UNIVERSITY OF CALIFORNIA, SAN DIEGO

A proteomics approach to study the DNA damage checkpoint in *Saccharomyces cerevisiae*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Chemistry and Biochemistry

by

Claudio Ponte de Albuquerque

Committee in charge:

Professor Huilin Zhou, Chair
Professor Xiang-Dong Fu
Professor Gourisankar Ghosh
Professor Partho Ghosh
Professor David Hendrickson

2010

The dissertation of Claudio Ponte de Albuquerque is approved, and it is acceptable in quality and form for publication of microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2010

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would first like to acknowledge my advisor, Huilin Zhou, for all the support I have received thought out the years, and for providing a good environment where I could learn. I will not forget all of our daily morning coffee breaks, and our discussion about science and life.

Next, I would like to acknowledge the current and past members of the Zhou lab, especially Marcus Smolka and Sheng-hong Chen for all the advice and help I received. I will never forget all the good times we had inside and outside the lab, including having to play Barbie with Julia, and trips to the Pub.

Finally, I would like to acknowledge all my friends and family for making my life pleasant and enjoyable outside of science, especially my parents Claudia & Fred and host parents Jim Hribar & Debora Sollohub.

Chapters 3, are reprints of the material as it appears in Mol Cell Proteomics, 2008, Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H. The dissertation author was the primary researcher and one of the primary author of this paper.

Chapter 4, is accepted and it will appears in J Biol Chem, 2010,Sheng-hong Chen, Claudio P. Albuquerque, Jason Liang, Raymond T. Suhandynata, and Huilin Zhou. The dissertation author was one of the primary researcher and author of this paper.

Vitae

2003 BS, Chemistry and Biochemistry, University of California, San Diego

2010 PhD, Chemistry and Biochemistry, University of California, San Diego

Publications

Chen SH*, **Albuquerque CP***, Liang J, Suhandynata RT, Zhou H. *Identification and Characterization of Substrates of DNA Damage Checkpoint Kinases in Yeast.* J Biol Chem
(* Equal contribution)

Albuquerque CP*, Smolka MB*, Payne SH, Bafna V, Eng J, Zhou H. *A multidimensional chromatography technology for in-depth phosphoproteome analysis*. Mol Cell Proteomics. 2008 Jul;7(7):1389-96.
(* Equal contribution)

Smolka MB, **Albuquerque CP**, Chen SH, Zhou H. *Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases*. Proc Natl Acad Sci U S A. 2007 Jun 19;104(25):10364-9.

Smolka MB, Chen SH, Maddox PS, Enserink JM, **Albuquerque CP**, Wei XX, Desai A, Kolodner RD, Zhou H. *An FHA domain-mediated protein interaction network of Rad53 reveals its role in polarized cell growth*. Journal Cell Biology. 2006 Dec 4;175(5):743-53.

Kats ES, **Albuquerque CP**, Zhou H, Kolodner RD. Checkpoint functions are required for normal S-phase progression in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2006 Mar 7;103(10):3710-5.

Smolka MB, **Albuquerque CP**, Chen SH, Schmidt KH, Wei XX, Kolodner RD, Zhou H. *Dynamic changes in protein-protein interaction and protein phosphorylation probed with amine-reactive isotope tag.* Mol Cell Proteomics. 2005 Sep;4(9):1358-69.

ABSTRACT OF THE DISSERTATION

A proteomics approach to study the DNA damage checkpoint in *Saccharomyces cerevisiae*

by

Claudio Ponte de Albuquerque

Doctor of Philosophy in Chemistry and Biochemistry

University of California, San Diego, 2010

Professor Huilin Zhou, Chair

The DNA damage checkpoint is a signal transduction pathway that is evolutionarily conserved from yeast to human. It is involved in the maintenance of the genome and regulation of cell cycle. Mutations in checkpoint proteins can lead to genomic instability and cancer in humans.

