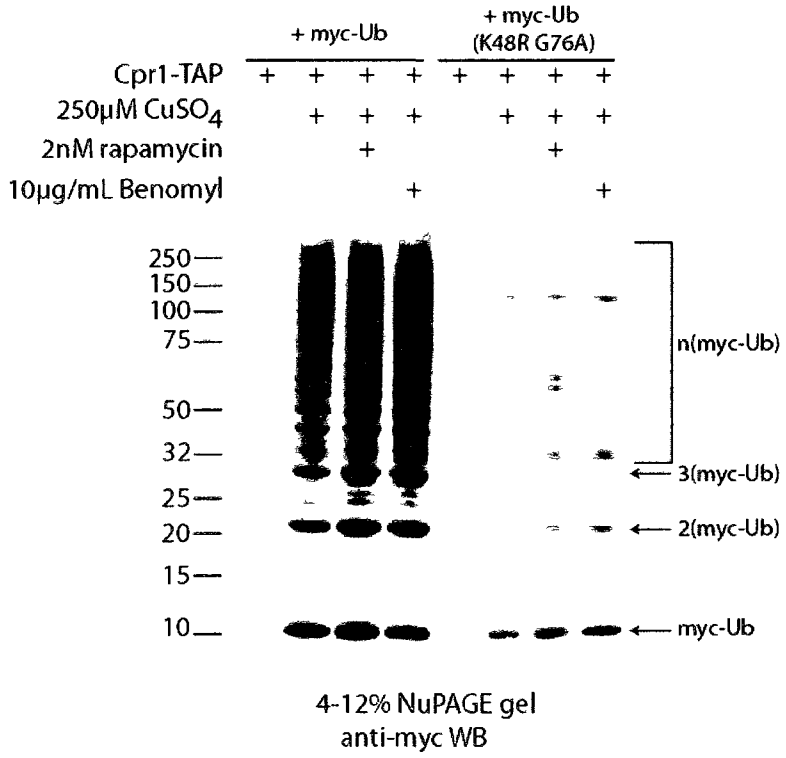


Figure 3-9: Ubiquitination levels are not globally increased following rapamycin or benomyl treatment. TCA extracted WCL from the indicated samples were resolved on a 4-12% NuPAGE gradient gel with MES buffer, transferred onto nitrocellulose and immunoblotted for myc-Ub. Higher molecular bands observed corresponds to mono- and poly-ubiquitinated moieties as marked and are not detected in a strain which cannot form polyubiquitin chains.

3.3.6 Dissection of physical interplay among histone H3/H4 chaperones

I previously used mChIP to show that the histone H3/H4 chaperone Rtt106 associates with two other histone chaperone complexes, HIR and CAF-1 (Fillingham et al., 2009). Because HIR and CAF-1 are both known to interact with Asf1 (Sharp et al., 2001; Sutton et al., 2001), mChIP was used to further characterize the chromatin associated protein networks of Hir1-TAP, Rtt106-TAP, Asf1-TAP, and Cac1-TAP (Figure 3-10A). MS analysis of these four baits revealed that HIR, Rtt106, and Asf1 are associated, whereas Rtt106 and CAF-1 compose another well characterized complex (Huang et al., 2005; Huang et al., 2007; Li et al., 2008). The association between Rtt106, HIR, and Asf1 was further dissected by testing whether Asf1 was required for Rtt106 association with HIR. Rtt106-TAP mChIP followed by Western blotting (mChIP-WB) for Hir1-myc showed a strong association which was abolished in the absence of *asf1* (Figure 3-10B). One of my collaborators previously demonstrated that Hir1 binding to the *HTA1-HTB1* promoter is not affected by deleting *asf1* or *rtt106*, whereas Rtt106 binding to the same promoter requires both Hir1 and Asf1 (Fillingham et al., 2009). Taken together, these finding suggest a central role for Asf1 in the association among Rtt106, HIR, and Asf1. I thus focused on Asf1 to further unravel the physical associations among these histone chaperones.

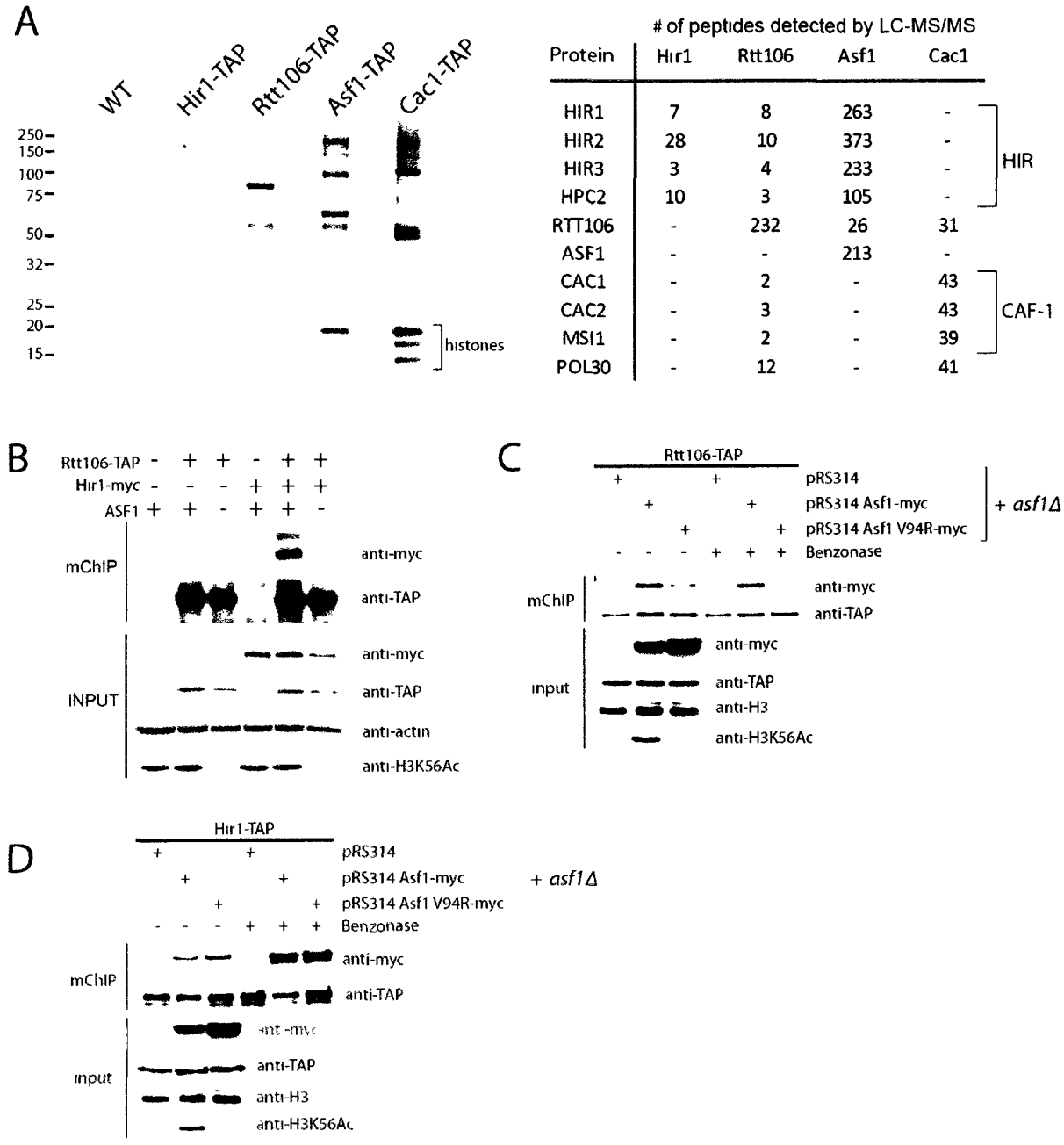


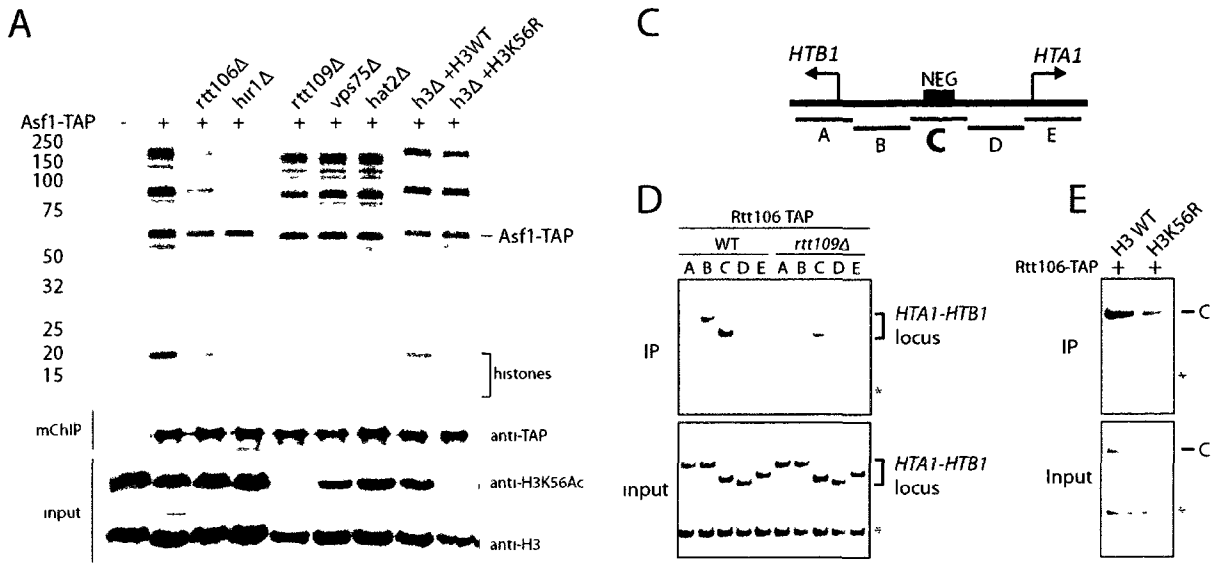
Figure 3-10: Dissection of the physical associations between histone H3/H4 chaperones by mChIP. (A) Rtt106, HIR, and Asf1 are physically associated. mChIP purifications from the indicated strains were performed as per *materials and methods*, and the purified proteins were resolved on a 4-12% NuPAGE gradient gel with MES buffer and stained with Colloidal Blue. Each sample was in-gel digested with trypsin and analyzed by LC-MS/MS. The number of peptides detected by LC-MS/MS in each mChIP sample for HIR, Rtt106, Asf1, CAF-1, and Pol30 are reported. **(B)** The Rtt106-HIR association requires ASF1. Immunoblotting against the TAP, myc tag, actin, or H3K56Ac were performed from the appropriate mChIP purifications or from WCL of the indicated strains following 8% SDS-PAGE gel (TAP, myc and actin WB) or 15% SDS-PAGE gel (H3K56Ac WB). **(C)** Rtt106 association with Asf1 is dependent on H3/H4 dimers but are independent of the presence of DNA. The appropriate strains were grown in SC–TRP media and subjected to mChIP in the presence or

absence of benzonase. The purified materials or WCL were resolved on 8% or 15% SDS-PAGE gel and immunoblotted against TAP, myc, H3, and H3K56Ac as indicated. **(D)** Hir1 association with Asf1 is independent of the presence of histones or DNA. Hir1-TAP mChIP were performed as per Figure 3-10C.

To directly probe the association between Rtt106 and Asf1, Rtt106-TAP mChIP-WB experiments were performed from strains containing a myc-tagged version of wild-type Asf1 or the Asf1 V94R mutant (Figure 3-10C). The V94R mutation was previously shown to cause a greatly reduced affinity for histone H3/H4 (Mousson et al., 2005) and, therefore, it is a good tool for defining the role of histones H3/H4 in these associations. The Rtt106-Asf1 association was found to be significantly reduced in the V94R mutant compared to the wild-type Asf1 (Figure 3-10C). This suggests that the ability of Asf1 to bind histone H3/H4 is critical for efficient interaction with Rtt106. Another alternative is that the association between Rtt106 and Asf1 is dependent on the presence of chromatin and thus is reduced in the V94R mutant. To directly test this alternative, mChIP-WB experiments were performed in the presence of benzonase, a promiscuous endonuclease that digests both DNA and RNA (Figure 3-10C). In the absence of DNA, wild-type Asf1 was co-purified with Rtt106-TAP, but the V94R mutant was not detected (Figure 3-10C). This suggests that Asf1 and Rtt106 interact through histone H3/H4. This indirect association between Asf1 and Rtt106 is consistent with the lack of interaction observed between recombinant Asf1 and Rtt106 in *in vitro* binding assays (Huang et al., 2005). On the other hand, the well documented interaction between the HIR complex and Asf1 (Sharp et al., 2001; Sutton et al., 2001) is not affected in the V94R point mutant or in the absence of DNA (Figure 3-10D).

The nucleosome assembly factor Asf1 has been extensively studied by AP-MS and possesses well-defined interaction partners such as the HIR complex (Green et al., 2005), Rad53 (Emili et al., 2001; Hu et al., 2001), and the histones H3/H4 (Munakata et al., 2000). Further, mChIP-MS experiments of Asf1-TAP successfully identified these known interaction

partners and also revealed an extended network of proteins associated with Asf1 such as transcription factors (Pdr1 and Pho2), proteins involved in DNA replication (Sld3, Fob1) and Mnr2, a putative magnesium transporter (Appendix 3-7). I next tested how this network of associated proteins was affected by the absence of genes previously linked to Asf1 (Figure 3-11A). Lack of *hir1* resulted in a drastic reduction of Asf1's network of associated proteins, including the loss of the HIR complex, Rtt106, and the transcription factors Pdr1 and Pho2 (Appendix 3-7). In contrast, deletion of *rtt106* appeared only to have a marginal impact on the proteins associated with Asf1-TAP by mChIP-MS (Appendix 3-7). These finding are consistent with our view that HIR functions upstream of Asf1 and Rtt106, whereas Rtt106 functions downstream of both HIR and Asf1 (Fillingham et al., 2009).



B

# of peptides detected by LC-MS/MS

	WT	*rtt106Δ*	*hir1Δ*	*rtt109Δ*	*vps75Δ*	*hat2Δ*	*h3Δ* +H3 WT	*h3Δ* +H3K56R	
Asf1	131	162	196	69	60	72	119	127	
Hir1	189	173	-	66	91	116	217	268	HIR
Hir2	276	300	-	183	207	259	373	536	
Hir3	188	168	-	128	150	192	279	148	
Hpc2	116	111	-	62	94	99	210	326	
Rad53	59	47	94	25	34	45	57	51	
Rtt106	39	-	-	**7**	31	33	41	**7**	

Figure 3-11: H3K56Ac influence Rtt106 interactions with Asf1 and Rtt106 binding to the *HTA1-HTB1* promoter. The Asf1-Rtt106 association is reduced in strains lacking H3K56Ac. **(A)** mChIP purifications of Asf1-TAP from the indicated background were performed, and 90% of the purified proteins were resolved on a 4-12% NuPAGE gel with MES buffer and then stained with Colloidal Blue. ~1% of the purified proteins or WCL were resolved on a 15% SDS-PAGE gel and immunoblotted for TAP, H3, or H3K56Ac. **(B)** The number of peptides detected by LC-MS/MS for Asf1, Rtt106, and HIR subunits are shown for each sample mChIP in Figure 3-11A. See Appendix 3-7 for complete MS data. **(C)** Schematics of the *HTA1-HTB1* divergent promoter and of the five PCR primer pairs used in this study. **(D)** Rtt106 binding to the *HTA1-HTB1* promoter is influenced by the H3K56Ac mark. Rtt106-TAP ChIP from a WT or *rtt109* background was performed as per *material and methods.* Top bands correspond to a region of the *HTA1-HTB1* locus; whereas, the bottom bands are from an untranscribed region of chromosome V. **(E)** Rtt106 binding to region "C" of the *HTA1-HTB1* promoter is reduced but not abrogated in the absence of the H3K56Ac mark. Rtt106 ChIP from the indicated strains was performed as per Figure 3-11D.

Asf1 was previously shown to be required for the acetylation of lysine 56 of histone H3 (H3K56Ac) (Recht et al., 2006). Tests were performed to define how this histone mark affects the network of associated proteins with Asf1. To do so, mChIP-MS purifications of Asf1-TAP in strains lacking *RTT109* (the KAT responsible for H3K56Ac) were performed (Figure 3-11A). In this background, a slight reduction in the Asf1-associated protein network was observed (Appendix 3-7). Interestingly, the number of Rtt106 peptides sequenced by MS (an indication of protein concentration) was significantly reduced in the *rtt109Δ* background. This reduction was not observed in a strain lacking *vps75* (Figure 3-11B). Vps75 is a chaperone previously shown to stabilize Rtt109 but not known to affect the levels of H3K56Ac (Fillingham et al., 2008). This observation points to an important role for H3K56Ac in the interaction between Rtt106 and Asf1. The mChIP-MS of Asf1-TAP in a strain where all histone H3 contained the K56R mutation also exhibited a lower number of Rtt106 peptides (Figure 3-11B). This supports the notion of a reduced association between Asf1 and Rtt106 in the absence of H3K56Ac. Previous work has shown that H3K56Ac (catalyzed by Rtt109) greatly increases the affinity of Rtt106 for H3-H4 and promotes Rtt106-based replication-coupled nucleosome assembly (Li et al., 2008). In addition, my collaborators have demonstrated that Rtt106 binds to the *HTA1-HTB1* divergent promoter and enables proper replication-independent nucleosome assembly (Fillingham et al., 2009). Using chromatin immunoprecipitation (ChIP), my collaborators tested whether H3K56Ac affected Rtt106 binding to the *HTA1-HTB1* promoter (Figure 3-711C). As expected, ChIP experiments revealed that Rtt106 binding to this promoter is reduced in the absence of rtt109, the enzyme responsible for the H3K56Ac mark (Figure 3-11D). This reduced binding of Rtt106

to the *HTA1-HTB1* promoter is also observed in a H3K56R strain background (Figure 3-11E). Taken together, these pieces of data support a model where Rtt106 interact with chromatin via Asf1/HIR in most cases (Figure 3-12). Moreover, the association with chromatin is more prominent when histone H3/H4 is previously acetylated at K56 which results in the proper assembly of nucleosomes at the *HTA1-HTB1* promoter.

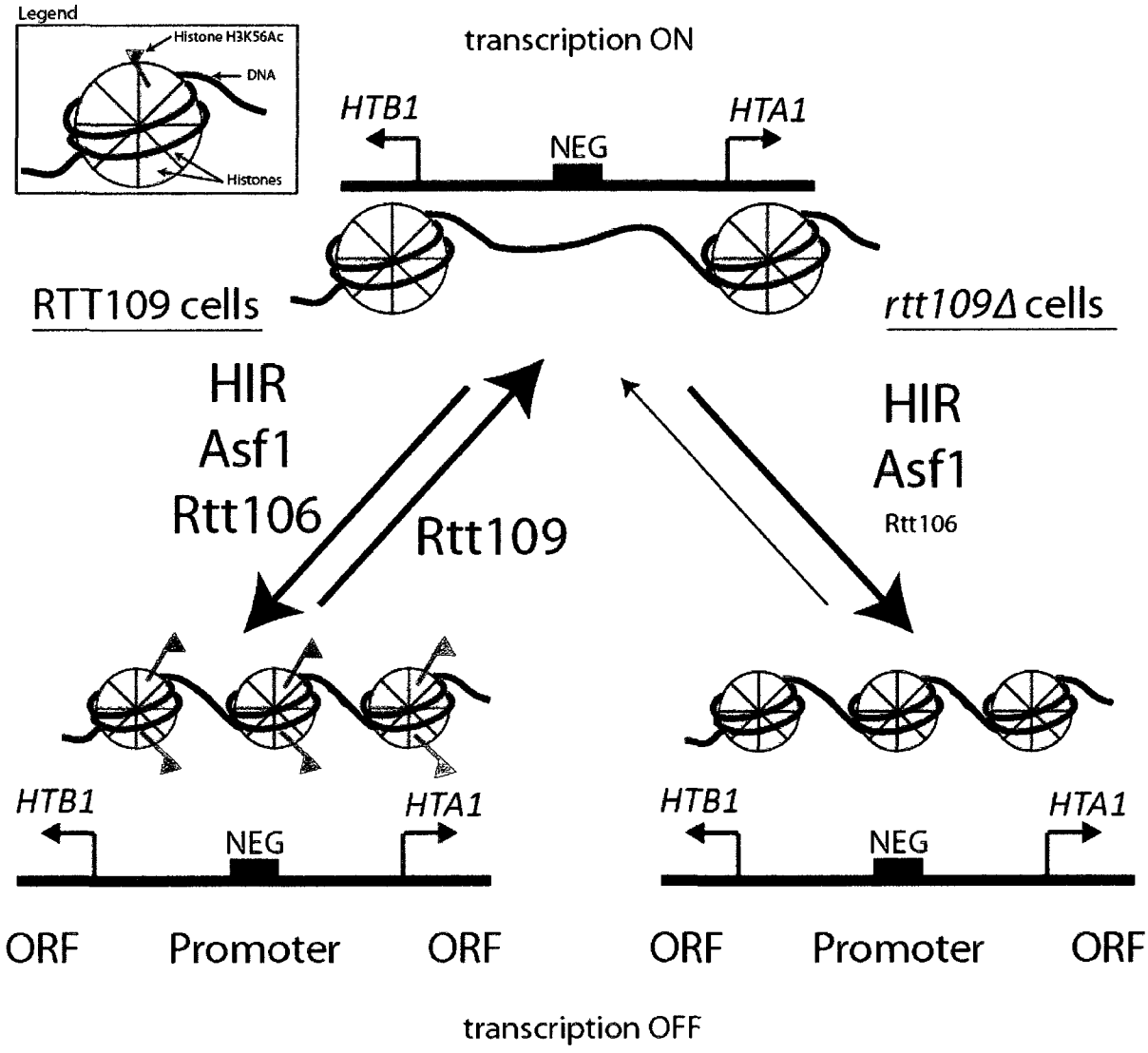


Figure 3-12: Model of histone chaperone activity at the *HTA1-HTB1* promoter. In WT cells, the HIR, Asf1 and Rtt106 histone chaperone promotes nucleosome deposition at the *HTA1-HTB1* promoter. In S phase, the Rtt109 KAT acetylate histone H3 at K56, increasing Rtt106 affinity for histone H3 and promoting nucleosome disassembly and transcription at the *HTA1-HTB1* promoter. In the absence of RTT109, there is no H3K56Ac resulting in a reduced binding of Rtt106 to the *HTA1-HTB1* promoter. Consequently, nucleosome disassembly during S phase is reduced increasing the repression of the *HTA1-HTB1* genes.

3.3.7 mChIP for other proteins

The solubility and stability of protein complexes can be a problem in AP-MS experiments. The spindle pole body is a very large macromolecule between 300 to 500 MDa and acts as the only microtubule organizing center in budding yeast (reviewed in (Jaspersen and Winey, 2004)). Unsurprisingly, such a macromolecule is refractory to common AP-MS protocols, forcing traditional biochemical approaches to be employed. For instance, the classical purification of the spindle pole body by gradient centrifugation followed by MS analysis revealed most of the components of this large organelle. Unfortunately, this technique is not suitable for large-scale studies because it require extensive manipulations and 40 litres of yeast culture (Wigge et al., 1998). As part of my work, several proteins associated with the spindle pole body were used as bait for mChIP-MS. They were successfully analyzed without the need for further optimization of the method (Figure 3-6D). Interestingly, some proteins not previously linked to the spindle pole body (for example the putative lysine methyltransferase Set5 or the poorly characterized peptidyl proline isomerase Fpr2) were found to co-purify with spindle pole body components by mChIP (supplementary Table S4). Numerous spindle pole body components are phosphorylated and those post-translational modifications are essential for proper spindle pole body function (Donaldson and Kilmartin, 1996; Stirling and Stark, 1996). The mChIP data obtained for the spindle pole body raises the possibility that other post-translation modifications such as lysine methylation might also be critical for this organelle's function.

I aimed to characterize chromatin associated proteins that were poorly studied by AP-MS. In doing so, however, I was also successful in studying proteins that possess other functions. One such example is Crp1, a poorly characterized nuclear protein (Huh et al., 2003) reported to bind cruciform DNA (Rass and Kemper, 2002). The known interaction partners of Crp1 include Pep4 and Prc1 which are two proteins involved in vacuolar degradation (Van Den Hazel et al., 1996). mChIP-MS of Crp1-TAP successfully detected Pep4 and Prc1. In addition, mChIP-MS identified another vacuolar proteinase Prb1, the glycogen synthases Gsy1 and Gsy2, as well as the phosphatase Glc7 and its targeting subunit Pig2 (Figure 3-6E). Surprisingly, six proteins associated with Crp1-TAP (Glc7, Pep4, Gsy2, Pig2, Htd2 and Prb1) are required for proper glycogen accumulation (Francois and Parrou, 2001) which suggests a critical role of Crp1 in this process. Interestingly, Crp1 is an ortholog to the mammalian AMP-activated protein kinase beta 2 subunit, which is known to directly bind glycogen and coordinate cellular metabolism in response to energy demands (Polekhina et al., 2003). Mutations in the AMPK genes in human have been reported to result in improper glycogen accumulation and numerous diseases (Arad et al., 2002). Although the exact role of Crp1 in the glycogen synthesis pathway is still undefined, my results clearly reinforce the need to further study this gene.

3.4 Discussion:

In this chapter, I have reported the characterization of the protein interactomes of 102 chromatin associated proteins. This was done using the mChIP-MS procedure, which I

developed to facilitate the purification of chromatin-bound protein networks (Chapter 2; (Lambert et al., 2009)). The application of mChIP-MS to these baits resulted in a substantial increase in the number of nodes in the network compared to conventional approaches (Figure 3-3). Many transcription factors notoriously difficult to study by conventional AP-MS methods were successfully analyzed by mChIP-MS (Appendix 3-3). One example is the study using mChIP of cell-cycle regulators involved in START (Figure 3-7). I was able to recapitulate the majority of the protein-protein interactions discovered over the last 10 years for the cell cycle transcription factors SBF and MBF as well as to considerably expand the network. For instance, the previously hypothesized association between Mcm1 and Bck2 (Ferrezuelo et al., 2009) was confirmed with the detection of 4 unique peptides for Bck2 after Mcm1 mChIP-MS analysis (Appendix 3-5). Physical interactions between transcription factors are recognized as a critical component of their regulation (Walhout, 2006). The ability of the mChIP-MS approach to identify these lower abundance interactions can be attributed to a reduction in sample loss as a consequence of maintaining chromatin in solution, a reduction in the number of processing steps as a consequence of using an efficient single-step affinity purification, and, finally, a fourfold reduction in the mass of cells required per purification. Therefore, the mChIP procedure has proven to be an efficient high-throughput method for studying numerous types of baits associated with chromatin.

By design, my mChIP approach enables the identification of both direct and indirect (mediated via chromatin) protein-protein interactions (Lambert et al., 2009). My final dataset contains both direct and indirect protein associations, which produce a more

holistic view of these bait interactomes. For instance, extensive literature links the process of histone H2B ubiquitination (requiring the action of Rad6, Bre1 and Lge1 (Hwang et al., 2003)) to the trimethylation of histone H3 on lysine 4 (H3K4) (performed by the Set1-containing COMPASS complex (Wood et al., 2003)). In particular, ubiquitination of histone H2B on lysine 123 was observed only when the E2 ubiquitin ligase Rad6, the E3 ubiquitin ligase Bre1 and their interaction partner, Lge1 (Hwang et al., 2003), were present. Deleting one of these factors resulted in the abrogation of both histone H2B ubiquitination and H3K4 trimethylation (Hwang et al., 2003; Wood et al., 2003). Recent work reported that Swd2, a subunit of the COMPASS complex, is recruited to chromatin in a manner that requires histone H2B ubiquitination (Lu et al., 2007), which suggests a direct physical link between Rad6/Bre1/Lge1, histone H2B ubiquitination and the COMPASS complex. My mChIP-MS analysis of Bre1 and Lge1 identified COMPASS components, whereas the reciprocal mChIP-MS of Set1 and Swd3 (two COMPASS components) identified Bre1 and Lge1 (Appendix 3-5). These physical associations are in accordance with the known links between histone H2B ubiquitination and H3K4 trimethylation and show that the study of large macromolecular complexes containing both direct and indirect associations can be very informative. This is especially relevant in light of recent work which challenged the classical linear view of chromatin architecture in favour of three dimensional models containing numerous intra- and inter-chromosomal interactions (Fraser, 2006; Schoenfelder et al., 2010a). For instance, the oestrogen receptor has been recently shown to cause extensive chromatin looping to bring together gene enhancers and their transcription start sites (Fullwood et al., 2009). More generally, co-regulated genes were

also shown to physically interact and to associate with "transcription factories", which are regions enriched for highly phosphorylated (i.e. active) RNAPII (Schoenfelder et al., 2010b). It is now clear that chromatin architecture is not random, but rather adopts preferred three-dimensional conformations which are now being discovered (Duan et al., 2010). Thus, our ability to study protein complexes associated with DNA in their native environment should prove invaluable for the study of chromatin.

mChIP-MS analyses of well characterized proteins, such as the nucleosome assembly factor Asf1, also revealed numerous novel protein associations. For instance, the association between Asf1 and transcription factors (e.g. Pho2, Pdr1) is likely indirect and is lost in the absence of *hir1* (Appendix 3-7). Interestingly, Asf1 and Pho2 have been previously localized to the *PHO5* promoter, and both proteins are essential for proper *PHO5* activation (Adkins et al., 2007). Moreover, nucleosome assembly at *PHO5* was found to be delayed in the absence of Hir1 (Schermer et al., 2005) which raises the possibility of a direct action of HIR in the association between Asf1 and Pho2. It is unknown whether Pho2, a transcription factor, can directly recruit Asf1 via Hir1 to *PHO5* in order to properly evict nucleosomes and thus promote *PHO5* expression. More generally, I found that the HIR requirement for mediating Asf1 interactions was reflected in the HIR requirement for Asf1 to recruit the H3/H4 chaperone Rtt106.

Even though most open reading frames in *Saccharomyces cerevisiae* have been analyzed by AP-MS, I was able to detect numerous novel protein-protein interactions for many baits associated with chromatin. This reinforces the need to further analyze protein-

protein interactions in model organisms, such as budding yeast, using novel techniques designed for a specific class of baits. For example, proteins associated with membranes would greatly benefit from improved protocols. In addition, technical improvements of affinity reagents and of the sensitivity of mass spectrometers will contribute to the detection of many more protein-protein interactions.

Bibliography:

Adkins, M.W., Williams, S.K., Linger, J., and Tyler, J.K. (2007). Chromatin disassembly from the PHO5 promoter is essential for the recruitment of the general transcription machinery and coactivators. Mol Cell Biol *27*, 6372-6382.

Akiyoshi, B., Nelson, C.R., Ranish, J.A., and Biggins, S. (2009). Quantitative proteomic analysis of purified yeast kinetochores identifies a PP1 regulatory subunit. Genes Dev *23*, 2887-2899.

Alper, B.J., Rowse, J.W., and Schmidt, W.K. (2009). Yeast Ste23p shares functional similarities with mammalian insulin-degrading enzymes. Yeast *26*, 595-610.

Arad, M., Benson, D.W., Perez-Atayde, A.R., McKenna, W.J., Sparks, E.A., Kanter, R.J., McGarry, K., Seidman, J.G., and Seidman, C.E. (2002). Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. J Clin Invest *109*, 357-362.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bienvenu, F., Jirawatnotai, S., Elias, J.E., Meyer, C.A., Mizeracka, K., Marson, A., Frampton, G.M., Cole, M.F., Odom, D.T., Odajima, J., *et al.* (2010). Transcriptional role of cyclin D1 in development revealed by a genetic-proteomic screen. Nature *463*, 374-378.

Bonenfant, D., Coulot, M., Towbin, H., Schindler, P., and van Oostrum, J. (2006). Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. Mol Cell Proteomics *5*, 541-552.

Burckstummer, T., Baumann, C., Bluml, S., Dixit, E., Durnberger, G., Jahn, H., Planyavsky, M., Bilban, M., Colinge, J., Bennett, K.L., and Superti-Furga, G. (2009). An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. Nat Immunol *10*, 266-272.

Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. Annu Rev Genet *43*, 559-599.

Chen, G.I., and Gingras, A.C. (2007). Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. Methods *42*, 298-305.

Cipak, L., Spirek, M., Novatchkova, M., Chen, Z., Rumpf, C., Lugmayr, W., Mechtler, K., Ammerer, G., Csaszar, E., and Gregan, J. (2009). An improved strategy for tandem affinity purification-tagging of Schizosaccharomyces pombe genes. Proteomics *9*, 4825-4828.

Cloutier, P., Al-Khoury, R., Lavallee-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. Methods *48*, 381-386.

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Donaldson, A.D., and Kilmartin, J.V. (1996). Spc42p: a phosphorylated component of the S. cerevisiae spindle pole body (SPD) with an essential function during SPB duplication. J Cell Biol *132*, 887-901.

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A., and Noble, W.S. (2010). A three-dimensional model of the yeast genome. Nature.

Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat Methods *2*, 667-675.

Emili, A., Schieltz, D.M., Yates, J.R., 3rd, and Hartwell, L.H. (2001). Dynamic interaction of DNA damage checkpoint protein Rad53 with chromatin assembly factor Asf1. Mol Cell *7*, 13-20.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495-508.

Errede, B., and Ammerer, G. (1989). STE12, a protein involved in cell-type-specific transcription and signal transduction in yeast, is part of protein-DNA complexes. Genes Dev *3*, 1349-1361.

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

Ferrezuelo, F., Aldea, M., and Futcher, B. (2009). Bck2 is a phase-independent activator of cell cycle-regulated genes in yeast. Cell Cycle *8*, 239-252.

Fillingham, J., Kainth, P., Lambert, J.P., van Bakel, H., Tsui, K., Pena-Castillo, L., Nislow, C., Figeys, D., Hughes, T.R., Greenblatt, J., and Andrews, B.J. (2009). Two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. Mol Cell *35*, 340-351.

Fillingham, J., Recht, J., Silva, A.C., Suter, B., Emili, A., Stagljar, I., Krogan, N.J., Allis, C.D., Keogh, M.C., and Greenblatt, J.F. (2008). Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol Cell Biol *28*, 4342-4353.

Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A., *et al.* (2010). A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. Cell *141*, 407-418.

Francois, J., and Parrou, J.L. (2001). Reserve carbohydrates metabolism in the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev *25*, 125-145.

Fraser, P. (2006). Transcriptional control thrown for a loop. Curr Opin Genet Dev *16*, 490-495.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., *et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *462*, 58-64.

Garcia, B.A., Shabanowitz, J., and Hunt, D.F. (2007). Characterization of histones and their post-translational modifications by mass spectrometry. Curr Opin Chem Biol *11*, 66-73.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

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

Glatter, T., Wepf, A., Aebersold, R., and Gstaiger, M. (2009). An integrated workflow for charting the human interaction proteome: insights into the PP2A system. Mol Syst Biol *5*, 237.

Green, E.M., Antczak, A.J., Bailey, A.O., Franco, A.A., Wu, K.J., Yates, J.R., 3rd, and Kaufman, P.D. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Curr Biol *15*, 2044-2049.

Gstaiger, M., and Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. Nat Rev Genet *10*, 617-627.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180-183.

Hochstrasser, M., Ellison, M.J., Chau, V., and Varshavsky, A. (1991). The short-lived MAT alpha 2 transcriptional regulator is ubiquitinated in vivo. Proc Natl Acad Sci U S A *88*, 4606-4610.

Howson, R., Huh, W.K., Ghaemmaghami, S., Falvo, J.V., Bower, K., Belle, A., Dephoure, N., Wykoff, D.D., Weissman, J.S., and O'Shea, E.K. (2005). Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. Comp Funct Genomics *6*, 2-16.

Hu, F., Alcasabas, A.A., and Elledge, S.J. (2001). Asf1 links Rad53 to control of chromatin assembly. Genes Dev *15*, 1061-1066.

Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E., and Zhang, Z. (2005). Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proc Natl Acad Sci U S A *102*, 13410-13415.

Huang, S., Zhou, H., Tarara, J., and Zhang, Z. (2007). A novel role for histone chaperones CAF-1 and Rtt106p in heterochromatin silencing. EMBO J *26*, 2274-2283.

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

Hwang, W.W., Venkatasubrahmanyam, S., Ianculescu, A.G., Tong, A., Boone, C., and Madhani, H.D. (2003). A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. Mol Cell *11*, 261-266.

Jaspersen, S.L., and Winey, M. (2004). The budding yeast spindle pole body: structure, duplication, and function. Annu Rev Cell Dev Biol *20*, 1-28.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., *et al.* (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol Cell *27*, 262-274.

Jiang, L., Smith, J.N., Anderson, S.L., Ma, P., Mizzen, C.A., and Kelleher, N.L. (2007). Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. J Biol Chem *282*, 27923-27934.

Kao, C.F., and Osley, M.A. (2003). In vivo assays to study histone ubiquitylation. Methods *31*, 59-66.

Kim, H.S., Vanoosthuyse, V., Fillingham, J., Roguev, A., Watt, S., Kislinger, T., Treyer, A., Carpenter, L.R., Bennett, C.S., Emili, A., *et al.* (2009a). An acetylated form of histone H2A.Z regulates chromosome architecture in Schizosaccharomyces pombe. Nat Struct Mol Biol *16*, 1286-1293.

Kim, J., Cantor, A.B., Orkin, S.H., and Wang, J. (2009b). Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. Nat Protoc *4*, 506-517.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., *et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Kuo, M.H., and Allis, C.D. (1999). In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. Methods *19*, 425-433.

Lambert, J.P., Baetz, K., and Figeys, D. (2010). Of proteins and DNA--proteomic role in the field of chromatin research. Mol Biosyst *6*, 30-37.

Lambert, J.P., Mitchell, L., Rudner, A., Baetz, K., and Figeys, D. (2009). A novel proteomics approach for the discovery of chromatin-associated protein networks. Mol Cell Proteomics *8*, 870-882.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. Cell *134*, 244-255.

Liu, H., Sadygov, R.G., and Yates, J.R., 3rd (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem *76*, 4193-4201.

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.

Lu, K.P., Finn, G., Lee, T.H., and Nicholson, L.K. (2007). Prolyl cis-trans isomerization as a molecular timer. Nat Chem Biol *3*, 619-629.

Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature *436*, 294-298.

Meacham, G.C., Browne, B.L., Zhang, W., Kellermayer, R., Bedwell, D.M., and Cyr, D.M. (1999). Mutations in the yeast Hsp40 chaperone protein Ydj1 cause defects in Axl1 biogenesis and pro-a-factor processing. J Biol Chem *274*, 34396-34402.

Mitchell, L., Lambert, J.P., Gerdes, M., Al-Madhoun, A.S., Skerjanc, I.S., Figeys, D., and Baetz, K. (2008). Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 is essential for complex integrity. Mol Cell Biol *28*, 2244-2256.

Moll, T., Dirick, L., Auer, H., Bonkovsky, J., and Nasmyth, K. (1992). SWI6 is a regulatory subunit of two different cell cycle START-dependent transcription factors in Saccharomyces cerevisiae. J Cell Sci Suppl *16*, 87-96.

Mousson, F., Lautrette, A., Thuret, J.Y., Agez, M., Courbeyrette, R., Amigues, B., Becker, E., Neumann, J.M., Guerois, R., Mann, C., and Ochsenbein, F. (2005). Structural basis for the interaction of Asf1 with histone H3 and its functional implications. Proc Natl Acad Sci U S A *102*, 5975-5980.

Munakata, T., Adachi, N., Yokoyama, N., Kuzuhara, T., and Horikoshi, M. (2000). A human homologue of yeast anti-silencing factor has histone chaperone activity. Genes Cells *5*, 221-233.

Ooi, S.L., Henikoff, J.G., and Henikoff, S. (2010). A native chromatin purification system for epigenomic profiling in Caenorhabditis elegans. Nucleic Acids Res *38*, e26.

Pijnappel, W.W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Seraphin, B., Aasland, R., and Stewart, A.F. (2001). The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev *15*, 2991-3004.

Pinkham, J.L., and Guarente, L. (1985). Cloning and molecular analysis of the HAP2 locus: a global regulator of respiratory genes in Saccharomyces cerevisiae. Mol Cell Biol *5*, 3410-3416.

Polekhina, G., Gupta, A., Michell, B.J., van Denderen, B., Murthy, S., Feil, S.C., Jennings, I.G., Campbell, D.J., Witters, L.A., Parker, M.W., *et al.* (2003). AMPK beta subunit targets metabolic stress sensing to glycogen. Curr Biol *13*, 867-871.

Pramila, T., Miles, S., GuhaThakurta, D., Jemiolo, D., and Breeden, L.L. (2002). Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle. Genes Dev *16*, 3034-3045.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218-229.

Rass, U., and Kemper, B. (2002). Crp1p, a new cruciform DNA-binding protein in the yeast Saccharomyces cerevisiae. J Mol Biol *323*, 685-700.

Recht, J., Tsubota, T., Tanny, J.C., Diaz, R.L., Berger, J.M., Zhang, X., Garcia, B.A., Shabanowitz, J., Burlingame, A.L., Hunt, D.F., *et al.* (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. Proc Natl Acad Sci U S A *103*, 6988-6993.

Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., *et al.* (2000). Genome-wide location and function of DNA binding proteins. Science *290*, 2306-2309.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol *17*, 1030-1032.

Robinson, M.D., Grigull, J., Mohammad, N., and Hughes, T.R. (2002). FunSpec: a web-based cluster interpreter for yeast. BMC Bioinformatics *3*, 35.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314.

Schermer, U.J., Korber, P., and Horz, W. (2005). Histones are incorporated in trans during reassembly of the yeast PHO5 promoter. Mol Cell *19*, 279-285.

Schoenfelder, S., Clay, I., and Fraser, P. (2010a). The transcriptional interactome: gene expression in 3D. Curr Opin Genet Dev *20*, 127-133.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., *et al.* (2010b). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet *42*, 53-61.

Schultz-Norton, J.R., Ziegler, Y.S., Likhite, V.S., Yates, J.R., and Nardulli, A.M. (2008). Isolation of novel coregulatory protein networks associated with DNA-bound estrogen receptor alpha. BMC Mol Biol *9*, 97.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Sharp, J.A., Fouts, E.T., Krawitz, D.C., and Kaufman, P.D. (2001). Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. Curr Biol *11*, 463-473.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Stark, C., Breitkreutz, B.J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. Nucleic Acids Res *34*, D535-539.

Stirling, D.A., and Stark, M.J. (1996). The phosphorylation state of the 110 kDa component of the yeast spindle pole body shows cell cycle dependent regulation. Biochem Biophys Res Commun *222*, 236-242.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Sutton, A., Bucaria, J., Osley, M.A., and Sternglanz, R. (2001). Yeast ASF1 protein is required for cell cycle regulation of histone gene transcription. Genetics *158*, 587-596.

Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabanowitz, J., Chait, B.T., Hunt, D.F., and Allis, C.D. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. Proc Natl Acad Sci U S A *104*, 2086-2091.

Thomas, C.E., Kelleher, N.L., and Mizzen, C.A. (2006). Mass spectrometric characterization of human histone H3: a bird's eye view. J Proteome Res *5*, 240-247.

Treger, J.M., Magee, T.R., and McEntee, K. (1998a). Functional analysis of the stress response element and its role in the multistress response of Saccharomyces cerevisiae. Biochem Biophys Res Commun *243*, 13-19.

Treger, J.M., Schmitt, A.P., Simon, J.R., and McEntee, K. (1998b). Transcriptional factor mutations reveal regulatory complexities of heat shock and newly identified stress genes in Saccharomyces cerevisiae. J Biol Chem *273*, 26875-26879.

Unnikrishnan, A., Gafken, P.R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. Nat Struct Mol Biol *17*, 430-437.

Van Den Hazel, H.B., Kielland-Brandt, M.C., and Winther, J.R. (1996). Review: biosynthesis and function of yeast vacuolar proteases. Yeast *12*, 1-16.

Veraksa, A., Bauer, A., and Artavanis-Tsakonas, S. (2005). Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. Dev Dyn *232*, 827-834.

Walhout, A.J. (2006). Unraveling transcription regulatory networks by protein-DNA and protein-protein interaction mapping. Genome Res *16*, 1445-1454.

Wigge, P.A., Jensen, O.N., Holmes, S., Soues, S., Mann, M., and Kilmartin, J.V. (1998). Analysis of the Saccharomyces spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. J Cell Biol *141*, 967-977.

Willems, A.R., Lanker, S., Patton, E.E., Craig, K.L., Nason, T.F., Mathias, N., Kobayashi, R., Wittenberg, C., and Tyers, M. (1996). Cdc53 targets phosphorylated G1 cyclins for degradation by the ubiquitin proteolytic pathway. Cell *86*, 453-463.

Wittenberg, C., and Reed, S.I. (2005). Cell cycle-dependent transcription in yeast: promoters, transcription factors, and transcriptomes. Oncogene *24*, 2746-2755.

Wood, A., Krogan, N.J., Dover, J., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J.F., *et al.* (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell *11*, 267-274.

Workman, J.L. (2006). Nucleosome displacement in transcription. Genes Dev *20*, 2009-2017.

Zhou, L., Ma, X., Arbeitman, M.N., and Sun, F. (2009). Chromatin regulation and gene centrality are essential for controlling fitness pleiotropy in yeast. PLoS One *4*, e8086.

Chapter 4

Summary and Future Directions

4.1 Summary

The overall hypothesis of protein interaction mapping is that proteins found to interact must be involved in common processes and common localisation i.e. guilt by association. The large scale mapping of proteins interactions allows to annotate protein of unknown functions, implicate protein of known function in different processes and derive new hypothesis. This is possible because most proteins do not act in isolation but rather as part of complexes and thus possess interaction partners that can now be detected with the right tools. As I have described in chapter 1, technical progress in the field of proteomics fuelled numerous large-scale studies enabling the determination of protein-protein interactions at an improved throughput and sensitivity. These successes have led to the *proposal of a systematic characterization of the protein complement of the genome* with regards to its expression, localization and interconnection (2010). It remains to be seen whether such a large endeavour will come to fruition.

Despite the numerous benefits of the AP-MS approach in the characterization of protein-protein interactions, some classes of proteins remained particularly challenging to study. For instance, membrane embedded proteins with their large hydrophobic regions tend to be difficult baits for affinity purification. Their low solubility in aqueous buffer requires the use of high levels of detergents which in turn disrupts some protein-protein interaction (Lin and Guidotti, 2009). Another class of protein which remains underrepresented in many large-scale studies are structural proteins, such as mitotic microtubules, which are large polymer-like proteins of low abundance (Tao and Scholey,

2010). Improvements in their characterization were achieved when the microtubules were first stabilized with Taxol facilitating their affinity purification (Gache et al., 2007). Protein complex solubility and stability are thus critical factors that must be controlled in large-scale AP-MS studies.