This signal transduction pathway is composed by several kinases and adaptor proteins, and only a few substrates of these signaling transduction pathway are known. In *Saccharomyces cerevisiae* (budding yeast) Mec1 and Tel1 are the two upstream kinases, homologues to ATR and ATM in humans, involved in the initial sensing of the damage DNA or the stalled replication fork. Rad9 and Mrc1, orthologs of human 53BP1 and Claspin, are the main adaptor proteins of Mec1 and Tel1, and they activate the effector kinase Rad53 and Dun1.

Recent technological advances in the field of proteomics have enabled us to identify and quantify thousands of peptides in a single experiment. In these studies I developed two separate multidimensional chromatography methods which when combined with quantitative proteomics can be used to examine the global changes in protein expression or phosphorylation. In Chapter 2, my first mutidimensional chromatography method was used to quantify changes in protein expression after DNA damage. In Chapter 3, a second mutidimensional chromatography method was developed, and it allowed an in-depth mapping of the phosphoproteome. Finally, we applied these methods to identify *in vivo* kinase substrates of the DNA damage checkpoint.

Chapter 1: Biological application of application of tandem mass spectrometry to biology

In the past few years tandem mass spectrometry (MS) has become a preferred method to identify and quantify changes in protein complexes and post translation modifications from purified samples. Don Hunt was the first to show that MS was able to sequence and identify a small number of peptides from a complex peptide mixture(Hunt et al., 1992). Recent advantages in MS instrumentation have increased the scanning speed, sensitivity and resolution enabling the identification of thousands of peptides in a single liquid chromatography and tandem mass spectrometry (LC-MS/MS) run. With this observation, several groups have postulated that global proteomic analysis can be achieved. By combining the modern quantitative techniques and multidimensional chromatography it is possible to identify changes in protein expression and *in vivo* kinases substrates from whole cell-lysate.

Quantitative Proteomics is used to quantify relative changes between samples, and the most commonly used method is isotope labeling which comes in two types: chemical and metabolic. In chemical labeling, an isotopic tag is added to peptides by a chemical reaction, pioneered by the use of the ICAT reagent (Gygi et al., 1999). In this approach, cysteine-containing peptides are isotopically labeled, purified and analyzed by MS. The ICAT approach has the advantage that only cysteine-containing peptides are

purified and analyzed by MS, which may be a useful simplification strategy considering the complexity of the proteome. However, a chemical labeling approach requires sufficient experience and optimization for proper use (Smolka et al., 2001). The second approach is the use of metabolic labeling via either N15 containing growth media (Oda et al., 1999), or media containing stable isotope containing amino acids (SILAC) (Ong et al., 2002). The SILAC approach, in particular, causes minimal perturbations to cell growth. Several studies have demonstrated its accuracy and applicability for many cell culture models including budding yeast (de Godoy et al., 2006; Graumann et , 2008; Kruger et al., 2008; Zhai et al., 2008).

Multidimensional Chromatography is used to reduce the complexity of a sample by fractionation based on physical characteristics such as charge, hydrophobic or hydrophilic interaction. It is necessary to reduce the complexity of the sample prior to its analysis by MS in order to allow time for the peptide to be sequenced and subsequently identified. Reverse-phase liquid chromatography (RP-LC) is the most commonly used method to fractionate peptides, and is based on hydrophobic interactions. This is usually the last fractionation step before delivering peptides into the MS. The second most commonly used fractionation method, strong cation exchange chromatography (SCX), relies on a charge based fractionation step. Hydrophilic columns are not commonly used to fractionate peptides because highly organic solvents are required and peptides are only slightly soluble in these solvents.

Chapter 2: A multidimensional chromatography technology for protein expression analysis.

2.1 Abstract

A major objective of proteomics is to characterize protein expression levels in cells and to study how they are perturbed in response to various environmental stimuli. Using the budding yeast *Saccharomyces cerevisiae*, a multi-dimensional chromatography technology was developed to study how the DNA alkylating agent, methyl methane sulfonate (MMS), could induce changes in protein abundance in cells. This technology entails the use of strong cation exchange (SCX) chromatography, hydrophilic interaction chromatography (HILIC) and reverse phase liquid chromatography (RP-LC) in a sequential order to fractionate proteolyzed cell extracts prior to their analysis by tandem mass spectrometry (MS/MS). To enable quantitative analysis, cells either untreated or treated with MMS were metabolically labeled using stable isotope labeling via amino acids in culture (SILAC). Among the 2302 proteins identified and quantified with at least three unique peptides, nine proteins were found to undergo MMS induced changes in abundance, which were confirmed using independent methods. This study thus revealed the potential of this technology for protein expression analysis in yeast and further identified areas for future improvement.

2.2 Introduction

Proteins perform the majority of the biological activities in cells. The expression levels of proteins in cells and how they change in response to various stimuli provide the basic information towards understanding their functions. The regulation of protein abundance is complex and controlled by multiple pathways including gene transcription, protein synthesis and selective protein degradation. While global analysis of mRNA expression has been commonly used in many biological investigations, the analysis of protein expression in cells and its applications has been more limited to date. Recent advances in mass spectrometry (MS), particularly the improvements in its scanning speed and sensitivity, have enabled the identification of hundreds of proteins in a single liquid chromatography and tandem mass spectrometry (LC-MS/MS) analysis. Additional fractionation methods, either at protein or peptide level, were shown to enable the identification of a few thousand proteins (de Godoy et al., 2006; Gygi et al., 2000; Washburn et al., 2001; Wei et al., 2005). In addition to identifying more proteins in a proteomic study, it is important to obtain quantitative information on the changes in protein abundance and their modifications in cells.

Tens of thousands of peptides are produced after enzymatic cleavage of total cell lysate, even for a relatively simple organism like yeast. It is therefore necessary to use sufficient protein/peptide separation methods to fractionate the sample prior to its analysis by MS. There are several

considerations for the choice of separation methods for proteins and peptides. First, it should be robust and can be automated. Liquid chromatography is ideal for this purpose because it is reproducible and can be readily automated. Second, it is important to consider the capacity of a microcapillary RP-LC-MS/MS system. Typically, microcapillary LC column is used to deliver peptides into MS at a flow rate of sub-microliters per minute. The low flow rate effectively raises peptide concentration, thereby increasing the sensitivity of its detection by MS. However, there is a limited amount of reverse phase C18 material in a microcapillary column, thus imposing a physical limit on the amount of peptides to be analyzed without overloading. For optimal sensitivity, one microgram of peptides is typically analyzed using a microcapillary LC-MS/MS system, which can yield the identification of over 300 proteins from yeast cell lysate (unpublished observations). Third, given the loading capacity of standard microcapillary LC-MS/MS method, it is necessary to use additional fractionation method that differs from RP-LC in their mechanisms of separation in order to identify more proteins. The combination of RP-LC and SCX methods has allowed the identification of a few thousand proteins in a single proteomic experiment (Washburn et al., 2001). Fourth, to further increase the number of proteins identified, additional fractionation and more starting material are required. Importantly, any additional fractionation method should offer a separation mechanism that differs from both SCX and RP-LC. Finally, with the use of multiple columns, many column fractions will be generated for

LC-MS/MS analysis. In this case, the MS instrument time becomes a bottleneck issue. With a typical LC-MS/MS experiment taking about two hours, the combination of SCX and RP-LC-MS can take more than one day of MS instrument time. Thus, introduction of any additional fraction step should significantly improve the coverage of the proteome with minimal increase in the amount of instrument time.

A large-scale proteomic experiment typically yields a vast amount of MS data, which presents significant challenges for computational analysis. For protein identification, many software tools are already available, including SEQUEST, MASCOT and INSPECT(Perkins et al., 1999; Tanner et al., 2005; Yates et al., 1995). On the other hand, peptide quantification is generally more difficult than protein identification. This is because quantification of peptides requires their detection with sufficient signal intensity over a period of time in the MS mode (Han et al., 2001). Compared to the MS/MS mode, the MS mode generally contains a higher noise level, which makes it more difficult to quantify peptides than to identify them. In addition, for complex peptide samples, co-elution of distinct peptides with different amino acid sequences yet with similar mass to charge ratios are relatively common and could cause erroneous quantification (unpublished observations). In principal, the latter situation can be improved with the use of high-resolution mass spectrometers and sample fractionations that reduce the complexity of the sample.