As I have also described in chapter 1, proteins binding to DNA are critical regulators of biological systems but remain difficult to study by AP-MS. Some of the principle problems affecting DNA binding proteins are their low abundance, solubility and stability. In chapter 2, I described in detail the development of the mChIP method which tackled these issues to facilitate the purification of proteins associated with DNA. To get around the inherent low solubility of chromatin associated proteins, the mChIP was optimized to function with minimal centrifugation steps reducing the irreversible precipitation of chromatin. It is interesting to note that an on-going study targeting human transcriptional coregulator complexes utilizes two 100 000 X *g* ultracentrifugation steps (Malovannaya et al., 2010). While this greatly reduces background contamination, it also removes DNA and the protein associated with it. Another aspect of the mChIP method that distinguishes it from common AP-MS approaches is the low level of salt (20 mM magnesium acetate) present in its lysis and wash buffer. The classical TAP-MS utilized 150mM sodium chloride (NaCl) which helps to reduce non-specific binding of protein to the affinity resin (Rigaut et al., 1999). Numerous laboratories utilize "high salt" wash step with NaCl concentration of 500mM or more to further reduce background contaminants (Wang et al., 2009). It is interesting to note that such high concentration of NaCl (>500mM) is extremely efficient at

extracting histones and other chromatin associated proteins rendering the identification of chromatin associated protein networks difficult, if not impossible.

Using the histone H2A or its variant Htz1, I have shown in chapter 2 that mChIP can purify distinct chromatin associated protein networks. Furthermore, the mChIP approach purifies direct interactors as well as indirect interactors that come along because of their association with DNA. Interestingly, the direct and indirect interactors can be distinguished by shrinking the size of DNA surrounding the bait protein. For instance, the Top2 DNA topoisomerase was a distal association observed only when Htz1-TAP was mChIP from sonicated cell extract, while the Nap1 histone chaperone was detected under all conditions tested. While I have not exploited this tool for numerous baits, I will describe in the next section how this could be coupled with mini-chromosome purification to define precisely the protein environment of a given genomic loci.

In Chapter 2, I have described how mChIP-MS can be utilized to identify interaction partners of non-histone chromatin associated proteins. Moreover, I reported in chapter 3 a large mChIP study targeting transcription factors and other poorly characterized DNA binding proteins. As well, other classes of baits of interest to me or my collaborators, with ties to chromatin biology were also targeted as part of this study. I have shown that for the majority of these baits, mChIP-MS produced high quality data which surpassed previously published TAP-MS data. In particular, transcription factors involve in cell-cycle regulation, such as Swi4, Swi6 and Mcm1, benefited from mChIP-MS analysis which significantly extended their known interaction partners.

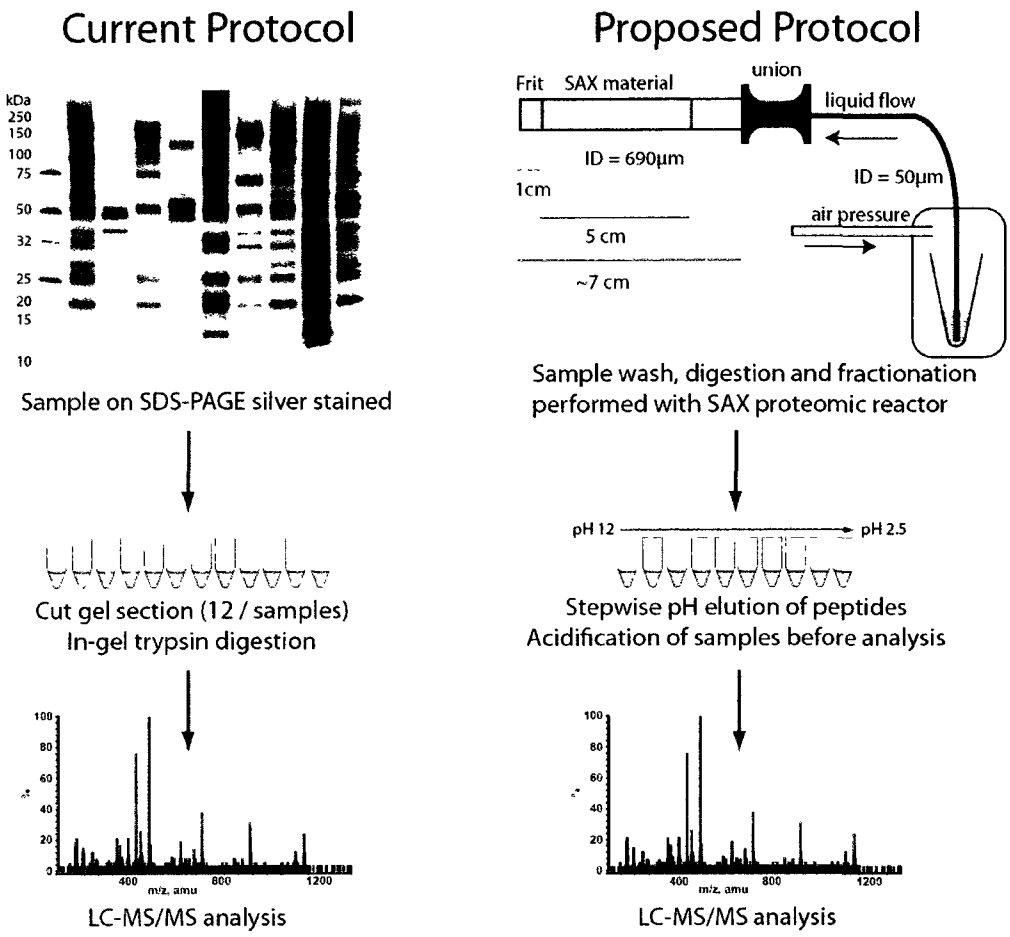
As part of a large study in 2009, I utilized the mChIP approach to detect novel physical associations between Rtt106 and two other histone chaperones, namely Asf1 and the HIR complex (Fillingham et al., 2009). In chapter 3, I followed-up on this work and described how mChIP can be utilized in targeted studies of histone chaperones with affinity for histones H3/H4. By utilizing a mChIP-WB approach, I demonstrated that *ASF1* is essential for the association between Rtt106 and the HIR complex. Moreover, the use of Asf1 V94R point-mutant enabled me to observe that the Rtt106 – Asf1 association requires the presence of histone H3/H4. This reflects the known Rtt106 dependency on *ASF1* for its binding to the *HTA1-HTB1* promoter (Fillingham et al., 2009) and is consistent with the absence of interaction between recombinant HIR and Rtt106 (Huang et al., 2005). Subsequent Asf1 mChIP-MS data revealed that the loss of H3K56Ac influenced the extent of the protein network associated with Asf1. In particular, a greater than two fold reduction in the Rtt106 – Asf1 association was observed by mChIP-MS in all samples lacking H3K56Ac. This is consistent with the presence of a PH-like domain in Rtt106 which possesses an affinity for H3K56Ac (Li et al., 2008). In summary, the use of mChIP enabled me to investigate novel aspects of histone chaperones biology and to better understand the mechanism by which they regulate the *HTA1-HTB1* promoter.

4.2 Future directions

As part of my graduate work, I have performed mChIP purification for more than 100 different baits. While this was a large endeavour requiring significant time and financial

investments, these baits represent about 10% of all nuclear proteins in budding yeast and about a quarter of all proteins predicted to bind DNA. Thus, a logical extension of my graduate work would be to continue exploiting the mChIP approach to fully characterize the budding yeast chromatin associated interactome. In particular the 120 predicted transcription factors in *S. cerevisiae* not already targeted by mChIP-MS in Chapter 3, would be targets of choice for future mChIP work.

As previously stated, a significant obstacle to increasing the number of mChIP-MS experiments performed is the overall low throughput of the in-gel digestion experimental procedure utilized in Chapter 2 and 3 (Figure 4-1). Briefly, following the successful mChIP of a given bait, 90% of the purified material is resolved on a SDS-PAGE gel and silver stained. Then, 12 gel sections are cut for each sample lane, in-gel digested with trypsin and subsequently analyzed by LC-MS/MS. For the analysis of 4 mChIP samples, this protocol requires approximately 9 days of work. An alternative to gel-based approaches is the proteomic reactor, an efficient gel-free device for proteins sample preparation developed in the Figeys laboratory (Ethier et al., 2006). The use of the proteomic reactor streamlines the sample preparation by combining sample wash, protein digestion and peptide fractionation into a single device (Zhou et al., 2010). Utilizing this streamline approach, the time required to perform 4 mChIP-MS analysis would be reduced to 6 days (Figure 4-1). This would facilitate future projects requiring mChIP-MS as well as extension of the mChIP that I will be describing in the following sections.



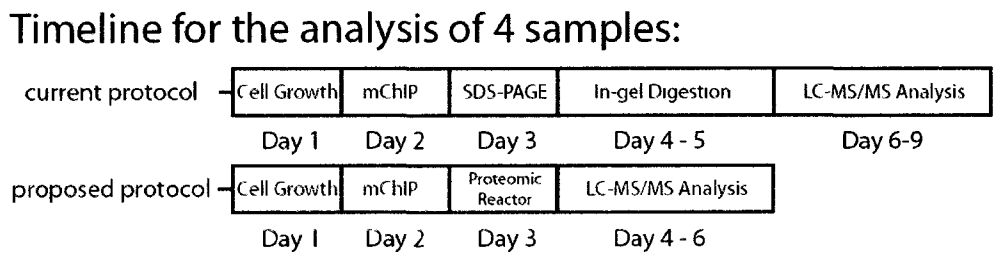


Figure 4-1: Proposed experimental pipeline for mChIP-MS analysis. The protocol utilized in Chapter 2 and 3 is based on 1D gel electrophoresis to visualize, clean-up and fractionate the protein samples. A "gel-free" protocol based on the proteomic reactor is proposed to improve the throughput of the mChIP-MS approach. The proteomic reactor is composed on strong anion exchange (SAX) material packed into a fritted fused silica capillary which will retain protein at high pH, enabling efficient sample clean-up and digestion. By circumventing the use of SDS-PAGE, it is possible to increase the sample preparation throughput by ~33%. See reference (Ethier et al., 2006) and (Zhou et al., 2010) for detailed description of the proteomic reactor.

4.2.1 Characterization of chromatin bound protein network associated with particular histone PTM

The histone code hypothesis states that "... *distinct histone modifications, on one or more tails, act sequentially or in combination to form a 'histone code' that is, read by other proteins to bring about distinct downstream events.* " (Strahl and Allis, 2000). As such histone PTM can be considered as a "binding signal" which will recruit proteins containing appropriate "binding module" (Figure 4-2A). As I have shown in chapter 3, H3K56Ac promotes the association between Asf1 and Rtt106 (Figure 3-11) and thus, can be considered a binding signal on chromatin. This histone PTM is known to mark newly synthesized histone H3 (Masumoto et al., 2005), participate in proper DNA damage response (Collins et al., 2007; Driscoll et al., 2007) and to be cell-cycle regulated (Masumoto et al., 2005). Despite the wealth of information currently known regarding H3K56Ac, we do not know if it recruits other proteins to chromatin. At the moment, researchers attempting to define which proteins bind to specific histone PTM rely on peptide immunopurification experiments. To do so, a peptide of interest is synthesized with and without a given PTM, immobilized onto an appropriate solid support and utilized in AP-MS experiments from nuclear extracts (Chan et al., 2009; Taverna et al., 2006; Vermeulen et al., 2007). This *in vitro* approach has led to the successful identification of some binding partners for PTM located on the N-terminal tails of histone proteins (Chan et al., 2009; Taverna et al., 2006; Vermeulen et al., 2007) but remains untested for PTM located in globular domains of histone, sections of the histone proteins in association with DNA and possessing precise conformations. Moreover, this approach is incapable of studying proteins indirectly recruited in proximity to a given histone mark.

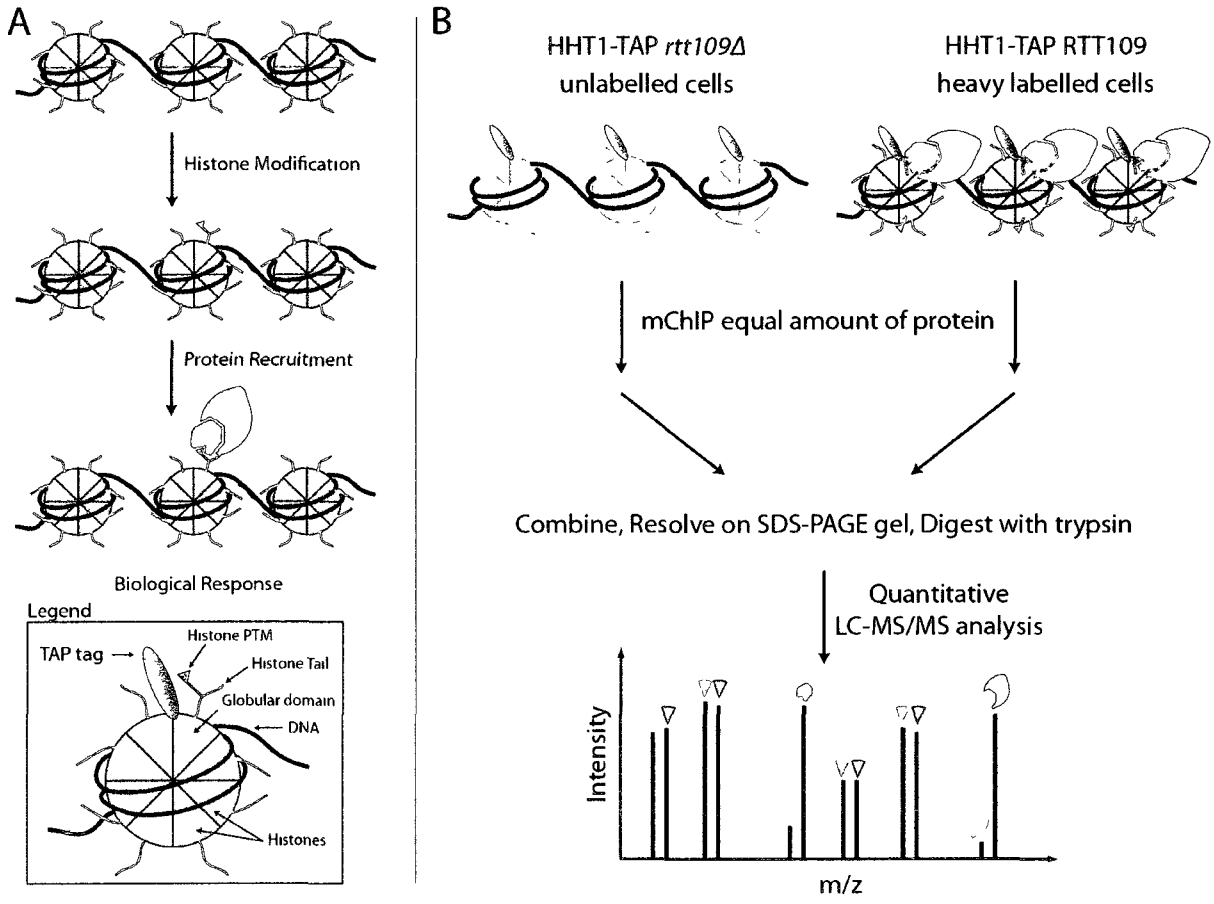


Figure 4-2: Direct detection of protein networks associated with specific histone PTM. (A) Post-translational modifications on histone proteins can act as binding signal recruiting specific chromatin binding proteins to elicit appropriate biological responses. **(B)** Proposed experimental scheme based on the mChIP approach to determine the protein networks associated with specific histone PTM.

To circumvent these issues, I propose to adapt the mChIP procedure to the study of chromatin bound protein networks associated with H3K56Ac (Figure 4-2B). To do so, strains expressing TAP tagged histone H3 (HHT1-TAP) in conjunction, or not, with *rtt109* deletion will be prepared. As Rtt109 is the only enzyme acetylating H3 at K56, this will generate two samples distinguished at the chromatin level only by the absence or presence of H3K56Ac. Then, each strain will be grown in either light or heavy SILAC media and mChIP individually from equal amounts of yeast extract. The purified samples will then be mixed, resolved on a SDS-PAGE gel, in-gel digested and quantitative LC-MS/MS analysis will be performed. Proteins whose association with chromatin is dependent on H3K56Ac will be detected at a significant higher level in heavy label cells than from the unlabelled cells. Rtt106 will act as a positive control as its binding to chromatin and Asf1 is reduced in the absence of H3K56Ac ((Fillingham et al., 2009), Chapter 3). The proposed procedure will also be applicable to other PTM localized in the globular region of histone proteins such as H3 K79 methylation and H2B K123 ubiquitination. It is also interesting to note that modulation of H3K56Ac levels in fungal pathogens has been proposed as a therapeutic strategy (Wurtele et al., 2010) (Lopes da Rosa et al., 2010). As H3K56Ac is also present in human cells, albeit at a very low level (Das et al., 2009) (Yuan et al., 2009), it is critical to fully understand the biological implication of this histone mark in fungus and mammalian cells.

4.2.2 Characterization of the chromatin environment at specific genomic loci

As I have described in Chapter 1, significant challenges remain in the identification of proteins associated with specific genomic locus. Specifically, the use of DNA probes for the isolation of specific genomic loci *in vivo* lacks sensitivity and requires an unruly amount of starting material (Dejardin and Kingston, 2009) whereas purification of a genomic loci of interest within a minichromosome suffers from high level of background contamination (Akiyoshi et al., 2009; Unnikrishnan et al., 2010). The approach I have taken in my graduate work and described in Chapter 2 and 3 was to focus on the optimization of an AP-MS method for chromatin associated proteins. One drawback of the mChIP approach is that it remains a genome-wide tool where the bait can be, and most likely is, associated with chromatin at numerous locations. For instance, the Asf1 mChIP-MS data described in Chapter 3 identified protein association from multiple genomic regions such as the *HTA1-HTB1* and *PHO5* promoters.

I propose here to combine the mChIP approach with minichromosome purification to enable a detail proteomic study of the *HTA1-HTB1* promoter (Figure 4-3). Briefly, a minichromosome containing the *HTA1-HTB1* genes and promoter, a selectable marker (*TRP1*), an origin of replication (*ARS1*) and eight repeats of the lactose operators (*LacO*) will be constructed based on a design previously reported (Figure 4-3; (Akiyoshi et al., 2009; Unnikrishnan et al., 2010)). This minichromosome will then be transformed into host cells constitutively expressing a tagged version of LacI (such as LacI-FLAG (Akiyoshi et al., 2009)). In addition, the host strain will express Rtt106-TAP under the control of its native promoter, a histone chaperone that was shown to bind to the central regulatory region of the *HTA1-HTB1* promoter ((Fillingham et al., 2009), Figure 3-11D) growth in media lacking tryptophan

to mid-log phase, cells will be lyzed and the yeast extract clarified by centrifugation at high speed to pellet chromatin. The soluble minichromosome will be purified with anti-FLAG M2 antibody crosslink to magnetic beads and eluted with an excess of 3X FLAG peptide. Then, the sample will be divided in three equal aliquots and subsequently sonicated, treated with MNase or DNase as described in Chapter 2 (Figure 4-3). The protein network associated with Rtt106 from each sample will be purified by mChIP and identified by LC-MS/MS enabling a precise mapping of the chromatin environment of the *HTA1-HTB1* promoter. In addition, it would be worthwhile to perform these studies in mutant background, such as *rtt109Δ* and *yta7Δ*, which have been previously shown to affect protein "spreading" at the *HTA1-HTB1* promoter (Fillingham et al., 2009).

Another benefit of this approach is that it would permit me to study the nucleosome assembly/disassembly at the *HTA1-HTB1* promoter throughout the cell-cycle. To do so, the strain described above would be grown to mid-log phase and arrested in G_1 phase with α-factor. Then, the cells would be washed and resuspended in fresh media without α-factor and grown for 120 minutes with aliquots taken every 15 minutes to cover a complete cell-cycle. mChIP-MS of Rtt106-TAP, which was previously shown to bind the *HTA1-HTB1* promoter throughout the cell-cycle (Fillingham et al., 2009), from each aliquots would allow me to identify novel factors participating in the regulation of this promoter. Putative factors associated with the *HTA1-HTB1* promoter identified in this way could be subsequently validated by ChIP.

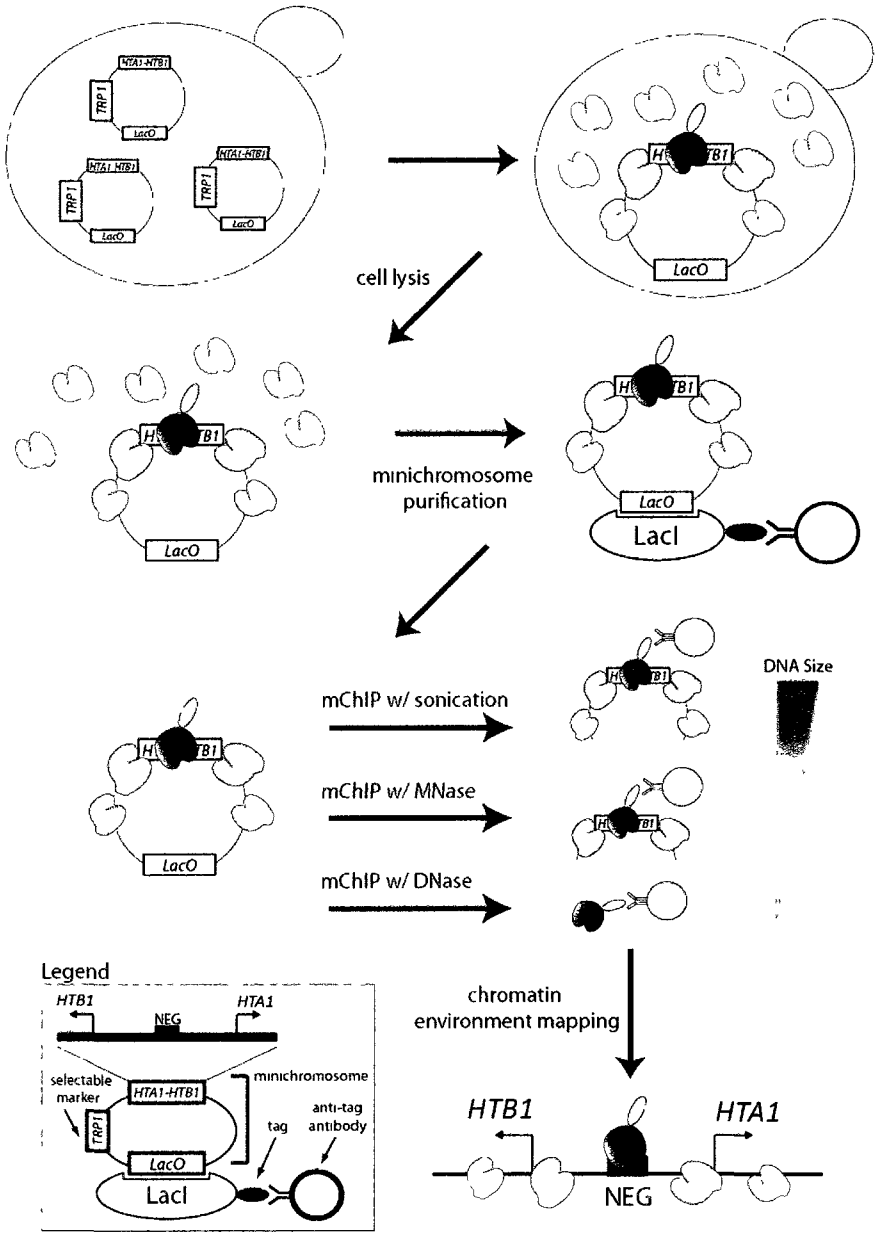


Figure 4-3: Proposed method to investigate the chromatin environment surrounding the *HTA1-HTB1* promoter in *S. cerevisiae.* A minichromosome containing the *HTA1-HTB1* genes and promoter, the *TRP1* selectable marker, an origin of replication and multiple repeats of the lactose operators will be constructed. The minichromosome will be transformed in RTT106-TAP strains and grown in –TRP media until mid-log phase. Cells will be lyzed and centrifuged at high speed to pellet chromatin, but not the more soluble minichromosome. Subsequently, the cell extract will be divided in three equal aliquots and the minichromosome sheared with either sonication, MNase and DNase I. mChIP-MS will be performed on each aliquot enabling the precise characterization of the protein environment of the *HTA1-HTB1* promoter.