Cells are highly responsive to environmental stimuli, particularly to damages to their genome. Indeed, several proteins, including ribonucleotide reductase (Elledge and Davis, 1989), Ddi1(Zhu and Xiao, 2001), Gpx2, Rfa2, and others (Lee et al., 2007), have been identified to undergo drastic changes in their expression levels in cells when treated with genotoxic agents. Importantly, these proteins are of relatively lower abundance in cells, therefore it was unclear whether they could be identified using a proteomic approach. The budding yeast *S. cerevisiae* is an ideal model system for such investigation, given the wealth of available information about yeast protein expression (Ghaemmaghami et al., 2003; Jelinsky and Samson, 1999; Lee et al., 2007). We thus sought to develop and evaluate a new multi-dimensional chromatography technology for protein expression analysis. The data presented here shows that both previously known and unknown DNA damage induced protein abundance changes were detected and confirmed using alternative methods.

2.3 Experimental procedures

2.3.1 Cell growth, treatment and sample preparation methods

SILAC media was prepared in the following manner: CSM-Arg-Lys-His was supplemented with 20 mg/L of histidine, 80 mg/L proline, 6.7 g/L yeast nitrogen base (Bacto Lab), and either 30 mg/L lysine and 20 mg/L arginine or 30mg/L (^{13}C6, ^{15}N7) lysine and 20 mg/L (^{13}C6, ^{15}N7) arginine (Sigma-Aldrich)

for “light” or “heavy” culture respectively. Two 50 mL cultures, containing either “heavy” or “light” amino acids, were inoculated with yeast strain SCY251 (MATa, *ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ*) for 6 generations. When cell density reached an OD of 0.3, 0.05% MMS was added to the yeast culture containing the heavy lysine and arginine for 3 hours and then the cells were harvested. The cell pellets were ground by glass beads at 4 degrees Celsius in a 2 mL microcentrifuge tube in 1.5 mL of lysate buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.2% NP-40, 5 mM EDTA, 5 mM NaF, 10 mM β-glycerophosphate, and 1 mM PMSF. Lysates were clarified by centrifugation at 13,200 RPM for 5 minutes. The protein concentration was measured using Bradford assay (Bio-Rad, Hercules, CA). The same amounts of light and heavy cell lysates prepared from untreated and MMS treated cells, respectively, were combined. The protein extract was precipitated by adding three volumes of 50% (v/v) ethanol/acetone with 0.1 % acidic acid. The protein pellet was air-dried and then re-suspended in 8 M urea and 100 mM Tris-HCl (pH 8.0), followed by reduction with 10 mM DTT for 30 mins at 50 °C. The reduced sample was then alkylated at room temperature with 50 mM iodoacetamide for 30 min. The sample was then diluted 5-fold by adding TBS buffer containing 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl and digested by trypsin overnight at 37 °C. The resulting peptides were then desalted using a 200 mg Sep-Pac C18 column and dried under reduced pressure.

2.3.2 Fractionation of peptides using SCX chromatography

A polysulfoethyl A column (The Nest Group, Inc**,** Southborough, MA, Cat # 051SE0502) was used to fractionate the peptides derived from the cell lysate above. Two buffers were used: Buffer A contains 5 mM KHPO4 (pH 2.7) and 25% ACN. Buffer B contains 5 mM KHPO4 (pH 2.7), 25% acetonitrile and 0.5 M KCl. For SCX fractionation, 0.5 mg of the salt-free peptides were re-suspended in 30 uL of Buffer A and injected via 100 uL loop with a flow rate of 200 uL/min. One-minute fractions were collected, dried, and desalted again using a 50 mg Sep-Pac C18 cartridge column. The SCX fractions 13-27 were then analyzed by LC-MS/MS. The loading from each SCX fraction varied from 10 % to 50% based on its UV absorption signal. The salt gradient used in the SCX fractionation is shown in Figure 2.2 (100 % Buffer A at time = 0 min, 100 % Buffer A at time = 5min, 5 % Buffer B at time = 10 min, 7 % Buffer B at time = 15 min, 15 % Buffer B at time = 18 min, 35 % Buffer B at time = 30 min, 100 % Buffer B at time = 35 min).