4.3 Overall significance

Our understanding of biological systems has greatly benefited from the advent of high-throughput technologies enabling the emergence of systems biology. Genomic tools have been the foundation of systems biology but other "omics" technologies, such as proteomics, are now routinely used to characterize biological systems. The sequencing and identification of genes, proteins and of other biomolecules is now common in life science laboratories. As the quality of the tools used by scientists increases, so does the complexity and scope of the questions pursued. Still, critical biological questions remain unanswered due to the lack of appropriate tools. However, immense progress was achieved for the identification of protein-protein interaction through the development of novel tools. The work presented in this thesis describes the development of specialized tools to efficiently study chromatin associated proteins in the model organism *S. cerevisiae*. The mChIP approach allowed me and my collaborators to effectively characterize the interactome of numerous DNA binding protein onto chromatin, something not previously possible. Moreover, extensions of the mChIP approach proposed here will further improve the characterization of chromatin and our understanding of gene expression regulation. Past experience tells us that the continuous development and applications of new tools provides further insight in the basis of life. My work contributed to this by facilitating the study of DNA binding proteins and shows the importance of specialized methodologies based on the physical characteristics of the proteins under study. Therefore future large-scale efforts to map protein-protein interaction should not utilized a single approach but rather be integrative and combine the advantages offered by numerous tools, including mChIP.

Bibliography:

(2010). A gene-centric human proteome project: HUPO--the Human Proteome organization. Mol Cell Proteomics *9*, 427-429.

Akiyoshi, B., Nelson, C.R., Ranish, J.A., and Biggins, S. (2009). Quantitative proteomic analysis of purified yeast kinetochores identifies a PP1 regulatory subunit. Genes Dev *23*, 2887-2899.

Chan, D.W., Wang, Y., Wu, M., Wong, J., Qin, J., and Zhao, Y. (2009). Unbiased proteomic screen for binding proteins to modified lysines on histone H3. Proteomics *9*, 2343-2354.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., *et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature *459*, 113-117.

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Driscoll, R., Hudson, A., and Jackson, S.P. (2007). Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science *315*, 649-652.

Ethier, M., Hou, W., Duewel, H.S., and Figeys, D. (2006). The proteomic reactor: a microfluidic device for processing minute amounts of protein prior to mass spectrometry analysis. J Proteome Res *5*, 2754-2759.

Fillingham, J., Kainth, P., Lambert, J.P., van Bakel, H., Tsui, K., Pena-Castillo, L., Nislow, C., Figeys, D., Hughes, T.R., Greenblatt, J., and Andrews, B.J. (2009). Two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. Mol Cell *35*, 340-351.

Gache, V., Waridel, P., Luche, S., Shevchenko, A., and Popov, A.V. (2007). Purification and mass spectrometry identification of microtubule-binding proteins from Xenopus egg extracts. Methods Mol Med *137*, 29-43.

Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E., and Zhang, Z. (2005). Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proc Natl Acad Sci U S A *102*, 13410-13415.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. Cell *134*, 244-255.

Lin, S.H., and Guidotti, G. (2009). Purification of membrane proteins. Methods Enzymol *463*, 619-629.

Lopes da Rosa, J., Boyartchuk, V.L., Zhu, L.J., and Kaufman, P.D. (2010). Histone acetyltransferase Rtt109 is required for Candida albicans pathogenesis. Proc Natl Acad Sci U S A *107*, 1594-1599.

Malovannaya, A., Li, Y., Bulynko, Y., Jung, S.Y., Wang, Y., Lanz, R.B., O'Malley, B.W., and Qin, J. (2010). Streamlined analysis schema for high-throughput identification of endogenous protein complexes. Proc Natl Acad Sci U S A *107*, 2431-2436.

Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature *436*, 294-298.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol *17*, 1030-1032.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Tao, L., and Scholey, J.M. (2010). Purification and assay of mitotic motors. Methods *51*, 233-241.

Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T., *et al.* (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol Cell *24*, 785-796.

Unnikrishnan, A., Gafken, P.R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. Nat Struct Mol Biol *17*, 430-437.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M., and Timmers, H.T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58-69.

Wang, J., Cantor, A.B., and Orkin, S.H. (2009). Tandem affinity purification of protein complexes in mouse embryonic stem cells using in vivo biotinylation. Curr Protoc Stem Cell Biol *Chapter 1*, Unit1B 5.

Wurtele, H., Tsao, S., Lepine, G., Mullick, A., Tremblay, J., Drogaris, P., Lee, E.H., Thibault, P., Verreault, A., and Raymond, M. (2010). Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. Nat Med.

Yuan, J., Pu, M., Zhang, Z., and Lou, Z. (2009). Histone H3-K56 acetylation is important for genomic stability in mammals. Cell Cycle *8*, 1747-1753.

Zhou, H., Hou, W., Lambert, J.P., Tian, R., and Figeys, D. (2010). Analysis of low-abundance proteins using the proteomic reactor with pH fractionation. Talanta *80*, 1526-1531.

Appendix

Appendix 1-1: Published or submitted work performed during my graduate studies not included in this thesis.

- Mitchell L., Lau A., **Lambert J.P.**, Zhou H., Fong Y., Couture J.F., Figeys D., Baetz K. (2010) A NuA4 genome-wide synthetic dosage lethal screen uncovers a novel link between KATs and septin dynamics. Submitted to Molecular and Cellular Biology on August 4th 2010.

  In this report, septins proteins were found to be substrates of the NuA4 KAT complex through genomic and proteomic experiments. My role in this effort was to performed MS analysis and its subsequent analysis to identify precise location of lysine acetylation sites as well as protein identity.

- Kennedy M.A., Kabani N., **Lambert J.P.**, Swayne L.A., Figeys D., Bennett S.A.L., Bryan J., Baetz K. (2010) Ydl133w is a Novel Regulator of Phospholipase D Activity in Saccharomyces cerevisiae. Submitted to PLOS genetics on July 12th 2010.

  As part of this report, I identified a novel physical association between the uncharacterized protein Ydl133w and Spo14, a yeast phospholipase D (Figure 3). This association had never been detected before and explain many of the phenotypes associated with the deletion of the YDL133W gene.

- Zhou H., Hou W., **Lambert J.P.**, Figeys D. (2010) New ammunition for the proteomic reactor: strong anion exchange beads and multiple enzymes enhance protein identification and sequence coverage. Analytical and Bioanalytical Chemistry. 2010 Jun 3 in press.

  I participated in this study by providing reagents and samples as well as supporting the manuscript preparation and editing.

- Zhou H., Hou W., **Lambert J.P.**, Tian R., Figeys D. (2010) Analysis of low-abundance proteins using the proteomic reactor with pH fractionation. Talenta. 80(4):1526-1531.

  I participated in this study by providing reagents and samples as well as supporting the manuscript preparation and editing.

- Fillingham J., Kainth P., **Lambert J.P.**, Bakel H.V., Tsui K., Peña-Castillo L., Nislow C., Figeys D., Hughes T.R., Greenblatt J., Andrews B.J. (2009) A two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. Molecular Cell. 35(3):340-351.

  As part of this report, I identified a novel physical association between Rtt106, HIR and Asf1 by mChIP-MS (Figure 4). These associations had never been detected before by classival AP-MS.

- Mitchell L., **Lambert J.P.**, Gerdes M., Al-Madhoun A., Skerjanc I.S., Figeys D., Beatz K. (2008) Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 is essential for complex integrity. Molecular and Cellular Biology. 28:2244-2246.

  In this report, a detail characterization of the NuA4 KAT complex was performed. I was responsible for performing all the MS analysis reported in this paper.

- Abu-Farha M., **Lambert J.P.**, Al-Madhoun A., Elsma F., Skerjanc I.S., Figeys D. (2008) The Tale of Two Domains: Proteomic and Genomic Analysis of SMYD2 a New Histone Methyltransferase. Molecular & Cellular Proteomics. 7:560-572.

  In this paper, SMYD2 an uncharacterized KMT was study to determine its interaction partners as well as the sites of the methylation of histone H3. I performed all the MS analysis reported in this paper.

- Smith J.C.*, **Lambert J.P.***, Elisma F., Figeys D. (2007) Proteomics in 2005/2006: developments, applications and challenges. Analytical Chemistry. 79:4325-4343. * Authors participated equally to work.

  Biannual proteomics review published in Analytical Chemistry. I was responsible for writing ~40% of the manuscript.

- Ewing R.M., Chu P., Elisma F., Li H., Taylor P., Climie S., McBroom-Cerajewski L., Robinson M.D., O'Connor L., Li M., Taylor R., Dharsee M., Ho Y., Heilbut A., Moore L., Zhang S., Ornatsky O., Bukhman Y.V., Ethier M., Sheng Y., Vasilescu J., Abu-Farha M., **Lambert J.P.**, Duewel H.S., Stewart I.I., Kuehl B., Hogue K., Colwill K., Gladwish K., Muskat B., Kinach R., Adams S.L., Moran M.F., Morin G.B., Topaloglou T., Figeys D. (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. Molecular Systems Biology. 3:89

  Large high throughput study of protein-protein interaction in human cells. I was one of many students involved in the manual curration of the protein-protein interaction data for the manuscript preparation.

- Denis N.J., Vasilescu J., **Lambert J.P.**, Smith J.C., Figeys D. (2006) Tryptic digestion of ubiquitin standards reveals an improved strategy for identifying ubiquitinated proteins by mass spectrometry. Proteomics. 7:868-874.

  This paper reported the identification of multiple ubiquitination sites using novel microfluidic devices. I participated in the manual curration of the MS data for the identification of ubiquitination sites.

- Ethier M.*, **Lambert J.P.***, Vasilescu J.*, Figeys D. (2006) Mass Spectrometric Analysis of Protein Interaction Networks. Analytica Chimica Acta. 564: 10-18. *Authors participated equally to work.

  Detailed review of techniques for the mapping of the protein-protein interaction. I was responsible for writing one third of the manuscript.

- **Lambert J.P.**, Ethier M., Smith J.C., Figeys D. (2005) Proteomics: from gel based to gel free. Analytical Chemistry. 77: 3771-3787.

  Biannual proteomics review published in Analytical Chemistry. I was responsible for writing ~60% of the manuscript.

Appendix 2-1: Complete mChIP-MS data generated during method development

mChIP bait*	Associated Preys
Untagged control (1)	Act1, Adh1, Cdc19, Eno2, Hhf1, Pdc1, Pma1, Rpl12b, Rps0a, Ssa1, Ssb2, Tdh3
Untagged control (2)	Act1, Cdc19, Eno2, Hta2, Rpp0, Rps3, Ssa1, Tdh3, Tef2, Tif2, Vma2, Yef3
Untagged control (3)	Rpl4a, Rpl7a, Ssb1
HTA2 (1)	Abf2, Adh1, Arb1, Arp4, Arp8, Asc1, Atp1, Bmh1, Bms1, Brx1, Cbf5, Cdc19, Chc1, Chd1, Cic1, Cka2, Dbp10, Dbp7, Dbp9, Ded1, Dhh1, Dip2, Eaf3, Ebp2, Eft2, Emg1, Eno2, Erb1, Fba1, Fpr3, Gar1, Gpm1, Has1, Hhf1, Hho1, Hmo1, Hos3, Hsc82, Hsn1, Hsp26, Hta2, Htb2, Idh1, Idh2, Imd3, Ino80, Ioc2, Isw1, Isw2, Itc1, Kap114, Kap123, Kre33, Kri1, Lst8, Mak21, Mcm2, Mcm4, Mcm5, Mcm7, Mgm101, Mir1, Mot1, Mpp10, Mrt4, Msh2, Mss116, Nan1, Nap1, Net1, Nip7, Noc2, Noc3, Nog1, Nop1, Nop12, Nop15, Nop2, Nop58, Nop7, Nop8, Npl6, Nsr1, Nto1, Nug1, Ola1, Orc1, Pad1, Pck1, Pdc1, Pet9, Pgi1, Pil1, Pma1, Por1, Pox1, Prp43, Psa1, Puf6, Pwp1, Pwp2, Rcl1, Ret1, Rfa1, Rfc1, Rfc2, Rfc3, Rfc4, Rfc5, Rlp7, Rpc34, Rpc40, Rpc82, Rpf1, Rpf2, Rpg1, Rpl11b, Rpl11b, Rpl12b, Rpl13a, Rpl15a, Rpl16b, Rpl17a, Rpl18b, Rpl19b, Rpl20b, Rpl21a, Rpl25, Rpl2b, Rpl4a, Rpl5, Rpl6b, Rpl7a, Rpl8a, Rpo31, Rpp0, Rps1a, Rps2, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rps9b, Rrn6, Rrp5, Rsc2, Rsc3, Rsc58, Rsc6, Rsc9, Sam1, Sas3, Sfc1, Sik1, Spb1, Spt16, Ssa1, Ssb2, Sse1, Stt4, Sup35, Tdh1, Tdh3, Tef2, Tef4, Tfc1, Tif2, Tif6, Top2, Tsa1, Uaf30, Ura2, Urb1, Urb2, Utp10, Utp15, Utp22, Utp4, Utp6, Vma2, Vps1, Yef3, Ykt6, YLR455W, Yra1, Yta7, Ytm1
HTA2 (2)	Abf2, Arb1, Asc1, Brx1, Cbf5, Cdc19, Cic1, Ded1, Ebp2, Erb1, Fpr3, Fpr4, Has1, Hhf1, Hht1, Hmo1, Hos3, Hta2, Htb2, Htz1, Imd3, Ioc2, Isw1, Isw2, Kap114, Kem1, Kre33, Mak21, Nap1, Net1, Noc2, Nog1, Nop1, Nop12, Nop2, Nop4, Nop58, Nug1, Pil1, Pma1, Prp43, Puf6, Rlp7, Rpa135, Rpa190, Rpa49, Rpf2, Rpl11b, Rpl12b, Rpl13a, Rpl14a, Rpl15a, Rpl16a, Rpl16b, Rpl17a, Rpl18a, Rpl18b, Rpl19b, Rpl1b, Rpl20b, Rpl21a, Rpl25, Rpl27b, Rpl2b, Rpl3, Rpl30, Rpl31a, Rpl33a, Rpl36b, Rpl5, Rpl6a, Rpl7a, Rpo31, Rpp0, Rpp2b, Rps0a, Rps10b, Rps13, Rps14b, Rps16b, Rps17b, Rps18a, Rps1b, Rps2, Rps22a, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rps9b, Rrp5, Rvb2, Sik1, Spb1, Srp40, Ssa1, Ssb2, Sth1, Tdh3, Tef2, Top2, Urb1, Utp10, Vma2, Vps1, Yef3, YHR214C-B, Yra1, Yta7, Ytm1
HTA2 (3)	Abf2, Acc1, Arb1, Arp4, Arp9, Asc1, Atp2, Bfr2, Bms1, Brx1, Cbf5, Cdc19, Cdc48, Chd1, Cic1, Cop1, Dbp10, Dbp6, Dbp7, Dbp9, Ded1, Dip2, Drs1, Ebp2, Eft2, Emg1, Enp2, Erb1, Fpr3, Fpr4, Fun12, Gar1, Gcn1, Has1, Hhf1, Hho1, Hht1, Hir2, Hmo1, Hos3, Hsc82, Hta2, htb2, Htz1, Imd3, Imp4, Ioc2, Isw1, Isw2, Itc1, Kap114, Kap123, Kar2, Kem1, Kre33, Kri1, Krr1, Mak16, Mak21, Mak5, Mcm2, Mcm3, Mcm4, Mcm5, Mcm7, Mgm101, Mis1, Mot1, Mrt4, Mss116, Nan1, Nap1, Net1, Nip1, Noc2, Noc3, Nog1, Nop1, Nop12, Nop14, Nop15, Nop2, Nop4, Nop58, Nop7, Nop8, Npl6, Nsr1, Nug1, Orc1, Pab1, Pdc1, Pfk1, Pil1, Pma1, Pob3, Prp4, Prp43, Psh1, Puf6, Pwp1, Pwp2, Rcl1, Ret1, Rfa1, Rfc1, Rfc2, Rfc3, Rfc5, Rlp7, Rpa135, Rpa190, Rpa43, Rpa49, Rpb2, Rpb5, Rpc40, Rpc82, Rpf1, Rpf2, Rpg1, Rpl10, Rpl11b, Rpl12b, Rpl13b, Rpl14b, Rpl15a, Rpl16a, Rpl16b, Rpl17a, Rpl18a, Rpl18b, Rpl19b, Rpl1b, Rpl20b, Rpl21a, Rpl23a, Rpl25, Rpl27b, Rpl28, Rpl2b, Rpl3, Rpl30, Rpl31a,

	Rpl33a, Rpl36a, Rpl4b, Rpl5, Rpl6b, Rpl7b, Rpl9a, Rpo21, Rpo31, Rpp0, Rps0a, Rps11b, Rps12, Rps13, Rps14b, Rps16b, Rps17b, Rps18a, Rps19b, Rps1b, Rps2, Rps22a, Rps24a, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rps9b, Rrp12, Rrp14, Rrp5, Rrp9, Rsc2, Rsc3, Rsc4, Rsc58, Rsc6, Rsc9, Rvb1, Rvb2, Sac6, Sas10, Sas3, Scp160, Sen1, Sfh1, Sik1, Snu13, Spb1, Spt15, Spt16, Spt5, Srp40, Ssa1, Ssb2, Ssf1, Sth1, Taf14, Tdh3, Tef2, Tef4, Tif2, Tif4631, Tif6, Top2, Ura2, Urb1, Urb2, Utp10, Utp13, Utp14, Utp15, Utp18, Utp21, Utp22, Utp4, Utp5, Utp7, Utp8, Utp9, Vma2, Vps1, Yef3, Yer138c, Yra1, Yta7
HTZ1 (1)	Abf2, Adh1, Arp4, Cdc19, Fpr3, Hhf1, Hho1, Hsc82, Hta2, Htb2, Htz1, Ioc2, Isw1, Kap114, Nap1, Pdc1, Rfa1, Rfa3, Ssa2, Ssb2, Ssc1, Tdh3, Top2, Yef3, Yta7
HTZ1 (2)	Asc1, Brx1, Cic1, Erb1, Fpr3, Fpr4, Has1, Hhf1, Hho1, Hht1, Hmo1, Hta2, Htb2, Htz1, Kap114, Nap1, Net1, Nop1, Rlp7, Rpf1, Rpf2, Rpl11b, Rpl12b, Rpl13a, Rpl16b, Rpl17b, Rpl18b, Rpl20b, Rpl21a, Rpl25, Rpl27b, Rpl2b, Rpl3, Rpl35a, Rpl4a, Rpl6b, Rpl7a, Rpl8a, Rpp0, Rpp2b, Rps13, Rps14b, Rps16b, Rps17b, Rps18a, Rps4b, Rps5, Rvb1, Rvb2, Snu13, Ssb1, Swr1, Top2
HTZ1 (3)	Abf2, Acc1, Ade5, Adh1, Arp4, Arp6, Asc1, Atp2, Bms1, Brx1, Bud21, Cdc19, Chc1, Chd1, Cic1, Dbp7, Dip2, Drs1, Ecm16, Eft2, Emg1, Eno2, Erb1, Faa4, Fas1, Fas2, Fba1, Fpr3, Fpr4, Fun12, Gar1, Gas1, Gcn1, Gpm1, Hhf1, Hho1, Hht1, Hmo1, Hsc82, Hta2, Htb2, Htz1, Ilv5, Imd3, Imp4, Ioc2, Isw1, Isw2, Itc1, Kap114, Kap123, Kem1, Kre33, Krr1, Mak21, Mak5, Mcm2, Mcm4, Mcm7, Mgm101, Mir1, Mot1, Mrt4, Nan1, Nap1, Net1, Noc2, Noc3, Nog1, Nop1, Nop12, Nop14, Nop15, Nop16, Nop2, Nop4, Nop58, Nop7, Pdc1, Pet9, Pil1, Pma1, Por1, Prp43, Puf6, Pwp1, Pwp2, Rcl1, Ret1, Rlp7, Rpa135, Rpa190, Rpa43, Rpa49, Rpb2, Rpc40, Rpc82, Rpf1, Rpf2, Rpl10, Rpl11b, Rpl12b, Rpl13a, Rpl14b, Rpl15a, Rpl16b, Rpl17a, Rpl18b , Rpl19a, Rpl19b, Rpl1b, Rpl20b, Rpl21a, Rpl25, Rpl27b, Rpl28, Rpl2b, Rpl3, Rpl30, Rpl31a, Rpl33a, Rpl34b, Rpl35a, Rpl36b, Rpl4a, Rpl5, Rpl6b, Rpl7b, Rpl8b, Rpo21, Rpo31, Rpp0, Rps0a, Rps11b, Rps12, Rps13, Rps14b, Rps16b, Rps17b, Rps18a, Rps19b, Rps1b, Rps2, Rps20, Rps22a, Rps24a, Rps25a, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rps9b, Rrp12, Rrp5, Rsc2, Rsc3, Rsc4, Rvb1, Rvb2, Scp160, Sen1, Shm2, Sik1, Smc3, Snu13, Srp40, Ssa2, Ssb2, Ssc1, Sse1, Sth1, Swc4, Swc5, Swr1, Taf14, Tdh3, Tef2, Tfc3, Tif2, Top1, Top2, Ura2, Urb1, Utp10, Utp13, Utp14, Utp18, Utp21, Utp22, Utp4, Utp8, Utp9, Vma2, Vps1, Vps71, Yef3, Yer138c, Yra1, Yta7
LGE1 (1)	Act1, Adh1, Akl1, Arb1, Ask10, Bem2, Boi1, Bre1, Bre2, Bub2, Cbk1, Cdc14, Cdc19, Cdc31, Cdc33, Cnm67, Dbp1, Dbp2, Ded1, Eft2, Fir1, Glc7, Hhf1, Hos4, Hrb1, Hta2, Imd3, Ipp1, Kar1, Kin2, Kin3, Kin4, Kog1, Kre33, Ksp1, Lge1, Lst8, Mgm101, Mrm1, Msh2, Msh3, Msh6, Mss116, Mud2, Myo2, Nan1, Net1, Nog2, Nop1, Npl3, Nsp1, Nsr1, Nud1, Pet127, Pil1, Prp43, Psp2, Rad52, Rfa1, Rfa3, Rga1, Rpa135, Rpa190, Rpa43, Rpa49, Rpc40, Rpl15a, Rpl16b, Rpl17a, Rpl19b, Rpl1b, Rpl20a, Rpl2a, Rpp0, Rps0a, Rps17b, Rps1b, Rps24a, Rps3, Rps4b, Rps5, Rps6b, Rps8a, Rps9b, Rrp5, Rsp5, Sdc1, Set1, Sif2, Snt1, Sod2, Spc110, Spc29, Spc42, Spc72, Spc97, Spc98, Spp1, Spt5, Sro9, Srp1, Ssa2, Ssb2, Ssd1, Stt4, Tao3, Tco89, Tdh3, Tef2, Tif2, Tof2, Top2, Tor1, Tub4, Vps1, Ybl104c, Ypl141c
LGE1 (2)	Bre1, Cbk1, Ded1, Kog1, Ksp1, Lge1, Npl3, Rpl27b, Rps6b, Snt1, Spt5, Ssa2,

	Ssd1, Tao3, Tef2
LGE1 (3)	Adh1, Akl1, Asc1, Bem2, Bfa1, Boi1, Bre1, Bub2, Cbk1, Cdc14, Cdc19, Cdc31, Cdc48, Cnm67, Dbp2, Ded1, Eft2, Eno2, Erb1, Fir1, Hhf1, Hsc82, Ics2, Imd4, Kar2, Kel1, Kem1, Kin2, Kin3, Kin4, Kog1, Kre33, Ksp1, Lge1, Lre1, Mak21, Mgm101, Mob2, Msh2, Msh3, Msh6, Mss116, Nam7, Net1, Nop1, Nop4, Npl3, Nud1, Pam1, Pet127, Prp43, Psp2, Rad52, Rfa3, Rga1, Rlp7, Rom2, Rpa135, Rpa190, Rpa49, Rpb2, Rpc40, Rpl11b, Rpl15a, Rpl17b, Rpl19b, Rpl20a, Rpl27b, Rpl28, Rpl2a, Rpl3, Rpl4a, Rpl7a, Rpl8a, Rpo21, Rpp0, Rps0a, Rps13, Rps17b, Rps1b, Rps3, Rps4b, Rps5, Rps6a, Rps7b, Rps8a, Rps9b, Snt1, Spc110, Spc29, Spc42, Spc97, Spc98, Spt5, Ssa2, Ssb2, Ssd1, Sse1, Tao3, Tco89, Tdh3, Tef2, Tif2, Top2, Tor1, Utp10, Utp22, Ybl104c, Ypl141c, Ypp1, Yra1
MCM5 (1)	Cdc45, Ctf4, Hhf1, Hta2, Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, Mcm7, Sap185, Ssa1, Ssb1, Tof1, Top2
MCM5 (2)	Ctf4, Hhf1, Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, Mcm7
MCM5 (3)	Arp9, Cdc14, Cdc45, Ctf4, Hhf1, Hsp60, Hta2, Htb2, Isw1, Kar2, Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, Mcm7, Mrc1, Net1, Nsr1, Rfa1, Rpc40, Rpl2a, Rpl4a, Rpo31, Rpp0, Rps3, Rps5, Rps6b, Rsc8, Sap185, Spt16, Ssa1, Ssb1, Ssc1, Tdh3, Tef2, Tif2, Tof1, Top2, Yar009c, Yta7
YTA7 (1)	Abf2, Acc1, Adh1, Arp9, Asc1, Atp2, Cbf5, Cdc19, Cdc48, Chc1, Chd1, Ckb1, Ded1, Dip2, Eft2, Eno2, Erb1, Fba1, Fob1, Fpr3, Fpr4, Fti1, Gpm1, Has1, Hhf1, Hho1, Hht1, Hir2, Hmo1, Hos3, Hta2, Htb2, Htz1, Idh1, Idh2, Imd2, Ino80, Ioc2, Isw1, Itc1, Kem1, Lsm12, Mcm2, Mcm3, Mcm5, Mcm7, Mgm101, Mif2, Mot1, Nan1, Net1, Ngg1, Nop1, Nop12, Nop2, Npl6, Nto1, Orc1, Orc4, Pbp1, Pbp4, Pdc1, Pil1, Pma1, Pob3, Prp43, Psa1, Rad16, Ret1, Rfa1, Rfa2, Rfc1, Rfc2, Rfc3, Rfc5, Rlp7, Rpa135, Rpa190, Rpa34, Rpa43, Rpa49, Rpb2, Rpc19, Rpc40, Rpc82, Rpl10, Rpl11b, Rpl12b, Rpl13b, Rpl14b, Rpl15a, Rpl16b, Rpl17a, Rpl18b, Rpl19b, Rpl1b, Rpl20b, Rpl21a, Rpl2b, Rpl3, Rpl30, Rpl31a, Rpl33a, Rpl36b, Rpl4a, Rpl5, Rpl6b, Rpl7a, Rpl8b, Rpo21, Rpo31, Rpp0, Rpp2b, Rps0a, Rps11b, Rps13, Rps14b, Rps16b, Rps17b, Rps18a, Rps19b, Rps1b, Rps2, Rps20, Rps24a, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rps9b, Rrp5, Rsc2, Rsc3, Rsc4, Rsc58, Rsc6, Rsc9, Rvb1, Rvb2, Sam1, Sas3, Scp160, Sec21, Sen1, Sfh1, Sik1, Snf7, Snu13, Snz2, Spt16, Srp40, Ssa1, Ssb2, Ssc1, Sth1, Sum1, Taf60, Taf61, Taf90, Tdh3, Tef2, Tif2, Tif34, Tof2, Top1, Top2, YBR048W, Ura2, Urb1, Utp10, Utp13, Vid22, Vma2, Vps1, Yap1, Ydl156w, Yhr020w, Yhr054c, Ylr455w, Yra1, Yta7
YTA7 (2)	Acc1, Adh1, Arp9, Asc1, Cdc19, Cdc48, Chd1, Cop1, Ded1, Eft2, Erb1, Fpr3, Fpr4, Hhf1, Hho1, Hht1, Hmo1, Hos3, Hsc82, Hta2, Htb2, Htz1, Imd3, Ioc2, Isw1, Isw2, Mak21, Mcm2, Mcm3, Mcm4, Mcm5, Mcm7, Nan1, Net1, Nop1, Nop12, Nop2, Npl6, Nto1, Orc1, Pbp1, Pma1, Pob3, Prp43, Puf6, Ret1, Rfc1, Rpa135, Rpa190, Rpa49, Rpb2, Rpc40, Rpc82, Rpl12b, Rpl13b, Rpl16b, Rpl17b, Rpl19b, Rpl20b, Rpl21a, Rpl2b, Rpl3, Rpl4a, Rpl6b, Rpl7a, Rpl8a, Rpo31, Rpp0, Rps0a, Rps13, Rps18a, Rps1b, Rps2, Rps3, Rps4b, Rps5, Rps6b, Rps7b, Rps8a, Rps9b, Rrn6, Rrp5, Rsc2, Rsc3, Rsc4, Rsc58, Rsc9, Rvb1, Rvb2, Sas3, Sec27, Sik1, Spt16, Spt5, Srp40, Ssa1, Ssb2, Ssc1, Sth1, Tdh3, Tef2, Top2, Ura2, Vma2, Vps1, YBR012W-B, Yef3, Yhr020w, Ylr455w, Yra1, Yta7
YTA7 (3)	Arp9, Asc1, Cdc19, Chd1, Cka2, Ckb1, Cop1, Ded1, Eft2, Eno1, Fob1, Fpr3,

	Hhf1, Hho1, Hht1, Hmo1, Hos3, Hta2, Htb2, Imd4, Ioc2, Ioc4, Isw1, Kap123, Lsm12, Mak21, Mcm2, Mcm3, Mcm4, Mcm5, Mcm7, Mot1, Net1, Nop1, Npl6, Nto1, Pab1, Pdc1, Pob3, Prp43, Ret1, Rfa1, Rfc1, Rfc2, Rfc3, Rpa135, Rpa190, Rpa43, Rpa49, Rpb2, Rpc40, Rpc82, Rpl10, Rpl13a, Rpl16b, Rpl18b, Rpl19b, Rpl2b, Rpl3, Rpl4a, Rpl5, Rpl6b, Rpl7a, Rpl7b, Rpl8a, Rpo21, Rpo31, Rpp0, Rps0a, Rps14b, Rps1b, Rps2, Rps3, Rps4b, Rps5, Rps6b, Rps7a, Rps8a, Rrn9, Rsc2, Rsc3, Rsc4, Rsc58, Rsc6, Rsc9, Rvb1, Rvb2, Sas3, Snz2, Spt16, Srp40, Ssa1, Ssb2, Sth1, Taf60, Tdh3, Tef2, Tif2, Tof2, Top1, Top2, Vma2, Vps1, Yef3, Yml045w, Yta7

* Number in parentheses denotes biological replicate of the mChIP-MS.

Appendix 2-2: High confidence mChIP-MS dataset generated during mChIP method development.

mChIP bait	Associated Preys
HTA2	ABF2, ARB1, ARP4, ASC1, BMS1, BRX1, CBF5, CIC1, DBP10, DBP7, DBP9, DED1, DIP2, EBP2, EFT2, EMG1, ERB1, FPR3, FPR4, GAR1, HAS1, HHF1, HHO1, HHT1, HMO1, HOS3, HSC82, HTA2, HTB2, HTZ1, IMD3, IOC2, ISW1, ISW2, ITC1, KAP114, KAP123, KEM1, KRE33, KRI1, MAK21, MCM2, MCM4, MCM5, MCM7, MGM101, MOT1, MRT4, MSS116, NAN1, NAP1, NET1, NOC2, NOC3, NOG1, NPL6, NSR1, NUG1, ORC1, PIL1, PRP43, PUF6, PWP1, PWP2, RCL1, RET1, RFA1, RFC1, RFC2, RFC3, RFC5, RLP7, RPA135, RPA190, RPA49, RPC40, RPC82, RPO31, RSC2, RSC3, RSC58, RSC6, RSC9, RVB2, SAS3, SIK1, SPB1, SPT16, SRP40, STH1, TOP2, URA2, URB1, URB2, VPS1, YRA1, YTA7, YTM1
HTZ1	ABF2, ARP4, ASC1, BRX1, CIC1, ERB1, FPR3, FPR4, HHF1, HHO1, HHT1, HMO1, HSC82, HTA2, HTB2, HTZ1, IOC2, ISW1, KAP114, NAP1, NET1, RLP7, RVB1, RVB2, SNU13, SWC3, SWC4, SWC5, SWR1, TOP2, VPS71, VPS72, YTA7
HTZ1 *swr1Δ*	FPR3, FPR4, HHF1, HTB2, HTZ1, KAP114, MPS3, NAP1, STM1
LGE1	AKL1, BEM2, BOI1, BRE1, BUB2, CBK1, CDC14, CDC31, CNM67, DBP2, DED1, EFT2, FIR1, HHF1, KIN2, KIN3, KIN4, KOG1, KRE33, KSP1, LGE1, MGM101, MSH2, MSH3, MSH6, MSS116, NET1, NPL3, NUD1, PET127, PRP43, PSP2, RAD52, RFA3, RGA1, RPA135, RPA190, RPA49, SNT1, SPC110, SPC29, SPC42, SPC97, SPC98, SPT5, SSD1, TAO3, TCO89, TOP2, TOR1, YBL104C, YPL141C
MCM5	CDC45, CTF4, HHF1, HTA2, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, SAP185, TOF1, TOP2
YTA7	ACC1, ARP9, ASC1, CDC48, CHD1, CKB1, COP1, DED1, EFT2, ERB1, FOB1, FPR3, FPR4, HHF1, HHO1, HHT1, HMO1, HOS3, HTA2, HTB2, HTZ1, IOC2, ISW1, LSM12, MAK21, MCM2, MCM3, MCM4, MCM5, MCM7, MOT1, NAN1, NET1, NPL6, NTO1, ORC1, PBP1, POB3, PRP43, RET1, RFA1, RFC1, RFC2, RFC3, RPA135, RPA190, RPA43, RPA49, RPC40, RPC82, RPO21, RPO31, RSC2, RSC3, RSC4, RSC58, RSC6, RSC9, RVB1, RVB2, SAS3, SIK1, SNZ2, SPT16, SRP40, STH1, TAF6, TDH3, TOF2, TOP1, TOP2, VPS1, YHR020W, YLR455W, YRA1, YTA7

Appendix 3-1: Common background contaminants observed by mChIP-MS

ORF Name	Gene Name	ORF Name	Gene Name	ORF Name	Gene Name
YOL086C	ADH1	YLR312W-A	MRPL15	YDL014W	NOP1
YMR116C	ASC1	YBL038W	MRPL16	YOL041C	NOP12
YOL077C	BRX1	YNL252C	MRPL17	YDL148C	NOP14
YAL038W	CDC19	YNL185C	MRPL19	YNL110C	NOP15
YDL031W	DBP10	YKR085C	MRPL20	YER002W	NOP16
YGL078C	DBP3	YNL177C	MRPL22	YNL061W	NOP2
YNR038W	DBP6	YOR150W	MRPL23	YPL043W	NOP4
YKR024C	DBP7	YMR193W	MRPL24	YOR310C	NOP58
YLR276C	DBP9	YGR076C	MRPL25	YGR103W	NOP7
YLL008W	DRS1	YBR282W	MRPL27	YOL144W	NOP8
YAL026C	DRS2	YDR462W	MRPL28	YGR159C	NSR1
YGR254W	ENO1	YMR024W	MRPL3	YGR178C	PBP1
YHR174W	ENO2	YKL138C	MRPL31	YDL053C	PBP4
YAL035W	FUN12	YCR003W	MRPL32	YLR044C	PDC1
YDR024W	FYV1	YMR286W	MRPL33	YOR158W	PET123
YPL240C	HSP82	YDR322W	MRPL35	YGL008C	PMA1
YCR046C	IMG1	YBR122C	MRPL36	YOR361C	PRT1
YCR071C	IMG2	YBR268W	MRPL37	YOL080C	REX4
YHR085W	IPI1	YKL170W	MRPL38	YOR119C	RIO1
YNL182C	IPI3	YML009C	MRPL39	YNL207W	RIO2
YPR169W	JIP5	YLR439W	MRPL4	YHR197W	RIX1
YJL034W	KAR2	YPL173W	MRPL40	YLR009W	RLP24
YHR121W	LSM12	YMR225C	MRPL44	YNL002C	RLP7
YKL021C	MAK11	YJL096W	MRPL49	YEL050C	RML2
YDR060W	MAK21	YNR022C	MRPL50	YPR010C	RPA135
YBR142W	MAK5	YPR100W	MRPL51	YOR341W	RPA190
YLR106C	MDN1	YHR147C	MRPL6	YNL248C	RPA49
YGL068W	MNP1	YDR237W	MRPL7	YHR088W	RPF1
YDR347W	MRP1	YJL063C	MRPL8	YKR081C	RPF2
YDL045W-A	MRP10	YGR220C	MRPL9	YBR079C	RPG1
YGR084C	MRP13	YNR036C	MRPS12	YLR075W	RPL10
YKL003C	MRP17	YPL013C	MRPS16	YPR102C	RPL11A
YPR166C	MRP2	YMR188C	MRPS17	YGR085C	RPL11B
YDR405W	MRP20	YNL306W	MRPS18	YEL054C	RPL12A
YBL090W	MRP21	YDR337W	MRPS28	YDR418W	RPL12B
YHL004W	MRP4	YGR165W	MRPS35	YDL082W	RPL13A
YKL167C	MRP49	YBR251W	MRPS5	YMR142C	RPL13B
YPL118W	MRP51	YMR158W	MRPS8	YKL006W	RPL14A
YNL005C	MRP7	YBR146W	MRPS9	YHL001W	RPL14B
YDR116C	MRPL1	YNL137C	NAM9	YLR029C	RPL15A
YNL284C	MRPL10	YMR309C	NIP1	YMR121C	RPL15B

YDL202W	MRPL11	YPL211W	NIP7	YIL133C	RPL16A
YKR006C	MRPL13	YHR170W	NMD3	YNL069C	RPL16B
YKL180W	RPL17A	YKR094C	RPL40B	YGL123W	RPS2
YJL177W	RPL17B	YDL184C	RPL41A	YHL015W	RPS20
YOL120C	RPL18A	YDL133C-A	RPL41B	YKR057W	RPS21A
YNL301C	RPL18B	YNL162W	RPL42A	YJL136C	RPS21B
YBR084C-A	RPL19A	YHR141C	RPL42B	YKR057W	RPS21A
YBL027W	RPL19B	YPR043W	RPL43A	YJL136C	RPS21B
YPL220W	RPL1A	YJR094W-A	RPL43B	YJL190C	RPS22A
YGL135W	RPL1B	YBR031W	RPL4A	YLR367W	RPS22B
YMR242C	RPL20A	YDR012W	RPL4B	YGR118W	RPS23A
YOR312C	RPL20B	YPL131W	RPL5	YPR132W	RPS23B
YBR191W	RPL21A	YML073C	RPL6A	YER074W	RPS24A
YPL079W	RPL21B	YLR448W	RPL6B	YIL069C	RPS24B
YLR061W	RPL22A	YGL076C	RPL7A	YGR027C	RPS25A
YFL034C-A	RPL22B	YJL177W	RPL7B	YLR333C	RPS25B
YBL087C	RPL23A	YPL198W	RPL7B	YGL189C	RPS26A
YER117W	RPL23B	YHL033C	RPL8A	YER131W	RPS26B
YGL031C	RPL24A	YLL045C	RPL8B	YKL156W	RPS27A
YGR148C	RPL24B	YGL147C	RPL9A	YHR021C	RPS27B
YOL127W	RPL25	YNL067W	RPL9B	YOR167C	RPS28A
YLR344W	RPL26A	YLR340W	RPP0	YLR264W	RPS28B
YGR034W	RPL26B	YDL081C	RPP1A	YLR388W	RPS29A
YHR010W	RPL27A	YDL130W	RPP1B	YDL061C	RPS29B
YDR471W	RPL27B	YOL039W	RPP2A	YNL178W	RPS3
YGL103W	RPL28	YDR382W	RPP2B	YLR287C-A	RPS30A
YFR032C-A	RPL29	YGR214W	RPS0A	YOR182C	RPS30B
YFR031C-A	RPL2A	YLR048W	RPS0B	YLR167W	RPS31
YIL018W	RPL2B	YOR293W	RPS10A	YJR145C	RPS4A
YOR063W	RPL3	YMR230W	RPS10B	YHR203C	RPS4B
YGL030W	RPL30	YDR025W	RPS11A	YJR123W	RPS5
YDL075W	RPL31A	YBR048W	RPS11B	YPL090C	RPS6A
YLR406C	RPL31B	YOR369C	RPS12	YBR181C	RPS6B
YBL092W	RPL32	YDR064W	RPS13	YOR096W	RPS7A
YPL143W	RPL33A	YCR031C	RPS14A	YNL096C	RPS7B
YOR234C	RPL33B	YJL191W	RPS14B	YBL072C	RPS8A
YER056C-A	RPL34A	YOL040C	RPS15	YER102W	RPS8B
YIL052C	RPL34B	YMR143W	RPS16A	YPL081W	RPS9A
YDL191W	RPL35A	YDL083C	RPS16B	YBR189W	RPS9B
YDL136W	RPL35B	YML024W	RPS17A	YHR065C	RRP3
YMR194W	RPL36A	YDR447C	RPS17B	YMR229C	RRP5
YPL249C-A	RPL36B	YDR450W	RPS18A	YCL031C	RRP7
YLR185W	RPL37A	YML026C	RPS18B	YPL193W	RSA1
YDR500C	RPL37B	YOL121C	RPS19A	YLR221C	RSA3
YLR325C	RPL38	YNL302C	RPS19B	YCR072C	RSA4

YDR041W	RSM10
YER050C	RSM18
YNR037C	RSM19
YKL155C	RSM22
YGL129C	RSM23
YDR175C	RSM24
YIL093C	RSM25
YJR101W	RSM26
YGR215W	RSM27
YDR494W	RSM28
YJR113C	RSM7
YPL183W-A	RTC6
YLR197W	SIK1
YFL002C	SPB4
YIR012W	SQT1
YAL005C	SSA1
YLL024C	SSA2
YDL229W	SSB1
YNL209W	SSB2
YJR045C	SSC1
YPL106C	SSE1
YHR066W	SSF1
YDR312W	SSF2
YHR064C	SSZ1
YDR346C	SVF1
YNL081C	SWS2
YJL052W	TDH1
YGR192C	TDH3
YPR080W	TEF1
YBR118W	TEF2
YKL081W	TEF4
YBR123C	TFC1
YAL001C	TFC3
YMR260C	TIF11
YJL138C	TIF2
YGR162W	TIF4631
YPR016C	TIF6
YDL060W	TSR1
YJL130C	URA2
YKL014C	URB1
YJL109C	UTP10
YLR222C	UTP13
YML093W	UTP14
YMR093W	UTP15
YJL069C	UTP18
YLR409C	UTP21
YGR090W	UTP22
YDR449C	UTP6
YER082C	UTP7
YGR128C	UTP8
YHR196W	UTP9
YBR127C	VMA2
YBR116C	YBR116C
YBR190W	YBR190W
YDR115W	YDR115W
YLR249W	YEF3
YEL008W	YEL008W
YJR146W	YJR146W
YLR184W	YLR184W
YML025C	YML6
YFR049W	YMR31
YNL122C	YNL122C
YPL080C	YPL080C
YPL260W	YPL260W

Appendix 3-2: Non-specific mChIP binders observed as part of this study

ORF Name	Protein Name	mChIP Abundance Factor
YMR072W	ABF2	29
YNR016C	ACC1	7
YER036C	ARB1	5
YDR101C	ARX1	6
YJR121W	ATP2	5
YDR299W	BFR2	12
YPL217C	BMS1	16
YOR078W	BUD21	8
YLR175W	CBF5	18
YCL064C	CHA1	6
YGL206C	CHC1	4
YHR052W	CIC1	17
YNL112W	DBP2	6
YOR204W	DED1	20
YDL160C	DHH1	3
YKL078W	DHR2	7
YKR035W-A	DID2	4
YLR129W	DIP2	15
YKL172W	EBP2	21
YMR128W	ECM16	5
YGR271C-A	EFG1	7
YOR133W	EFT1	13
YDR385W	EFT2	4
YLR186W	EMG1	14
YBR247C	ENP1	10
YGR145W	ENP2	13
YMR049C	ERB1	24
YKL182W	FAS1	4
YKL060C	FBA1	4
YLR051C	FCF2	5
YHR089C	GAR1	21
YKL152C	GPM1	4
YMR290C	HAS1	27
YJL033W	HCA4	14
YDR174W	HMO1	45
YMR186W	HSC82	12
YLR259C	HSP60	4
YOR136W	IDH2	3
YLR432W	IMD3	28
YML056C	IMD4	11
YHR148W	IMP3	5
YNL075W	IMP4	10

YLR247C	IRC20	3
YGL173C	KEM1	19
YNL132W	KRE33	37
YNL308C	KRI1	16
YCL059C	KRR1	8
YER127W	LCP5	5
YDL051W	LHP1	4
YFR001W	LOC1	12
YPL004C	LSP1	9
YAL025C	MAK16	10
YIL070C	MAM33	13
YJR144W	MGM101	28
YBR084W	MIS1	3
YJR002W	MPP10	10
YPR112C	MRD1	6
YKL009W	MRT4	12
YDR194C	MSS116	27
YMR080C	NAM7	10
YPL126W	NAN1	26
YDL208W	NHP2	14
YOR206W	NOC2	18
YLR002C	NOC3	13
YPR144C	NOC4	3
YPL093W	NOG1	15
YNR053C	NOG2	5
YHR072W-A	NOP10	11
YNL175C	NOP13	9
YDL213C	NOP6	12
YJL010C	NOP9	3
YDR432W	NPL3	7
YGL111W	NSA1	9
YER126C	NSA2	4
YER006W	NUG1	12
YER165W	PAB1	18
YOR017W	PET127	3
YBL030C	PET9	3
YGR240C	PFK1	3
YCR012W	PGK1	8
YGR086C	PIL1	27
YPL036W	PMA2	4
YOR145C	PNO1	8
YNL055C	POR1	4
YGL120C	PRP43	50
YDR496C	PUF6	19
YLR196W	PWP1	14

YCR057C	PWP2	24
YGR280C	PXR1	4
YAL036C	RBG1	3
YOL010W	RCL1	11
YBR049C	REB1	12
YOR207C	RET1	27
YAR007C	RFA1	51
YNL312W	RFA2	16
YJL173C	RFA3	33
YOR217W	RFC1	11
YJR068W	RFC2	15
YNL290W	RFC3	21
YOL094C	RFC4	8
YBR087W	RFC5	8
YCR028C-A	RIM1	42
YCL028W	RNQ1	12
YGL171W	ROK1	4
YJR063W	RPA12	14
YDR156W	RPA14	9
YJL148W	RPA34	5
YOR340C	RPA43	38
YOL005C	RPB11	3
YOR151C	RPB2	26
YJL140W	RPB4	5
YBR154C	RPB5	23
YOR224C	RPB8	22
YGL070C	RPB9	6
YHR143W-A	RPC10	13
YDR045C	RPC11	4
YJL011C	RPC17	5
YNL113W	RPC19	19
YKL144C	RPC25	3
YNL151C	RPC31	5
YNR003C	RPC34	8
YKR025W	RPC37	12
YPR110C	RPC40	49
YDL150W	RPC53	14
YPR190C	RPC82	32
YDL140C	RPO21	29
YPR187W	RPO26	5
YOR116C	RPO31	44
YBL014C	RRN6	3
YDR087C	RRP1	7
YPL012W	RRP12	8
YKL082C	RRP14	6

YPR143W	RRP15	5
YDR412W	RRP17	3
YPR137W	RRP9	15
YLR180W	SAM1	5
YDL153C	SAS10	17
YJL080C	SCP160	5
YLR430W	SEN1	11
YNL007C	SIS1	10
YEL026W	SNU13	22
YLL011W	SOF1	11
YCL054W	SPB1	9
YKR092C	SRP40	32
YDR172W	SUP35	4
YJR009C	TDH2	30
YKR059W	TIF1	14
YOL006C	TOP1	35
YNL088W	TOP2	53
YJR041C	URB2	15
YKL099C	UTP11	3
YBL004W	UTP20	9
YOR004W	UTP23	5
YLR373C	VID22	12
YKR001C	VPS1	48
YAR010C	YAR010C	8
YBR012W-B	YBR012W-B	4
YCR016W	YCR016W	4
YNL064C	YDJ1	3
YGR027W-B	YGR027W-B	3
YGR038C-B	YGR038C-B	10
YGR283C	YGR283C	7
YHR214C-B	YHR214C-B	8
YLR003C	YLR003C	3
YML039W	YML039W	3
YML045W	YML045W	10
YOR060C	YOR060C	4
YPR158W-B	YPR158W-B	5
YDR381W	YRA1	54
YOR272W	YTM1	15

Appendix 3-3: Current literature data concerning the number of protein-protein interaction known for the 64 proteins manually curated in SGD to bind to DNA.