2.3.3 Separation of peptides in individual SCX fractions using HILIC.

The SCX fractions 14 to16 were used for further fractionation using a TSK gel Amide-80 column (2 mm x 15 cm, 5um; Tosoh, Montgomeryville, PA, CAT #19696**).** Three buffers were used: Buffer A containing 90 % ACN and 0.001 % TFA, Buffer B containing 0.005 % TFA in water, and Buffer C

containing 0.1 % phosphoric acid in water. The desalted peptides from SCX fractions 14, 15 and 16 were re-suspended in 30 uL buffer containing 90 % ACN and 1 % formic acid and then injected at a flow rate of 150 uL/min. One-minute fractions from HILIC were collected and analyzed by LC-MS/MS. The gradient used in HILIC is shown in Figure 3 (100 % Buffer A at time = 0 min, 100 % Buffer A at time = 5 min, 52 % Buffer B at time = 65 min, 95 % Buffer C at time= 70 min).

2.3.4 Mass Spectrometry and data analysis

The MS settings are similar to those previously described (Albuquerque et al., 2008) with the following minor modifications. For RP-LC gradient, a 140 min gradient from 15 % to 45 % Buffer B was used. On the MS setting, the minimal intensity threshold to trigger MS/MS scan was set to be 8000 counts and the dynamic exclusion time was lowered to 40 seconds. The charge state rejection option was enabled to reject singly charged peptides.

The tandem MS spectra were searched using a Sorcerer SEUQEST system (SageN, San Jose, CA) and searched against the budding yeast protein database (downloaded from SGD on January 12, 2007) with its reverse shuffle protein database appended. A 3.0 Da mass tolerance was allowed in the search. Two variable modifications of +8.0 and + 10.0 were added to lysine and arginine respectively. PeptideProphet probability was used to filter the results to less than 1 % false positive rate base on the decoy

database. XPRESS was used to quantify the ratio between the light and heavy peptides (Han et al., 2001). The XPRESS standard settings were changed to allow 1.5 Da mass tolerance, +8 Da or +10 Da for lysine and arginine respectively and +/-100 scans from peak apex as the fixed elution peak area. Peptides that did not contain lysine, arginine, or those that were assigned to more than one protein due to multiple isoforms of proteins were removed from further consideration, along with proteins identified with fewer than three unique peptides. Using this stringent filter, there were no proteins among the final list that were identified in the decoy database. Moreover, it is possible to determine the median of the abundance changes from at least three peptides obtained for each protein.

2.3.5 Western Blotting

TAP-tagged yeast strains were purchased from Open Biosystems, Huntsville, AL. Yeast cultures were first grown to saturation in CSM–His medium. 50 microliters of the saturated yeast culture was added to 75mL YPD overnight. At an OD of 0.5, the culture was split into 50 mLs and 25 mLs. 0.05% MMS was added to the 50 mL yeast culture for 3 hours and both yeast cultures were harvested after 3 hours of MMS treatment. The cell lysate was prepared by adding 800 uL of lysis buffer containing 50 mM Tris-HCl (pH8.0), 150 mM NaCl, 0.2% NP-40, 5 mM EDTA, 5 mM NaF, 10 mM β-glycerophosphate, 1 mM PMSF, and a protease inhibitor cocktail of 200 uM

benzamidine, 1 uM leupeptin, and 1.5 uM pepstatin A and grinding the cell pellet with glass beads in an ice-cold vortexer. Cell debris was removed by centrifugation at 13,200 rpm for 10 minutes in the cold room. The protein concentration was measured by Bradford assay. Twenty micrograms of the protein extracts were resolved on a 10% SDS-PAGE gel. The proteins were then transferred to a Hybond-P membrane (GE Healthcare Bio-Sciences Corp, Piscataway, NJ). The TAP-tagged protein was detected using an anti-TAP antibody (Open Biosystems, Huntsville, AL) followed by an HRP-conjugated goat anti-rabbit antibody (Millipore-Upstate, Billerica, MA) and then visualized by ECL (GE Healthcare Bio-Sciences Corp, Piscataway, NJ).