Gene Name	ORF #	# of protein-protein interaction identified by AP-MS*	Protein Function**
ARG80	YMR042W	0	transcription factor
AZF1	YOR113W	0	transcription factor
GSM1	YJL103C	0	transcription factor
HST3	YOR025W	0	lysine deacetylase
HST4	YDR191W	0	lysine deacetylase
MET28	YIR017C	0	transcriptional activator
MET32	YDR253C	0	transcriptional activator
MGA1	YGR249W	0	transcription factor
MIG3	YER028C	0	transcriptional repressor
PDR8	YLR266C	0	transcription factor
YOX1	YML027W	0	transcriptional repressor
BRF1	YGR246C	1	transcriptional activator
GCR1	YPL075W	1	transcriptional activator
GIS1	YDR096W	1	histone demethylase
MET31	YPL038W	1	transcriptional activator
MOT3	YMR070W	1	transcription factor
MSH4	YFL003C	1	meiotic recombination
NRG1	YDR043C	1	transcriptional repressor
PZF1	YPR186C	1	transcriptional activator
RGT1	YKL038W	1	transcription factor
TEA1	YOR337W	1	transcriptional activator
YFL052W	YFL052W	1	unknown
YHP1	YDR451C	1	transcriptional repressor
ZAP1	YJL056C	1	transcription factor
CIN5	YOR028C	2	transcription factor
RAP1	YNL216W	2	chromatin silencing
VHR1	YIL056W	2	transcriptional activator
ABF1	YKL112W	3	transcriptional activator and repressor
CBF1	YJR060W	3	transcription factor
RLM1	YPL089C	3	transcription factor
CRP1	YHR146W	4	unknown
HDA3	YPR179C	4	histone deacetylase
MCM1	YMR043W	4	transcription factor
YHM2	YMR241W	4	mitochondrial genome maintenance
MBP1	YDL056W	5	transcription factor
STB4	YMR019W	5	unknown
HDA2	YDR295C	6	histone deacetylase
TBF1	YPL128C	6	transcription factor

MEC3	YLR288C	7	DNA repair
SWI4	YER111C	8	transcription factor
CTF4	YPR135W	9	DNA repair and replication
XRS2	YDR369C	9	DNA repair
OAF1	YAL051W	10	transcriptional activator
HHO1	YPL127C	11	Histone
INO4	YOL108C	13	transcription factor
MGM101	YJR144W	13	mitochondrial genome maintenance
RSC30	YHR056C	15	Chromatin remodeling
CHL4	YDR254W	16	Chromosome segregation
UME6	YDR207C	16	transcription regulation
REF2	YDR195W	23	transcription
ABF2	YMR072W	24	mitochondrial DNA replication and recombination
SPT2	YER161C	25	transcription regulation
HHT2	YNL031C	31	Histone
MHR1	YDR296W	32	homologous recombination
HHF2	YNL030W	37	Histone
HTA1	YDR225W	41	Histone
RSC3	YDR303C	41	Chromatin remodeling
HHT1	YBR010W	46	Histone
HTB1	YDR224C	54	Histone
YKU80	YMR106C	63	DNA repair
HTB2	YBL002W	86	Histone
HTA2	YBL003C	117	Histone
STM1	YLR150W	119	Translation
HHF1	YBR009C	140	Histone

* Proteins with GO annotation #0003677 (DNA binding, manual curation) obtained from SGD on 15/03/10. Number of protein-protein interaction obtained from the *BioGRID* (www.thebiogrid.com) on 15/03/10.

** Data obtained from SGD (www.yeastgenome.org) compiled on 15/03/10.

Appendix 3-4: Baits tested as part of this study.

Bait studied	Detection by Western blot from cell extract (anti-TAP)	Detection by mass spectrometer following mChIP
Aft1	yes	yes
Aft2	no	no
Asf1	yes	yes
Asf2	no	no
Ash1	yes	yes
Azf1	yes	yes
Bre1	yes	yes
Bre5	yes	yes
Bub2	yes	yes
Cac1	yes	yes
Cbf1	no	no
Cdc14	yes	yes
Cdc45	yes	yes
Chz1	yes	yes
Cin5	yes	no
Cnm67	yes	yes
Cpr1	yes	yes
Crp1	yes	yes
Cse4	yes	yes
Ctf18	yes	yes
Ctf4	yes	yes
Cti6	yes	yes
Dot1	yes	yes
Eco1	yes	yes
Elp3	yes	yes
Esa1	yes	yes
Ess1	yes	yes
Fpr2	yes	yes
Fpr3	yes	yes
Fpr4	yes	yes
Gcn5	yes	yes
Gcr1	no	no
Gis1	yes	yes
Glc7	yes	yes
Gsm1	yes	no
Hap2	yes	yes
Hat2	no	no
Hda1	yes	yes
Hhf1	yes	yes
Hho1	yes	yes
Hht1	yes	yes
Hif1	yes	yes

Hir1	yes	yes
Hos1	yes	yes
Hos2	yes	yes
Hos3	yes	yes
Hst3	yes	yes
Hst4	yes	yes
Htb1	yes	yes
Htz1	yes	yes
Iml3	yes	yes
Ixr1	no	no
Kog1	no	no
Lge1	yes	yes
Lst8	no	no
Mbp1	yes	yes
Mcm1	yes	yes
Met28	no	no
Met31	yes	no
Met32	yes	yes
Mga1	yes	no
Mig3	yes	yes
Mot3	yes	yes
Msn2	yes	yes
Msn4	yes	yes
Nap1	yes	yes
Ndc10	yes	yes
Nrg1	yes	no
Pah1	no	no
Pdr8	no	no
Pob3	yes	yes
Pol30	no	no
Rad53	yes	yes
Rad6	yes	yes
Rad9	no	no
Rap1	yes	yes
Rco1	yes	yes
Rgt1	yes	yes
Rlm1	yes	yes
Rpd3	no	no
Rrd1	yes	yes
Rrd2	yes	yes
Rsc4	yes	yes
Rsc58	yes	yes
Rtt106	yes	yes
Rtt109	yes	yes
Sap185	yes	yes
Sas2	yes	yes

Sas3	yes	yes
Sch9	yes	yes
Scm3	yes	yes
Set1	yes	yes
Set2	yes	yes
Set3	yes	yes
Set5	yes	yes
Set6	yes	no
Sfp1	yes	yes
Sgf73	no	no
Sin3	yes	yes
Sir2	yes	yes
Sir3	yes	yes
Sir4	no	no
Slx5	yes	yes
Slx8	yes	no
Spc24	yes	yes
Spc72	yes	yes
Spt10	yes	yes
Spt16	yes	yes
Spt4	no	no
Spt5	yes	yes
Spt6	yes	yes
Stb4	no	no
Stu2	yes	yes
Swd2	yes	yes
Swd3	yes	yes
Swi4	yes	yes
Swi6	yes	yes
Swr1	yes	yes
Taf1	yes	yes
Tco89	no	no
Tea1	yes	yes
Tor1	yes	yes
Tor2	yes	no
Ubp3	yes	yes
Vps72	yes	yes
Vps75	yes	yes
Yfl052w	no	no
Yhp1	yes	yes
Yox1	yes	yes
Yta7	yes	yes

Appendix 3-5: High-confidence mChIP-MS data

mChIP bait*	Associated preys
AFT1	ADA2, AFT1, ARP8, ARP9, CDC14, EAF1, GRX3, HHF1, HHO1, HHT1, HTA2, HTB2, HTZ1, INO80, ISW1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MET18, MOT1, NET1, NGG1, NPL6, RAD3, RGR1, RSC2, RSC3, RSC6, RSC8, RVB1, RVB2, SGF29, SGF73, SNF12, SNF2, SNF5, SPT15, SPT20, SPT3, SPT7, SPT8, SSL2, STH1, SWI1, SWI3, SWP82, TAF10, TAF12, TAF2, TAF5, TAF6, UBP8, YGR071C, YLR278C
ASF1 (1)	ARP4, ARP8, ASF1, CDC14, CKA1, CKA2, CKB1, CKB2, DPB4, FOB1, HHF1, HHO1, HIR1, HIR2, HIR3, HPC2, HTA2, HTB2, HTZ1, IES1, IES3, INO80, ISW2, ITC1, KAP95, MCM6, MNR2, NET1, NPL6, PDR1, PHO2, RAD53, RSC2, RSC3, RSC4, RSC8, RTT106, RVB1, RVB2, SNF7, SRP1, STH1, TOF2, YOL086W-A
ASF1 (2)	ANP1, ARP4, ARP5, ARP7, ARP8, ASF1, ASG1, CDC14, CDC31, CKA1, CKA2, CKB1, CKB2, CSM1, DID2, DPB4, FOB1, HFD1, HHF1, HHO1, HHT1, HIR1, HIR2, HIR3, HPC2, HTA2, HTB2, HTZ1, IES1, IES3, INO80, ISW1, ISW2, ITC1, KAP95, MNR2, MOT1, MPH1, MSH2, NET1, NHP10, NPL6, PDC2, PDR1, PHO2, RAD53, RFM1, RPO41, RSC58, RSC6, RSC8, RSC9, RTT102, RTT106, RVB1, RVB2, SAS2, SAS4, SLD3, SNF12, SNF5, SNF7, SPC110, SPC29, SPT15, STH1, SUM1, SWI3, TAF12, TAF14, TAF5, TAF6, TBF1, TFB2, TFC4, TOF2, TRI1
ASF1 (3)	ARP4, ARP5, ARP8, ARP9, ASF1, BDP1, BRF1, CDC14, CDC31, CHD1, CKA1, CKA2, CKB1, CKB2, DLS1, DPB4, FOB1, HHF1, HHO1, HHT1, HIR1, HIR2, HIR3, HOS3, HPC2, HTA2, HTB2, HTZ1, IES1, IES3, IES4, INO80, IOC2, ISW1, ISW2, ITC1, KAP95, KOG1, MCM2, MCM4, MCM7, MIP1, MNR2, MOT1, MSH2, MSH6, NET1, NHP10, NHP6A, NPL6, ORC1, ORC3, ORC4, PDC2, PDR1, PHO2, RAD52, RAD53, RLF2, RRN10, RRN11, RRN3, RRN5, RRN9, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT106, RVB1, RVB2, SAS2, SAS4, SAS5, SIR2, SLD3, SNF12, SNF5, SNF7, SPT5, SRP1, STH1, SUM1, TAF10, TAF14, TAF5, TBF1, TFC4, TFC6, TFC7, TOA2, TOF2, TOR1, TRA1, TRI1, UAF30, YEL007W, YGR071C, YPR137C-B, YPR204W, YTA7
ASH1	ASH1, CTI6, DEP1, PHO23, RPD3, RXT2, SDS3, SIN3, UME1
AZF1	AZF1, CDC14, HHF1, HHT1, HTA2, HTB2, NET1, NPL6, RAP1, STH1, SWI6
BRE1 (1)	AKL1, BRE1, CBK1, CDC14, CDC31, CDC33, IPP1, KIN2, KIN3, KOG1, KSP1, LGE1, LST8, MOB2, NET1, NUD1, PPN1, PSP2, RAI1, RTG2, SDC1, SPC110, SPC29, SPC42, SSD1, TCO89, TOR1
BRE1 (2)	BFA1, BRE1, CDC14, CDC31, HHF1, HHO1, HHT1, HTA2, HTB2, IMD1, ISW1, KOG1, LGE1, LST8, MLC1, MOT1, MYO2, MYO4, NDC1, NET1, NUD1, RAD3, RAD52, SNF5, SPC110, SPC29, SPC42, SSD1, STH1, TAF10, TAF12, TAF14, TAF5, TFB2, TOR1
BRE5	BFR1, BRE5, CDC14, CLU1, HHF1, HTA2, HTB2, RLI1, SSD1, UBP3
BUB2	BFA1, BUB2, CDC31, CMD1, CNM67, NIC96, NSP1, NUD1, SPC110, SPC29, SPC42, SPC72, SPC97, SPC98
CAC1 (1)	ARP4, ARP9, CAC2, CDC14, EAF3, HHF1, HHO1, HHT1, HIR2, HTA2, HTB2, IOC3, ISW1, MSI1, NET1, NPL6, ORC1, POL30, RLF2, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RTT106, RVB1, RVB2, SNF12, SNF5, STH1, SUM1, SWI3, TOF2, TRI1, YGR126W
CAC1 (2)	ABF2, ARF1, ARP4, ARP5, ARP7, ARP8, ARP9, BDF1, BDP1, BFR2, BMS1, BRF1, BRN1, CAC2, CDC14, CHA1, CHD1, CIC1, CKA2, CSM1, DED1, DHR2, DID2, DID4, DIG1, DIP2, DPB4, EAF3, EAF5, ECM16, EFG1, EFT1, EMG1, ENP2, ERB1, ESC8, FCF2,

	FOB1, HAS1, HAT1, HAT2, HCA4, HHF1, HHO1, HHT1, HIR1, HIR2, HIR3, HMO1, HPC2, HSC82, HTA2, HTB2, HTZ1, IES1, IES3, IES4, IES5, IMP3, IMP4, INO80, IOC2, IOC3, IST1, ISW1, ISW2, ITC1, KAP95, KRE33, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MGM101, MIP1, MOT1, MPP10, MSH2, MSH3, MSH6, MSI1, MSS116, NAN1, NET1, NHP10, NHP2, NOC2, NOP9, NPL6, OAF3, ORC1, ORC2, ORC3, ORC4, ORC5, PDB1, PDC2, PDR1, PGK1, PNO1, POL30, PRP43, PUF6, PWP1, PWP2, RAD3, RAD52, RAP1, RCO1, REB1, RET1, RFA1, RFA2, RFA3, RFC1, RFC2, RFC3, RFC4, RFC5, RFM1, RIM1, RLF2, RPA12, RPA14, RPA43, RPB11, RPB2, RPB5, RPB8, RPB9, RPC11, RPC17, RPC19, RPC25, RPC34, RPC37, RPC40, RPC82, RPO21, RPO26, RPO31, RRP9, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RTT106, RVB1, RVB2, SAR1, SFH1, SGF29, SGF73, SIR2, SIS1, SNF12, SNF2, SNF5, SNF7, SNU13, SPT15, SPT16, SPT20, SRP1, SRP40, SSL2, STH1, SUA7, SUM1, SWC5, SWI1, SWI3, SWP82, TAF10, TAF12, TAF14, TAF3, TAF4, TAF5, TAF6, TAF9, TBF1, TFB1, TFB4, TFC4, TFC6, TFC7, TFC8, TIF1, TOA2, TOF2, TOP1, TOP2, TRA1, TRI1, VID22, VPS1, YAF9, YAP1, YCG1, YCS4, YDL156W, YER064C, YGR071C, YGR283C, YML045W, YMR310C, YOR060C, YPR204W, YRA1, YTA7
CDC14	BFA1, BUB2, CDC14, CDC31, CMD1, CNM67, FOB1, HHF1, HTB2, HTZ1, IMD1, NAP1, NET1, NIC96, NUD1, NUP170, PSE1, SPC110, SPC29, SPC42, SPC72, SPC97, SPC98, TOF2, YBL005W-B
CDC45	CDC45, CDC53, CTF4, DIA2, HHF1, HTZ1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MRC1, SKP1, SLD5, SPT16, TOF1
CHZ1	CHZ1, HTZ1
CNM67	CNM67, HHF1, NAP1, NUD1, SPC110
CNS1	AHA1, ARF2, BMH1, CCT2, CCT3, CCT4, CCT5, CCT6, CCT7, CCT8, CDC48, CNS1, DBP5, FAS2, GCD11, GCN1, HEK2, HGH1, HHF1, HOG1, HSP104, ILV6, KAP123, KAP95, LSC1, LYS12, NEW1, OLA1, PFK2, RNA1, RNR1, RNR2, RPN10, RPN7, SAC6, SAR1, SEC21, SEC23, SPT5, SRV2, SUB2, SUP45, TCP1, TFP1, TRP5, TRX2, TSA1, TUB1, TUB2, URA7, VMA13, VMA5, YKT6, YPK2
CPR1	AKL1, ASK10, AVO1, BBP1, BEM2, BFA1, BRE1, BUB2, CBK1, CDC14, CDC31, CNM67, CPR1, EFR3, HHF1, HOS2, HOS4, HTZ1, KIN2, KOG1, KSP1, LGE1, LST8, MLC1, MOB2, MPS2, MYO1, MYO2, MYO4, NET1, NUD1, NUP170, RSP5, SDC1, SEC26, SET3, SIF2, SNT1, SOD2, SPC110, SPC29, SPC42, SPC97, SPC98, SSD1, SUP45, TAO3, TCO89, TOR1, TOR2, YHR127W, YKT6
CRP1	CKB1, CRP1, FPR3, FPR4, GLC7, GSY1, GSY2, HHF1, HTA2, HTD2, NAP1, PEP4, PIG2, PRB1, PRC1, RSE1, YPP1
CTF18	CTF18, CTF8, DCC1, DPB2
CTF4	CDC45, CTF4, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MRC1, PSF2, PSF3, SLD5
CTI6	CTI6, DEP1, DOT6, RPD3, RXT2, SIN3, UME1
DOT1	DOT1, FPR3, HHF1
ECO1 (1)	ECO1
ECO1 (2)	ECO1, FPR3, HHF1
ELP3	ELP2, ELP3, ELP4, ELP6, HRR25, IKI1, IKI3, KTI12
ESA1	ACT1, ARP4, BDF1, CDC14, CHD1, EAF1, EAF3, EAF6, EPL1, ESA1, GDS1, HHF1, HHO1, HTA2, HTB2, HTZ1, MCM4, MSN4, NET1, RVB1, RVB2, SPT16, SWC4, TRA1, YAF9, YNG2
ESS1	ASM4, ASR1, ESS1, FPR3, GLE2, HHF1, MEX67, NIC96, NSP1, NUP100, NUP120,

	NUP133, NUP145, NUP159, NUP170, NUP188, NUP192, NUP49, NUP57, NUP82, NUP84, NUP85, RPO21, SXM1
GCN5	ADA2, AHC1, ARP7, BDF1, CDC14, CHD1, CRZ1, EAF1, GAL11, GAT1, GCN5, HFI1, HHF1, HHO1, HHT1, HSF1, HTA2, HTB2, HTZ1, ISW1, MED1, MED4, MED6, MED7, MED8, MGA2, MOT1, MSN2, MSN4, NAP1, NET1, NGG1, NPL6, NUT1, RGR1, RSC2, RSC30, RSC58, RSC6, RSC8, RSC9, RTG1, RTG3, RTT102, RVB2, SGF29, SGF73, SIN4, SNF12, SNF5, SPC29, SPT15, SPT20, SPT7, SPT8, SSN3, SSN8, STH1, SUM1, SWI3, SWI5, TAF10, TAF12, TAF14, TAF2, TAF5, TAF6, TAF9, TRA1, TRI1, YAP1, YML081W, YTA7
FPR2	AKL1, BBP1, BEM2, BFA1, BRE1, BUB2, CBK1, CDC14, CDC31, CDC33, CMD1, CNM67, DIP2 , FPR2, HHF1, HTZ1, KEL1, KIN2, KIN4, KOG1, KSP1, LGE1, LST8, MLC1, MOB2, MPS2, MYO2, NET1, NUD1, NUP192, PPN1, RTG2, SDC1, SEH1, SNF12, SPC110, SPC29, SPC42, SPC97, SPC98, SPT5, SSD1, TAO3, TCO89, TOR1, YBL104C
FPR3	BUD20, BUD22, CBF5, CKA1, CKA2, CKB1, FPR3, FPR4, GLC7, HHF1, HTZ1, KAP123, MTR4, NET1, NOP53, RIX7, RRB1, RRS1, SSA4, STT4, UTP7, YMR086W, YMR310C, YOR093C, YPP1
FPR4	AVO1, BUD20, CKA1, CKA2, CKB1, FPR3, FPR4, GLC7, HHF1, HTA2, HTB2, PPZ1, STT4, TOR2, YIL091C, YMR310C, YNL022C, YPP1
GIS1	ARP4, CDC14, FPR4, GIS1, HHF1, HHT1, HTB2, HTZ1, ISW1, ISW2, ITC1, MCM2, MCM4, MCM5, MCM6, NET1, NPL6, RSC2, RSC3, RSC6, RSC8, RSC9, RVB1, RVB2, SNF5, STH1, SWI3, SXM1, TAF12, TAF5, TAF6, TBF1
GLC7	ADE16, AIR1, AVO1, BMH1, BNI4, BUD14, CDC10, CDC11, CDC12, CDC3, CFT1, CFT2, CKA1, CKA2, CKB1, CKB2, FIP1, FPR3, FPR4, GIP3, GIP4, GLC7, HHF1, HTB2, IMP2, JIP4, KEL1, MHP1, MPE1, PAP1, PEP3, PEP5, PPZ1, PTA1, PTI1, REF2, SDS22, SEC1, SEC9, SHS1, SSD1, SSU72, SWD2, TOR2, VAM6, VPS41, YER158C, YGR237C, YHR127W, YMR310C, YOR227W, YPI1, YPP1, YSH1
HAP2	ADA2, ARP8, ARP9, ASF1, CDC14, CHD1, DAL81, EAF1, EPL1, FOB1, GLN3, HAP2, HAP3, HAP5, HFI1, HHF1, HHO1, HHT1, HIR2, HRR25, HTA2, HTB2, HTZ1, IKI3, INO80, IOC2, IOC3, ISW1, ISW2, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, NET1, NGG1, NPL6, RFM1, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SGF29, SGF73, SNF12, SNF2, SNF5, SPT15, SPT20, SPT3, SPT5, SPT6, SPT7, SPT8, STE12, STH1, STP1, STP2, SWI3, SWP82, TAF1, TAF10, TAF12, TAF14, TAF4, TAF5, TAF6, TAF8, TAF9, TBF1, TRA1, YAP5, YTA7
HDA1	FPR3, HDA1, HDA2, HDA3, HHF1, HTA2, HTB2
HHF1	ARP4, ARP8, ARP9, ASE1, CDC14, CHD1, CKA1, CKA2, CKB1, CKB2, DIS3, DPB4, EAF1, EAF3, EPL1, ESF1, FPR3, FPR4, GCD11, HAT1, HAT2, HHF1, HHO1, HHT1, HIF1, HIR1, HIR2, HIR3, HTA2, HTB2, HTZ1, INO80, IOC2, IOC3, ISW1, ISW2, ITC1, KAP123, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, NET1, NPL6, ORC1, ORC2, ORC4, ORC6, POB3, POL5, PSH1, RAP1, REP1, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT106, RVB1, RVB2, SAS3, SGD1, SGF73, SNF5, SPT16, SPT5, STH1, SWC4, SWI3, TAF10, TAF12, TAF4, TAF5, TAF6, TRA1, YDL156W, YER064C, YMR310C, YTA7
HHO1	ARP4, ARP7, ARP9, BDF1, CDC14, CHD1, EAF3, EPL1, FPR3, FPR4, GCD11, HHF1, HHO1, HHT1, HIR2, HTA2, HTB2, HTZ1, IMD2, INO80, IOC2, IOC3, IOC4, ISW1, ITC1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, NET1, NPL6, ORC1, ORC5, POB3, REP1, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SFH1, SIN3, SPT16, SPT5, SSL2, STH1, STM1, SUM1, TAF12, TAF14,