2.4 Results and Discussion

2.4.1 A multidimensional chromatography technology to identify and quantify changes in protein expression

Here we evaluated a three-dimensional chromatography technology for analysis of MMS induced protein expression changes in the budding yeast. As shown in Figure 2.1, cells were grown in either light Lys/Arg or heavy Lys/Arg containing media. To induce DNA damage, cells were either left untreated or treated by 0.05% MMS for 3 hours, a condition that is sufficient to activate the DNA damage checkpoint using Rad53 hyperphosphorylation as a marker (unpublished observation). Equal amounts of cell lysate from either untreated or MMS treated cells were combined, digested by trypsin, and then analyzed

using either LC-MS/MS alone, a combination of SCX and LC-MS/MS, or a combination of SCX, HILIC and LC-MS/MS (see Fig. 2.1). In each case, approximately 1 microgram of peptides was analyzed by LC-MS/MS based on the UV absorbance data. In this way, it is possible to evaluate the coverage of the proteome for each fractionation method used and identify an optimal strategy for large-scale proteomic analysis. It should be noted that HILIC was used immediately prior to the use of RP-LC-MS/MS. This is because the buffers used in HILIC are completely volatile, thus there is no need for sample clean-up steps after its usage. On the other hand, SCX was used as a first step to separate peptides derived from cell lysate. In this case, manual desalting steps are involved, which are preferred to be kept at a minimum.

2.4.2 Analysis of the budding yeast proteome using a combination of SCX and RP-LC-MS/MS

As shown in Figure 2.2A, peptides derived from cell lysate were first fractionated using SCX chromatography. They appear to elute in three major peaks. The first peak corresponds to the SCX-unbound fraction and it was discarded. To evaluate the separation of peptides, we chose to analyze SCX fractions 14 to 16 using LC-MS/MS. As shown in Figure 2.2B, overlap of peptides identified in the two non-adjacent SCX fractions 14 and 16 was less than 5%, although more overlaps were observed between adjacent fractions such as SCX fractions 14 and 15.

Next, to address how many proteins could be identified using this combination of SCX and RP-LC-MS/MS, a total of 15 SCX fractions were analyzed by LC-MS/MS. This led to the identification of 1,045 proteins and each protein was identified with at least three unique peptides. It should be noted that this protein identification criterion was chosen so that no false positives in protein identification are present in the final dataset, using the decoy database method (Elias and Gygi, 2007). To examine how the use of SCX method helps to increase the number of proteins identified, the cumulative number of the proteins identified from more SCX fractions appeared to reach a plateau by SCX fraction 22 (see Fig. 2.2C). Beyond that point, analysis of additional SCX fractions mostly increases the coverage of each protein with relatively minor gains in the identification of unique proteins. This is most likely due to the fact that each protein would generate peptides throughout the SCX gradient and their presence prevents the identification of lower abundance, and thus unique proteins. It should be noted that a comparable coverage of the yeast proteome was also obtained using the MudPIT method, which is essentially the same as the method used here (Washburn et al., 2001).

2.4.3 The use of HILIC as a third dimensional chromatography to complement SCX and RP LC-MS/MS

An additional fractionation method is necessary to further increase the coverage of the budding yeast proteome. To this end, HILIC was used to fractionate individual SCX fractions, including SCX14, SCX15, and SCX16 (see Fig. 2.3). It has been shown recently that HILIC provides an orthogonal separation compared to RP-LC (Gilar et al., 2005). Furthermore, the buffers used for HILIC can be made to be completely volatile and are thus compatible with RP-HPLC and MS/MS analysis (Albuquerque et al., 2008).