	TAF5, TAF6, TBF1, TFB2, TRA1, YDL156W, YPP1, YTA7
HHT1	ADA2, ARP4, ARP7, ARP8, ARP9, BDF1, BDP1, BRF1, BUD22, CDC14, CHD1, CKA1, CKA2, CKB1, CKB2, DBP8, EAF3, EPL1, ESF1, FOB1, FPR3, FPR4, GCD11, GCN5, GLC7, HHF1, HHO1, HHT1, HIR2, HTA2, HTB2, HTZ1, INO80, IOC2, IOC3, IOC4, IRR1, ISW1, ISW2, KAP123, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, MTR4, NAP1, NET1, NPL6, NTO1, ORC3, POB3, POL30, RAF1, RAP1, RCO1, REP1, RLF2, RRS1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SGD1, SMC3, SPT15, SPT16, SPT5, STH1, SUI2, SUI3, SUM1, SWC4, TAF12, TAF14, TAF4, TAF5, TAF6, TFC4, TRA1, UTP30, YIL127C, YTA7
HIF1	HAT1, HAT2, HIF1, HHF1, SPT16
HIR1 (1)	CDC14, HHF1, HIR1, HIR2, HIR3, HPC2, HTA2, RVB2
HIR1 (2)	ARP8, ARP9, CDC14, DPB4, HHF1, HHT1, HIR1, HIR2, HIR3, HPC2, HTA2, HTB2, IES4, INO80, ISW2, ITC1, NET1, RSC8, RTT106, RVB1, RVB2, SWI3, TFC4
HOS1	HOS1
HOS2	AKL1, CCT2, CCT3, CCT4, CCT5, CCT6, CCT7, CCT8, CDC14, CPR1, HHF1, HOS2, HOS3, HOS4, HTB2, ISW1, KOG1, MCM6, MKT1, NET1, RVB1, SET3, SIF2, SNT1, TCP1, TOF2, TOS4, YBL005W-B
HOS3	ARP9, CDC14, CDC31, CHD1, FPR3, HHF1, HHO1, HHT1, HOS3, HTA2, HTB2, HTZ1, IOC4, IRR1, ISW1, KOG1, NET1, NPL6, NUP57, REP1, RRN11, RRN7, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RVB1, RVB2, SPC110, SPT5, STH1, TAF12, TAF5, TAF6
HST3	HHF1, HST3, KAP95
HST4	HST4, NAP1, NMD5, PEP4, SXM1
HTA2	ARP4, ARP8, ARP9, ATP1, BMH1, CDC48, CHD1, CKA2, COP1, EAF3, FPR3, FPR4, GCN1, HHF1, HHO1, HHT1, HIR2, HOS3, HSN1, HSP26, HTA2, HTB2, HTZ1, IDH1, INO80, , IOC2, ISW1, ISW2, ITC1, KAP114, KAP123, LST8, MCM2, MCM3, MCM4, MCM5, MCM7, MIR1, MOT1, MSH2, NAP1, NET1, NPL6, NTO1, OLA1, ORC1, PAD1, PCK1, PGI1, POB3, POX1, PRP4, PSA1, PSH1, RSC2, RSC3, RSC4, RSC58, RSC6, RSC9, RVB1, RVB2, SAC6, SAS3, SFC1, SFH1, SPT15, SPT16, SPT5, STH1, STT4, TAF14, TSA1, UAF30, YER138C, YKT6, YLR455W, YTA7
HTB1	ADA2, ARP4, ARP5, ARP7, ARP8, ARP9, ASG1, BDF1, BDP1, BMH1, BRF1, BUR6, CAD1, CCL1, CDC14, CDC31, CHD1, CKA2, CKB1, CSE4, CSM1, DBP8, DIG1, DPB4, EAF1, EAF3, EAF5, EPL1, ESA1, FOB1, FPR3, FPR4, GAT1, GCN5, HAT1, HHF1, HHO1, HHT1, HIR1, HIR2, HOS3, HPC2, HST1, HTA2, HTB2, HTZ1, IES1, IES2, IES3, INO80, IOC2, IOC3, IOC4, IRR1, ISW1, ISW2, ITC1, MCD1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MED4, MOT1, MSN1, NCB2, NET1, NGG1, NHP10, NIC96, NPL6, NTO1, NUF2, NUT1, NUT2, OAF3, ORC1, ORC2, ORC3, ORC4, ORC5, ORC6, PDR1, PDS5, POB3, POL30, PSH1, RAD16, RAD3, RAD52, RAF1, RAP1, RCO1, REP1, RFM1, RPB3, RPD3, RPH1, RRN7, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RTT106, RVB1, RVB2, SAS3, SFH1, SGF29, SIN4, SIR2, SIR3, SKN7, SMC1, SMC2, SMC3, SNF12, SNF5, SNF6, SPC29, SPC42, SPN1, SPT15, SPT16, SPT3, SPT5, SPT6, SRB2, SRB4, SSL1, SSL2, STE12, STH1, SUA7, SUM1, SWC3, SWC4, SWC5, SWI1, SWI3, SWP82, SWR1, TAF1, TAF10, TAF11, TAF12, TAF14, TAF2, TAF3, TAF4, TAF5, TAF6, TAF7, TAF8, TAF9, TBF1, TFA1, TFA2, TFB1, TFB2, TFB3, TFB4, TFC4, TFC7, TFG1, TOA1, TOA2, TOF2, TRA1, TRI1, UME1, YAF9, YCS4, YDL156W, YER064C, YGR071C, YIL127C, YLR278C, YLR455W, YMR310C, YNG2, YNL050C, YOL103W-B, YTA7

HTZ1 (1)	ADE5, ARP4, ARP6, CHD1, FAA4, FAS2, FPR3, FPR4, GAS1, GCN1, HHF1, HHO1, HHT1, HTA2, HTB2, HTZ1, ILV5, IOC2, ISW1, ISW2, ITC1, KAP114, KAP123, MCM2, MCM4, MCM7, MIR1, MOT1, NAP1, NET1, RSC2, RSC3, RSC4, RVB1, RVB2, SHM2, SMC3, STH1, SWC4, SWC5, SWR1, TAF14, VPS71, YER138C, YTA7
HTZ1 (2)	ARP4, ARP6, ARP7, ARP8, ARP9, BDF1, BDF2, CDC14, CHD1, CKB1, EAF3, FPR3, FPR4, HHF1, HHO1, HOS3, HTA2, HTB2, HTZ1, IOC2, IOC3, IOC4, ISW1, ISW2, ITC1, KAP114, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, MPS3, NAP1, NET1, NPL6, ORC3, POB3, RAF1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SPT15, SPT16, STH1, SWC3, SWC4, SWC5, SWR1, TAF10, TAF12, TAF14, TAF4, TAF5, TAF6, TBF1, TRA1, VPS71, VPS72, YPP1, YTA7
IML3	AME1, CDC14, CHL4, CTF19, CTF3, DAM1, ELP2, ELP3, IKI3, IML3, MCM21, MPS2, NIC96, NUD1, NUP49, NUP53, NUP57, OKP1, SPC110, SPC42, SPC97, SPC98, TID3
LGE1	AKL1, ASK10, BEM2, BFA1, BOI1, BRE1, BRE2, BUB2, CBK1, CDC14, CDC31, CDC33, CDC48, CNM67, DBP1, FIR1, GLC7, HHF1, HOS4, HRB1, HTA2, ICS2, IPP1, KAR1, KEL1, KIN2, KIN3, KIN4, KOG1, KSP1, LGE1, LRE1, LST8, MOB2, MRM1, MSH2, MSH3, MSH6, MUD2, MYO2, NET1, NSP1, NUD1, PAM1, PSP2, RAD52, RGA1, ROM2, SDC1, SET1, SIF2, SNT1, SOD2, SPC110, SPC29 ,SPC42, SPC72, SPC97, SPC98, SPP1, SPT5, SRO9, SRP1, SSD1, STT4, TAO3, TCO89, TOF2, TOR1, TUB4, YBL104C, YPL141C, YPP1
MBP1	BMH1, CDC14, CDC48, HHF1, HHT1, HIR2, HTA2, HTB2, ISW1, MBP1, MCM2, NET1, NRM1, RSC6, RTT102, RVB2, SNF12 , STB1, SWI6, TAF12, TAF5, TFC6, YMR144W
MCM1	ARG81, ARP9, BCK2, CBK1, CDC14, DIG1, FOB1, GZF3, HHF1, HTZ1, IOC3, ISW1, MBP1, MCM1, MCM5, MCM6, MOT1, NET1, NMD5, NPL6, POG1, RSC2, RSC3, RSC4, RSC58, RSC8, RSC9, RVB1, RVB2, SKN7, SRL2, STE12, STH1, SUM1, SWI4, SWI6, TAF12, TAF6, TBF1, YOX1
MCM5	ARP9, CDC14, CDC45, CTF4, HHF1, HTA2, HTB2, ISW1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MRC1, NET1, RSC8, SAP185, SPT16, TOF1, YAR009C, YTA7
MOT3	CDC14, HHF1, HHO1, HHT1, HTB2, HTZ1, ISW1, ITC1, MOT3, NET1, RSC30, RSC8, RVB1, SPT15, STH1, YJR029W
MSN2	ACT1, ADA2, ARP9, CDC14, EAF1, GCN5, HFI1, HHF1, HHO1, MSN2, NET1, NGG1, NPL6, RVB2, SGF29, SGF73, SNF12, SPT15, SPT3, SPT7, SPT8, STE23, SWI1, SWI3, TAF10, TAF12, TAF4, TAF5, TAF6, TAF9, TRA1, UBP8
MSN4	ACT1, ADA2, ARP4, ARP9, BDF1, BMH1, CDC14, EAF1, EAF3, EPL1, ESA1, GCN5, GDS1, HFI1, HHF1, HHO1, MSN4, NET1, NGG1, RSC2, RSC3, RSC58, RSC8, RTT102, RVB1, SGF29, SGF73, SNF12, SNF5, SPT15, SPT20, SPT3, SPT7, SPT8, SWC4, SWI1, SWI3, SWP82, TAF1, TAF10, TAF11, TAF12, TAF2, TAF3, TAF4, TAF5, TAF6, TAF8, TAF9, TRA1, YNG2
NAP1	BMH1, BNI5, CDC10, CDC11, CDC12, CDC3, GIN4, KCC4, KOG1, MLC1, MYO2, NAP1, NBA1, SHS1, TOR1
NDC10	ARP9, CBF2, CDC14, CEP3, CTF13, FOB1, HHF1, HHT1, HIR2, HTA2, HTB2, ISW1, MCM6, MSH2, NET1, NHP10, NPL6, RSC6, RSC8, RVB1, RVB2, SKP1, SNF12, SPT15, STH1, SUM1, SWI3, TAF12, TAF14, TAF5, TBF1, TOF2
POB3	ARP9, CDC14, CDC73, CHD1, CKA1, CKA2, CKB1, CKB2, CTR9, HHF1, HTA2, HTZ1, ISW1, LEO1, MCM2, MCM5, NET1, NGG1, NPL6, PAF1, POB3, PSH1, RSC3, RSC4, RSC58, RSC6, RSC8, RSC9, RTF1, RTT102, RVB1, SPT16, SPT6, STH1, YNL054W-B, YTA7
RAD53	ASF1, KAP95, RAD53, SRP1

RAD6 (1)	BMH2, BRE1, CDC48, MUB1, RAD6, RAD18, UBR1, UBR2
RAD6 (2)	BRE1, LGE1, MIF2, MUB1, RAD18, RAD6, UBR1, UBR2, YDR186C
RAP1	CDC48, HHF1, MCM2, MCM3, MCM4, MCM5, NPL6, RAP1, RSC2, RSC8, RVB2, SNF12
RCO1	ARP4, ARP9, CDC14, CHD1, EAF3, HHF1, HHO1, HTA2, HTZ1, IOC3, ISW1, NET1, NPL6, RCO1, RPD3, RSC9, SIN3, SPT15, STH1, TAF5, UME1, YTA7
RGT1	CDC14, HHF1, HHO1, HTB2, NET1, RGT1
RLM1	FPR3, FPR4, RLM1
RRD1	FPR3, PPH3, PSY2, PSY4, RRD1, SAP190, SIT4, STM1
RRD2	RRD2, TPD3
RSC4	ADA2, AHC1, ARP4, ARP7, ARP8, ARP9, BDF1, BDP1, BRF1, CDC14, CDC31, CHD1, CSE4, DPB4, EAF1, EAF3, GAT1, GCN5, GLC7, HHF1, HHO1, HHT1, HIR2, HOS3, HTA2, HTB2, HTL1, HTZ1, IES3, IES5, INO80, IOC2, IOC3, IOC4, IRR1, ISW1, ISW2, ITC1, LDB7, MCD1, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOT1, NCB2, NET1, NPL6, NTO1, ORC3, ORC4, ORC5, PDS5, POB3, RAF1, RAP1, REP1, RFM1, RPB7, RPD3, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SAS3, SFH1, SGF29, SIR3, SPC29, SPN1, SPT15, SPT16, SPT5, SPT6, STH1, SUM1, SWR1, TAF10, TAF11, TAF12, TAF3, TAF5, TAF6, TAF9, TBF1, TFB2, TFB4, TFC4, TFC7, TRA1, UBP8, UTP30, YLR278C, YLR455W, YTA7
RSC58	ARP7, ARP8, ARP9, BDF1, CDC14, CHD1, EAF1, FOB1, FPR3, HHF1, HHO1, HHT1, HIR2, HOS3, HTB2, HTL1, HTZ1, IES3, INO80, IOC2, IOC3, ISW1, LDB7, MCM2, MCM3, MCM4, MCM5, MCM6, MOT1, NET1, NPL6, NTO1, POB3, RAP1, RSC1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SAS3, SFH1, SNF7, SPC110, SPN1, SPT15, SPT16, SPT5, SPT8, STH1, TAF10, TAF12, TAF5, TAF6, TAF9, TBF1, TOA2, TRA1, TRI1, YHR127W, YTA7
RTT106 (1)	CAC2, HHF1, HIR1, HIR2, HIR3, HPC2, HTB2, ISW1, ISW2, ITC1, MSI1, POL30, RLF2, RSC3, RTT106
RTT106 (2)	ARP9, CAC2, CDC14, CHD1, DPB4, HHF1, HHO1, HHT1, HIR1, HIR2, HIR3, HPC2, HTA2, HTB2, HTZ1, INO80, IOC2, ISW1, ISW2, ITC1, MCM5, MOT1, NET1, NPL6, ORC4, ORC6, POL30, RLF2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT106, RVB1, RVB2, SPT15, STH1, TAF5, TAF6, YTA7
RTT109 (1)	RTT109, VPS75
RTT109 (2)	RTT109, VPS75
SAS2	MAM33, NAP1, SAS2, SAS4, SAS5
SAS3	HHF1, HHT1, HTA2, HTB2, ISW1, NPL6, NTO1, RSC8, RVB2, SAS3, STH1
SAP185	CDC25, KIN2, MDS3, PMD1, RMD1, RMD8, SAP185, SIT4, TIP41, YBR225W
SCH9 (1)	SCH9
SCH9 (2)	SCH9
SCM3	CSE4, SCM3
SET1	BRE1, BRE2, CDC14, HHF1, HTB2, ISW1, KOG1, MCM2, NET1, NRP1, RSC8, SDC1, SET1, SHG1, SMC3, SPP1, STH1, SWD1, SWD2, SWD3, TFB2
SET2	CDC14, CKA1, CKB1, FPR3, FPR4, HHF1, HTA2, HTZ1, ISW1, POL5, RVB2, SET2, YPP1
SET3	ADA2, ARP4, ARP5, ARP7, ARP8, ARP9, BDP1, BRE1, BRF1, CDC14, CHD1, CKA2, CKB1, CPR1, DID4, DIG1, DPB4, EAF3, EPL1, FOB1, FPR3, FPR4, GCN5, GLC7, HHF1, HHO1, HHT1, HIR1, HIR2, HOS2, HOS3, HOS4, HST1, HTA2, HTB2, HTZ1, IES1, IES2, IES3, INO80, IOC2, IOC3, ISW1, ISW2, ITC1, KAP95, KIN28, KOG1, LST8, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MOB2, MOT1, MRT4 , MSH2, MSH6, NET1,

	NPL6, NUT1, ORC1, ORC4, ORC5, ORC6, PDR1, POB3, RAP1, REP1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SAS3, SET3, SFH1, SIF2, SIR2, SKI6, SNF5, SNT1, SOD2, SPT15, SPT16, SPT20, SPT3, SPT5, SPT6, SRP1, SSL2, STH1, STT4, SUM1, TAF12, TAF4, TAF5, TAF6, TAF9, TBF1, TFB2, TFC4, TFC7, TOF2, TOS4, TRA1, TRI1, VPS24, YDL156W, YER064C, YTA7
SET5	ACT1, CDC31, CMD1, GLC7, HHF1, MLC1, MYO2, NUD1, SET5, SPC110, SPC29, SPC42
SFP1 (1)	BAR1, CDC14, HHF1, MRS6, NET1, SFP1
SFP1 (2)	CKA2, CKB1, HHF1, MRS6, NET1, SFP1
SIN3	ARP4, ARP9, ASH1, BMH1, CDC14, CHD1, CTI6, DEP1, DOT6, EAF3, FPR3, HHF1, HHO1, HHT1, HOS3, HTA2, HTB2, ISW1, ISW2, ITC1, MCM3, MCM5, MCM6, MOT1, NET1, NPL6, PHO23, RCO1, RPD3, RSC3, RSC4, RSC58, RSC6, RSC8, RSC9, RVB1, RVB2, RXT2, RXT3, SAP30, SDS3, SIN3, SPT16, SPT5, SRP1, STH1, SUM1, TAF12, TFC4, TUB2, UME1, UME6, YPR158C-D, YTA7
SIR2	CDC14, FOB1, HHF1, HTA2, NET1, SIR2, SIR4, YNL054W-B
SIR3	CDC14, FPR3, HHF1, HHO1, HHT1, HTA2, HTB2, NET1, NPL6, RSC6, RVB1, RVB2, SIR3
SLX5	HHF1, IMD1, NET1, RVB1, RVB2, SLX5, YMR111C
SPC24	ALY2, ASK1, BBP1, BFA1, BUB2, CDC14, CDC31, CDC5, CNM67, DSN1, FIN1, FPR3, HHF1, HSL7, HTB2, KAP123, KAR1, MPS1, MPS2, MPS3, MTW1, NBP1, NDC1, NIC96, NSL1, NSP1, NUD1, NUF2, NUP120, NUP133, NUP145, NUP170, NUP188, NUP192, NUP49, NUP53, NUP57, NUP82, SFI1, SPC105, SPC110, SPC24, SPC25, SPC29, SPC42, SPC72, SPC97, SPC98, TID3, TUB4, VMA13
SPC72	BBP1, BFA1, CDC14, CDC31, CNM67, KAP95, KAR1, MPS1, NBP1, NIC96, NSP1, NUD1, NUP116, SFI1, SPC110, SPC29, SPC42, SPC72, SPC97, SPC98, TUB4, YLR146W-A
SPT10 (1)	ARP9, CDC14 , HHF1, HIR2, HPC2, HTA2, HTB2, ISW1, NET1, NPL6, RSC2, RSC3, RSC30, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SFH1, SPT10, STH1
SPT10 (2)	ARP9, HAT1, HAT2, HHF1, HTZ1, ISW1, MCM2, MCM3, MCM4, MCM6, MOT1, NET1, NPL6, RSC2, RSC3, RSC30, RSC4, RSC8, RSC9, RTT102, RVB2, SFH1, SPT10, STH1
SPT16	CDC14, CDC73, CKA1, CKA2, CKB1, CKB2, CSE4, CTR9, HHF1, HHO1, HTA2, HTB2, LEO1, NET1, PAF1, POB3, PSH1, RTF1, SPT16, YTA7
SPT5	CDC14, CKA1, CKA2, CKB2, FPR3, HHF1, HHO1, HTA2, HTB2, IOC2, KAP95, LST8, MPS2, MYO2, PSY2, RPB3, RPB7, RVB1, SHE3, SPN1, SPT16, SPT5, SSD1, TAF12, TCP1, TOR1
SPT6	HHF1, SPN1, SPT6
STU2	CDC14, CDC31, CKA2, CNM67, FOB1, FPR3, FPR4, HHF1, HHO1, HTA2, HTB2, HTZ1, IMD2, KAP123, KOG1, LST8, MSH6, NAP1, NET1, NIC96, NOP53, RAD52, RPO41, SEC27, SGS1, SIR2, SLH1, SPC110, SPC29, SPC42, SPC72, SPC97, SPC98, SSD1, STT4, STU2, SUM1, TCP1, TOR1, TSC11, YPP1
SWD2	BRE2, CFT1, CFT2, FIP1, GLC7, HHF1, HHO1, HSP10, HTA2, MPE1, NET1, PAP1, PFS2, PTA1, PTI1, REF2, SDC1, SET1, SSU72, SWD2, SYC1, YSH1
SWD3	BRE1, BRE2, DIG1, KOG1, KSP1, LGE1, SDC1, SET1, SHG1, SPP1, SWD1, SWD2, SWD3, TOR1
SWI4 (1)	BCK2, CDC14, CHD1, DIG1, HHF1, HHT1, HTA2, HTB2, HTZ1, ISW1, MCM4, NAP1, RSC2, RSC6, RSC8, RVB1, STH1, SWI4, SWI6, TAF12, TAF5, TRA1, WHI5

SWI4 (2)	ARP4, ARP7, ARP8, ARP9, BCK2, CDC14, CDC31, CKB1, FOB1, HHF1, HHO1, HHT1, HTA2, HTB2, HTZ1, IES1, ISW1, ISW2, ITC1, MCM1, MCM2, MCM3, MCM4, MCM5, MCM6, MOT1, MSA1, MSA2, NET1, NPL6, PDR1, RAD3, RLM1, RSC2, RSC3, RSC30, RSC4, RSC58, RSC6, RSC8, RSC9, RTT102, RVB1, RVB2, SKN7, SNF12, SNF5, SNF6, SPT15, SPT20, SPT8, STH1, SWC3, SWI3, SWI4, SWI6, SWP82, TAF12, TAF14, TAF5, TAF6, TAF8, TAF9, TBF1, TFB2, TOA2, TRX2, WHI5, YLR278C
SWI6	ARP8, ARP9, CDC14, DIG1, FOB1, HHF1, HHO1, HHT1, HTA2, HTB2, HTZ1, ISW1, MBP1, MCM2, MCM4, MOT1, NET1, NPL6, NRM1, RSC2, RSC3, RSC4, RSC58, RSC6, RSC8, RSC9, RVB1, RVB2, SNF5, SPT15, STB1, STH1, SWI4, SWI6, TAF12, TAF5, TAF6, WHI5, YDR261C-D, YMR144W
SWR1	ACT1, ARP4, ARP6, ARP9, BDF1, CDC14, CHD1, HHF1, HHO1, HHT1, HTB2, HTZ1, ISW1, MCM2, MCM3 , MCM4, MCM7, NET1, PIH1, RSC2, RSC3, RSC58, RSC6, RSC8, RSC9, RVB1, RVB2, STH1, SWC3, SWC5, SWC7, SWR1, TAF12, VPS71, VPS72, YAF9
TAF1	ARP9, BDF1, BDF2, CDC14, HHF1, HHT1, HSF1, HTA2, HTB2, HTZ1, ISW1, ISW2, MCM3, MCM6, NET1, NPL6, RSC2, RSC3, RSC30, RSC4, RSC6, RSC8, RSC9, RVB1, RVB2, SPT15, STH1, TAF1, TAF10, TAF11, TAF12, TAF13, TAF14, TAF2, TAF3, TAF4, TAF5, TAF6, TAF7, TAF8, TAF9, TOA2
TEA1	CDC14, HHF1, HTB2, RSC8, TEA1
TOR1 (1)	RVB2, TOR1
TOR1 (2)	TOR1
UBP3	BFR1, BRE5, CDC14, CLU1, ECM32, HHF1, HTA2, HTB2, MPD1, NET1, RLI1, RPN1, SRO9, SSD1, STT4, SUP45, TMA46, TOR1, TOR2, UBP3, YLR419W, YPP1
VPS72	ACT1, ARP6, HTZ1, RVB1, RVB2, SWC3, SWC4, SWC5, SWR1, VPS71, VPS72
VPS75	RTT109, VPS75
YHP1 (1)	CDC14, HHF1, HTA2, HTB2, KAP95, MSH2, NET1, YHP1
YHP1 (2)	CDC14, HHF1, HHT1, HTA2, HTB2, ITC1, KAP95, MSH2, MSH6, NET1, RAD52, RVB2, SRP1, YHP1
YOX1	CDC14, HHF1, HTB2, HTZ1, NET1, SNF5, YOX1
YTA7 (1)	ARP9, CDC48, CHD1, CKA2, CKB1, COP1, FOB1, FPR3, FPR4, FTI1, HHF1, HHO1, HHT1, HIR2, HOS3, HTA2, HTB2, HTZ1, IDH1, IMD2, INO80, IOC2, IOC4, ISW1, ISW2, ITC1, KAP123, MCM2, MCM3, MCM4, MCM5, MCM7, MIF2, MOT1, NET1, NGG1, NPL6, NTO1, ORC1, ORC4, POB3, PSA1, RAD16, RRN9, RSC2, RSC3, RSC4, RSC58, RSC6, RSC9, RVB1, RVB2, SAS3, SEC21, SEC27, SFH1, SNF7, SNZ2, SPT16, SPT5, STH1, SUM1, TAF60, TAF61, TAF90, TIF34, TOF2, YAP1, YDL156W, YHR020W, YHR054C, YLR157C-BP, YLR455W, YTA7
YTA7 (2)	APL6, APM3, ARP9, BDF1, BDF2, CDC14, CHD1, CKA1, CKA2, CKB1, CKB2, ESF1, FPR3, FPR4, GRX4, HHF1, HHO1, HHT1, HTA2, HTB2, HTZ1, IOC2, ISW1, MCM2, MCM4, MCM5, MCM6, NAB3, NET1, NPL6, POB3, RSC2, RSC3, RSC4, RSC6, RSC8, RSC9, RTT102, RTT106, RVB1, RVB2, SEC2, SFH1, SNZ3, SPT16, STH1, TAF14, TIF34, TIF35, TIF5, UBC8, YPL009C, YTA7

* Number in parentheses denotes biological replicate of the mChIP-MS.