In Figures 2.3A-C, the UV absorbance of these SCX fractions separated by HILIC are shown. To evaluate the fractionation of peptides by HILIC, three consecutive HILIC fractions from SCX fraction 16 were chosen and analyzed by LC-MS/MS. As shown in Figure 3D, there is less than 15% overlap of peptides identified for two non-adjacent HILIC fractions. Next, we examined the proteins identified from all the HILIC fractions from a single SCX fraction. As shown in Figure 3E, there is a continuing increase of proteins identified as additional HILIC fractions of SCX fraction 16 was analyzed. This increase in protein identification could be attributed to several factors. First, for the same weight of peptides loaded to LC-MS/MS, there is a higher percentage of sample-loading after more fractionation steps. Thus, lower abundant proteins should be better identified. Second, for a single SCX fraction, it is less likely to have too many peptides from a single abundant

protein. Thus, additional HILIC separation of this SCX fraction simplifies the peptide mixture and enables better identification of peptides from unique proteins instead. It should be noted, however, that the total number of proteins identified from the analysis of a single SCX fraction using HILIC and RP-LC-MS/MS is similar to the use of SCX and RP-LC-MS/MS above (see Fig. 2.2), suggesting that the capacity of the LC-MS/MS is likely limiting.

To further increase the coverage of protein identification, the HILIC fractions from the two other SCX samples were also analyzed by LC-MS/MS. The result is summarized in Figure 2.3F. Analysis of a total of 51 HILIC fractions derived from three SCX fractions led to the identification and quantification of 2302 proteins, all of which were identified based on at least three unique peptides with no false positives in protein identification, as estimated using the standard target decoy approach (Elias and Gygi, 2007). The lack of false positive identifications can be directly attributed to the stringent requirement that three unique peptides must be found for each identified protein. It is likely that the actual number of proteins identified in this experiment is considerably higher. However, quantification of proteins is significantly more demanding and requires the quantification of multiple peptides in the present study. In summary, the above analysis showed that the inclusion of HILIC as a separation technique helped to increase the coverage of the proteome substantially.

2.4.4 MS data analysis of MMS induced protein abundance changes

As discussed above, identification of proteins is typically less demanding than quantification of proteins. This is primarily because accurate quantification requires sufficient signal/noise ratios of the peptides identified. Thus, we only consider a protein to be identified in the screening if at least three unique peptides identified and quantified. This not only helps to eliminate potential false positives in protein identification, it also provides the minimal number of peptides for each protein for a statistical analysis. With the availability of abundance ratios for three peptides of the same protein, their median was used to represent protein abundance changes in response to MMS treatment (see Fig. 2.4A).

The medians of the quantification results of all 2,302 proteins are shown in Figure 2.4A. Most of them do not show significant changes in protein abundance after MMS treatment, suggesting that MMS-induced changes in protein abundance may occur to a few low abundance proteins. Among them, 35 proteins show a median abundance change of over 2-fold and are further subjected to manual analysis. Several criteria are used during manual inspection. First, the signal/noise ratio must be over 5-fold for the more abundant isotopically labeled peptide. Second, a Gaussian-like elution profile must be detected for the peptide of interest. Third, all the peptides quantified for the same protein must show consistent quantification results. In other

words, they must all increase or decrease in abundance. After manual inspection, only nine candidate proteins satisfied all of the above criteria.

2.4.5 Confirmation of MMS induced protein abundance changes

We used independent methods to confirm the changes in protein abundance detected by MS. Anti-TAP western blot was used to probe any MMS-induced changes in protein abundance of these nine candidate proteins. As shown in Figure 4B, results from WB analysis of these nine proteins are consistent with their MS results. Rnr1, Gpx2, Rfa2, Ddi1, Kap120, and a protein encoded by YIL154C show abundance increases after MMS treatment. Rsf2, Qcr7, and Erg5 show an MMS-induced decrease in abundance. Previously, only Rnr1, Rfa2 and Ddi1 were known to undergo protein abundance increase after MMS treatment (Elledge and Davis, 1989; Zhu and Xiao, 2001). YIL154C is involved in maintenance of ion homeostasis and protection against DNA damage. Its deletion has been shown to be sensitive to MMS and other DNA damage reagents (Masson and Ramotar, 1996). Qcr7 is a subunit of the ubiquinol cytochrome-c reductase complex (Hemrika and Berden, 1990). Kap120 is a karyopherin involved in nuclear transport (Stage-Zimmermann et al., 2000). Gpx2 is a Phospholipid Hydroperoxide Glutathione Peroxidases involved in oxidative stress response(Inoue et al., 1999).