Appendix 3-6: Protein associated with Hap2 using various AP-MS method and their cellular localization.

Prey	Interaction partner detected by			Prey Location***
	mChIP	TAP*	FLAG**	
AAC3			yes	mitochondria
ADA2	yes			nucleus
ARP4			yes	nucleus
ARP8	yes			nucleus
ARP9	yes			nucleus
ASF1	yes			nucleus
ATG17			yes	cytoplasm
ATG29			yes	cytoplasm
ATP3			yes	mitochondria
CDC14	yes			nucleolus
CDC33			yes	cytoplasm
CHD1	yes			nucleus
CIC1			yes	nucleus; nucleolus
CIS1			yes	cytoplasm
CYS4			yes	cytoplasm
DAL81	yes			nucleus
EAF1	yes			nucleus
EPL1	yes			nucleus
FOB1	yes			nucleus
FOL2			yes	cytoplasm; nucleus
GLN3	yes			cytoplasm; nucleus
GRH1			yes	cytoplasm
HAP3	yes	yes		cytoplasm; nucleus
HAP5	yes	yes	yes	cytoplasm; nucleus
HFI1	yes			nucleus
HHF1	yes			nucleus
HHO1	yes			nucleus
HHT1	yes			nucleus
HIR2	yes			nucleus
HRR25	yes			nucleus, bud neck, spindle pole body
HTA2	yes			nucleus
HTB2	yes			nucleus
HTZ1	yes			nucleus
IKI3	yes			cytoplasm
INO80	yes			nucleus
IOC2	yes			nucleus
IOC3	yes			nucleus
IPP1			yes	cytoplasm; nucleus

ISW1	yes		nucleus
ISW2	yes		nucleus
LOC1		yes	nucleus
MAM33		yes	mitochondria
MCM2	yes		nucleus
MCM3	yes		nucleus
MCM4	yes		nucleus
MCM5	yes		nucleus
MCM6	yes		nucleus
MCM7	yes		nucleus
MOT1	yes		nucleus
MTQ1		yes	mitochondria
NAP1		yes	cytoplasm
NET1	yes		nucleolus
NGG1	yes		nucleus
NMD5		yes	cytoplasm; nucleus
NOP16		yes	nucleus; nucleolus
NPL6	yes		nucleus
NSA2		yes	cytoplasm
POL5		yes	nucleus; nucleolus
PSE1		yes	cytoplasm; nucleus
RFM1	yes		cytoplasm; nucleus
RHR2		yes	ambiguous
RSC1	yes		nucleus
RSC2	yes		nucleus
RSC3	yes		nucleus
RSC30	yes		nucleus
RSC4	yes		nucleus
RSC58	yes		nucleus
RSC6	yes		nucleus
RSC8	yes		nucleus
RSC9	yes		nucleus
RTT102	yes		nucleus
RVB1	yes		nucleus
RVB2	yes		nucleus
SAH1		yes	cytoplasm
SAP190		yes	unknown
SGF29	yes		nucleus
SGF73	yes		nucleus
SNF12	yes		nucleus
SNF2	yes		nucleus
SNF5	yes		nucleus
SPE3		yes	cytoplasm; nucleus
SPT15	yes		nucleus
SPT20	yes		nucleus

SPT3	yes		nucleus
SPT5	yes		nucleus
SPT6	yes		nucleus
SPT7	yes		nucleus
SPT8	yes		nucleus
SSK2		yes	cytoplasm
STE12	yes		nucleus
STH1	yes		nucleus
STP1	yes		nucleus
STP2	yes		cytoplasm; nucleus
SWI3	yes		nucleus
SWP82	yes		nucleus
TAF1	yes		nucleus
TAF10	yes		nucleus
TAF12	yes		nucleus
TAF14	yes		nucleus
TAF4	yes		nucleus
TAF5	yes		nucleus
TAF6	yes		nucleus
TAF8	yes		nucleus
TAF9	yes		nucleus
TBF1	yes		nucleus
TIF2		yes	cytoplasm
TIF6		yes	nucleus; nucleolus
TRA1	yes		nucleus
YAP5	yes		cytoplasm; nucleus
YPR085C		yes	unknown
YRA1		yes	nucleus
YRA2		yes	cytoplasm; nucleus
YTA7	yes		nucleus
YTM1		yes	nucleus; nucleolus

* Information obtained from the *BioGRID* (www.thebiogrid.org) from TAP-MS genome-wide studies {Gavin, 2006 #119;Krogan, 2006 #109}.

** FLAG tagged baits were overexpressed using a galactose inducible promoter {Ho, 2002 #117}.

*** Prey localization was obtained from the *Yeast GFP fusion localization database* (yeastgfp.yeastgenome.org) {Huh, 2003 #321}.

Appendix 3-7: Protein associated with Asf1 in multiple strain background by mChIP-MS

	Asf1-TAP mChIP (number of peptide identified following mChIP-MS)									
Prey Name	WT-1	WT-2	WT-3	*rtt106Δ*	*hir1Δ*	*rtt109Δ*	*vps75Δ*	*hat2Δ*	*hht1Δhht2Δ* +H3	*hht1Δhht2Δ* +H3K56R
ABF2	5	7	6	6	0	11	6	8	15	9
ANP1	0	6	0	0	0	0	3	8	2	0
ARB1	0	3	0	0	0	0	0	0	0	0
ARP4	4	14	16	12	0	16	15	17	18	16
ARP5	0	11	6	0	0	4	7	11	8	6
ARP7	0	4	0	0	0	3	4	6	4	0
ARP8	8	19	30	16	0	21	17	26	32	55
ARP9	0	0	9	0	0	2	4	11	13	0
ASC1	0	20	5	4	4	3	5	11	7	4
ASF1	54	213	127	162	196	69	60	72	119	127
ASG1	0	3	0	0	0	0	0	0	0	0
ATP2	0	0	9	0	0	0	0	0	0	0
BDF1	0	0	0	0	0	0	3	0	4	0
BDP1	0	0	10	0	0	0	4	0	0	0
BFR2	0	0	0	0	0	0	0	0	5	0
BRF1	0	0	5	0	0	0	0	0	3	0
BRN1	0	0	0	0	0	0	0	0	2	0
BUR6	0	0	0	0	0	0	0	0	4	0
CBF2	0	0	0	0	0	3	2	0	0	0
CDC14	25	40	136	56	3	32	40	41	97	44
CDC31	0	7	3	3	7	0	0	0	6	0
CHA1	0	0	3	0	0	0	0	0	0	0
CHD1	0	0	3	3	0	8	6	25	3	0
CKA1	3	48	76	30	108	26	23	26	40	32
CKA2	10	88	61	48	96	39	18	25	38	28
CKB1	5	44	33	29	43	22	16	11	19	26
CKB2	6	43	45	31	72	21	16	18	24	28
CSM1	0	4	0	0	0	0	0	0	0	0
DED1	0	5	3	0	0	0	0	0	3	0
DID2	0	6	4	0	0	4	4	9	5	4
DID4	0	0	0	0	0	0	0	5	0	0
DLS1	0	0	5	0	0	0	0	0	0	0
DPB4	5	11	24	8	0	17	18	15	20	16
EBP2	0	0	0	0	0	0	0	0	2	0
FKS1	0	0	0	0	8	0	0	0	0	0
FOB1	3	12	30	12	0	3	8	8	14	8
FPR3	0	0	0	0	0	0	0	0	7	0
GPM1	0	0	0	0	0	0	0	0	3	0
HAS1	0	0	0	0	0	0	0	0	5	0

HAT1	0	0	0	0	4	0	0	0	5	0
HAT2	0	0	0	0	0	0	0	0	4	0
HFD1	0	6	0	0	0	0	0	0	0	4
HHF1	60	109	99	129	117	54	51	56	104	72
HHO1	5	6	15	13	0	4	8	8	12	0
HHT1	0	9	11	9	9	0	3	0	10	12
HIR1	52	263	253	173	0	66	91	116	217	268
HIR2	71	373	385	300	0	183	207	259	373	536
HIR3	55	233	277	168	0	128	150	192	279	148
HMO1	7	16	40	17	0	10	12	11	50	12
HOS3	0	0	2	0	0	0	0	0	0	0
HPC2	31	105	212	111	0	62	94	99	210	326
HSC82	0	0	0	0	0	0	0	0	4	0
HSP26	0	0	0	0	0	0	0	0	3	0
HTA2	28	16	21	15	0	21	20	19	17	17
HTB2	10	28	38	30	0	17	19	17	27	21
HTZ1	19	5	16	8	0	10	5	12	12	12
IES1	3	13	28	11	0	9	11	15	25	17
IES2	0	0	0	0	0	0	0	0	4	0
IES3	2	3	3	0	0	0	3	5	6	0
IES4	0	0	4	0	0	0	0	4	0	4
IES5	0	0	0	0	0	0	0	3	0	3
IMD3	0	0	0	0	0	4	10	0	3	8
IMD4	0	9	0	0	0	0	0	5	0	0
INO80	9	48	34	26	0	26	16	47	46	14
IOC2	0	0	4	0	0	0	0	0	3	0
IOC3	0	0	0	0	0	0	0	3	4	0
IOC4	0	0	0	0	0	0	2	0	2	0
IST1	0	0	0	0	0	0	0	4	0	0
ISW1	0	29	37	24	0	17	16	32	35	19
ISW2	25	79	131	73	0	61	53	70	102	46
ITC1	16	67	70	52	0	60	33	90	96	44
KAP95	2	20	44	31	21	18	13	16	33	48
KAR2	0	9	14	13	0	9	0	7	22	18
KEM1	0	0	4	0	0	0	0	0	0	0
KOG1	0	0	3	0	0	0	0	0	0	0
KRE33	0	0	5	0	0	0	0	0	7	0
LSM12	0	0	0	0	4	6	5	6	2	0
MAM33	0	11	12	21	35	0	4	0	0	0
MCM2	0	0	3	0	0	0	0	0	0	0
MCM3	0	0	0	0	0	0	0	3	6	0
MCM4	0	0	4	0	0	0	0	0	3	0
MCM6	2	0	0	0	0	0	0	0	0	0
MCM7	0	0	3	0	0	0	0	0	3	0

MGM101	3	11	16	6	0	12	15	9	15	8
MIP1	0	0	4	0	0	0	0	0	0	0
MIR1	0	0	0	0	0	0	0	0	3	0
MNN9	0	0	0	0	0	0	0	5	0	0
MNR2	4	108	64	54	126	54	34	39	53	221
MOT1	0	12	11	0	0	7	0	10	16	0
MPH1	0	9	0	0	0	0	0	0	3	0
MSH2	0	4	4	0	0	6	7	10	13	0
MSH3	0	0	0	0	0	0	0	0	5	0
MSH6	0	0	9	0	0	0	4	0	4	0
MSS116	0	7	19	0	0	0	0	0	7	0
NAN1	0	0	0	0	0	0	0	0	4	0
NET1	40	107	307	162	0	61	62	94	221	104
NHP10	0	7	5	0	0	3	7	7	17	11
NHP2	2	0	0	0	0	0	0	0	5	0
NHP6A	0	0	4	0	0	0	0	3	3	0
NPL6	3	10	12	3	0	7	16	9	10	3
NUD1	0	0	0	0	0	0	0	0	6	0
ORC1	0	0	3	0	0	5	7	7	7	0
ORC2	0	0	0	0	0	0	0	4	0	0
ORC3	0	0	11	0	0	0	0	3	8	5
ORC4	0	0	7	0	0	7	9	0	7	0
ORC5	0	0	0	0	0	2	0	4	4	0
PBP1	0	0	0	0	4	0	0	0	0	0
PBP4	0	0	0	0	0	0	2	3	0	0
PDC2	0	3	5	0	0	0	3	0	5	0
PDR1	3	9	17	9	0	15	16	16	24	11
PFK1	0	0	0	0	0	0	0	0	2	0
PHO2	10	20	60	26	0	7	12	14	35	72
PIL1	0	0	3	0	0	3	0	0	5	0
POL30	0	0	0	0	0	0	0	0	3	0
PRP43	0	0	8	7	0	8	5	0	23	3
PWP2	0	0	0	0	0	0	0	0	6	0
PXR1	0	0	0	0	0	0	0	0	2	0
RAD52	0	0	3	0	0	0	0	0	4	0
RAD53	4	85	88	47	94	25	34	45	57	51
REB1	0	0	0	0	0	0	0	0	3	0
RET1	3	6	19	6	0	0	0	8	22	11
RFA1	3	27	28	15	0	16	21	14	36	31
RFA2	0	6	5	0	0	0	0	6	11	6
RFA3	0	8	4	4	0	4	3	6	9	8
RFC1	0	0	7	2	0	0	6	7	12	0
RFC2	0	5	7	0	0	0	0	0	8	6
RFC3	3	3	9	0	0	0	0	2	5	10

RFC4	0	0	0	0	0	0	0	0	13	3
RFC5	0	0	6	0	0	0	0	0	9	9
RFM1	0	4	0	0	0	0	0	5	7	5
RIM1	5	12	21	13	0	13	12	10	11	18
RLF2	0	0	6	3	5	0	0	0	3	0
RNQ1	0	9	8	16	12	0	0	3	9	8
RPA12	2	0	7	0	0	0	0	0	7	0
RPA135	12	56	93	71	0	55	51	59	79	45
RPA14	0	6	8	4	0	5	4	6	9	6
RPA190	38	87	123	122	0	56	59	111	131	87
RPA34	0	0	3	7	0	0	0	0	10	5
RPA43	9	8	18	9	0	4	8	11	19	13
RPA49	9	15	46	18	0	12	18	18	50	16
RPB2	0	0	0	0	0	4	0	0	4	0
RPB5	0	14	20	10	0	5	8	9	31	13
RPB8	0	4	4	7	0	8	0	4	6	0
RPC10	3	0	0	0	0	0	0	0	0	0
RPC17	0	0	4	0	0	0	0	7	4	4
RPC19	3	5	4	4	0	4	0	0	6	4
RPC31	0	0	0	0	0	0	0	2	7	3
RPC34	0	0	3	0	0	0	0	5	6	0
RPC37	2	3	0	0	0	0	0	0	10	6
RPC40	9	19	30	9	0	13	14	15	32	19
RPC53	0	0	0	0	0	3	0	6	9	4
RPC82	0	8	17	13	0	8	13	18	18	44
RPL24A	3	0	0	0	0	0	0	0	6	0
RPO21	0	4	0	0	0	0	0	0	7	0
RPO26	0	0	0	0	0	0	5	0	6	0
RPO31	5	16	24	14	0	19	7	11	28	19
RPO41	0	2	0	0	0	0	0	0	0	0
RPS14A	2	8	7	6	7	7	7	4	7	9
RRN10	0	0	4	3	0	0	0	0	0	2
RRN11	0	0	2	0	0	0	0	0	0	0
RRN3	0	0	2	0	0	0	0	0	0	0
RRN5	0	0	4	0	0	0	0	0	0	0
RRN6	0	0	0	0	0	0	0	0	6	0
RRN7	0	0	0	0	0	0	0	0	3	0
RRN9	0	0	13	0	0	0	0	0	0	0
RRP9	0	0	0	0	0	0	0	0	5	0
RSC1	0	0	0	0	0	0	0	0	2	0
RSC2	3	0	11	0	0	5	5	4	11	0
RSC3	2	0	12	0	0	3	7	13	4	2
RSC30	0	0	8	3	0	0	0	0	6	3
RSC4	2	0	14	0	0	0	5	10	8	0

RSC58	0	5	9	3	0	8	9	8	12	3
RSC6	0	8	13	6	0	2	11	4	13	5
RSC8	5	20	27	14	0	11	20	20	22	26
RSC9	0	3	17	0	0	4	7	10	12	4
RTT102	0	2	0	0	0	0	0	0	4	0
RTT106	19	26	73	0	0	7	31	33	41	7
RVB1	9	32	34	18	0	34	28	35	52	25
RVB2	9	33	33	23	0	40	37	38	39	46
SAS2	0	6	7	12	23	0	0	0	0	0
SAS4	0	9	15	31	50	0	0	0	0	0
SAS5	0	0	5	5	11	0	0	0	0	0
SEN1	0	0	0	0	0	0	0	0	5	0
SFH1	0	0	0	0	0	0	0	0	3	0
SGF29	0	0	0	0	0	0	0	0	4	0
SIK1	0	0	7	4	0	0	0	0	23	7
SIR2	0	0	9	0	0	0	0	0	9	0
SIS1	0	0	8	5	0	0	0	0	3	0
SLD3	0	10	54	24	47	0	0	7	18	17
SNF12	0	9	4	0	0	0	3	0	0	0
SNF2	0	0	0	0	0	3	0	8	0	0
SNF5	0	2	2	0	0	4	0	0	0	0
SNF6	0	0	0	0	0	0	0	3	0	0
SNF7	3	6	5	0	0	3	4	6	4	4
SNU13	0	0	0	0	0	0	0	0	0	4
SPC110	0	3	0	0	4	0	0	0	8	0
SPC29	0	3	0	0	0	0	0	3	3	4
SPC42	0	0	0	0	0	0	0	0	6	0
SPT15	0	6	0	7	0	6	6	6	13	5
SPT16	0	0	0	0	0	0	0	2	0	0
SPT5	0	0	3	0	0	0	0	0	0	0
SRP1	5	0	22	13	8	9	4	7	16	18
SRP40	0	8	4	0	0	0	0	0	7	8
STH1	3	19	8	5	0	9	10	9	18	0
SUB2	0	0	0	0	0	0	0	0	2	0
SUM1	0	16	12	7	0	15	3	20	17	0
SUP35	0	0	0	0	0	0	0	0	7	0
SWC4	0	0	0	0	0	0	0	0	4	0
SWI1	0	0	0	0	0	4	0	0	2	0
SWI3	0	7	0	0	0	11	10	19	12	0
TAF10	0	0	2	0	0	5	0	3	8	0
TAF11	0	0	0	0	0	0	0	0	2	0
TAF12	0	9	0	2	0	0	11	5	13	6
TAF14	0	9	5	4	0	7	5	7	8	6
TAF4	0	0	0	0	0	0	0	0	5	0

TAF5	0	6	6	3	0	3	7	7	8	9
TAF6	0	3	0	0	0	12	6	8	14	0
TAF8	0	0	0	0	0	0	0	0	4	0
TBF1	0	2	4	0	0	0	4	0	5	0
TDH2	0	0	0	0	0	0	0	6	0	0
TFB1	0	0	0	0	0	0	0	0	4	0
TFB2	0	5	0	0	0	0	0	3	4	0
TFC4	0	3	5	0	0	6	7	6	3	0
TFC6	0	0	7	0	0	0	0	0	0	0
TFC7	0	0	4	0	0	0	0	3	0	0
TIF1	0	3	0	0	0	0	0	0	3	0
TOA2	0	0	4	0	0	0	0	0	0	3
TOF2	3	12	16	12	0	0	5	7	20	0
TOP1	13	41	42	38	0	21	29	27	62	141
TOP2	7	27	54	67	0	50	34	37	83	65
TOR1	0	0	11	0	0	0	0	0	0	0
TRA1	0	0	4	0	0	0	0	4	7	0
TRI1	0	7	5	3	0	0	0	6	11	0
TUB1	0	0	0	0	0	0	0	0	2	0
TUB2	0	0	0	0	0	0	0	0	8	0
UAF30	0	0	6	0	0	0	0	0	4	0
UME6	0	0	0	0	0	0	0	0	3	0
VID22	0	0	0	0	0	0	4	6	5	0
VPS1	5	59	92	35	0	38	29	46	52	420
VPS24	0	0	0	0	0	0	0	3	0	0
YCG1	0	0	0	0	0	0	0	0	3	0
YDL156W	0	0	0	0	0	0	3	0	3	0
YDR210C-D	0	0	0	0	0	0	0	0	0	7
YEL007W	0	0	4	0	0	0	0	0	0	0
YGR027W-B	0	0	0	0	0	4	0	0	0	0
YGR071C	0	0	2	0	0	0	0	0	0	0
YJR029W	0	0	0	27	0	0	0	0	0	0
YML045W	0	10	0	0	0	0	7	0	0	0
YNL054W-B	0	0	0	0	0	0	0	0	21	0
YOL086W-A	3	0	0	0	0	0	0	0	0	0
YOR060C	2	13	12	8	15	0	0	7	12	14
YPR137C-B	0	0	20	0	0	0	0	9	0	0
YPR204W	0	0	7	0	0	0	0	0	0	0
YRA1	0	8	8	0	0	4	5	10	6	8
YTA7	0	0	4	0	0	0	0	7	4	0