Next, the results of MMS-induced protein abundance changes, determined using MS and WB analysis, were compared to the results from

GFP tagged proteins and mRNA expression studies (Jelinsky and Samson, 1999; Lee et al., 2007). As shown in Figure 2.4B, these results appear to be consistent with each other, since the abundance increases of Gpx2 and Rfa2 were previously reported (Jelinsky and Samson, 1999; Lee et al., 2007). The only exception is Rsf2, a zinc-finger protein involved in transcriptional regulation (Bohm et al., 1997; de Lichtenberg et al., 2005). While the mRNA data of Rsf2 showed little change in abundance, our proteomic and WB results revealed a significant decrease in its abundance after MMS treatment.

2.5 Concluding remarks

In this study, we describe the development and evaluation of a three-dimensional liquid chromatography technology for analysis of protein expression in cells. Its application to the budding yeast led to the identification and quantification of 2302 proteins. Among them, nine proteins were found to undergo appreciable changes in protein abundance in response to MMS-induced DNA damages in cells and were confirmed using MS-independent method.

Despite its promises, there are technical challenges for global protein expression analysis. First, even after the use of three orthogonal chromatography methods, i.e., SCX, HILIC and RP-LC, the samples loaded to MS are still relatively complex. This likely limits the identification of lower abundant proteins and could also compromise their quantification. Second, a

low mass resolution MS instrument such as LTQ-MS, despite its popularity in protein identification, is limited for quantitative analysis of complex peptide mixtures. This has been noted in several studies (Venable et al., 2007). From our experience, the use of stringent criteria and manual examination of such MS data by experienced users is critical to ensure the quantification results are reliable. With the use of higher mass resolution MS, some of these difficulties would be alleviated and it would be possible to identify abundance changes to even lower abundant proteins in cells.

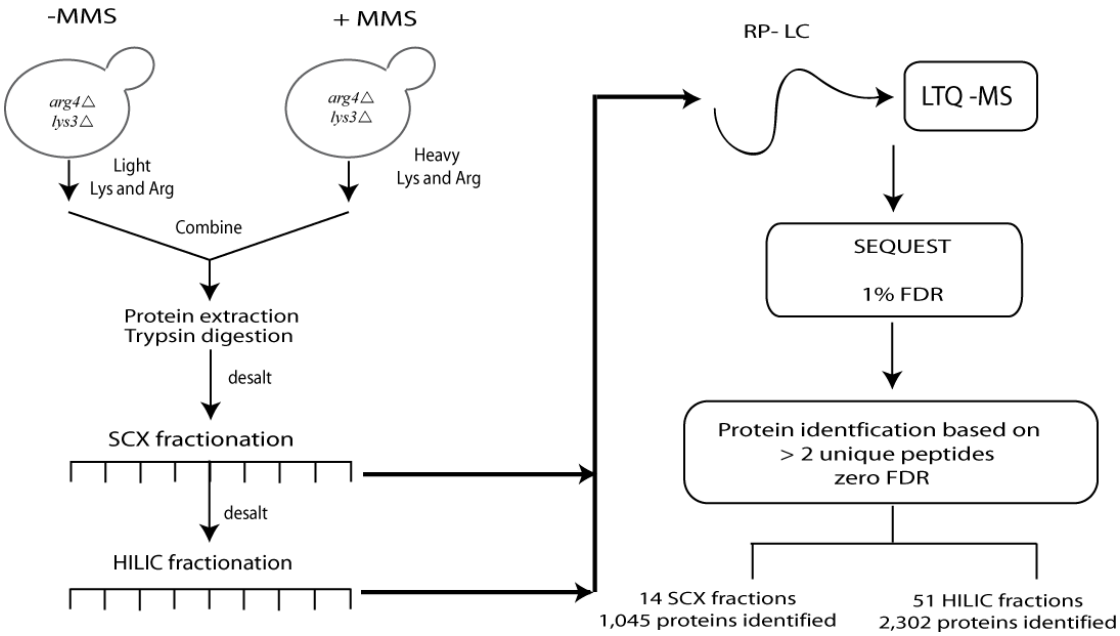

Figure 2.1 Schematics of a multi-dimensional chromatography and data analysis used to identify changes in protein abundance

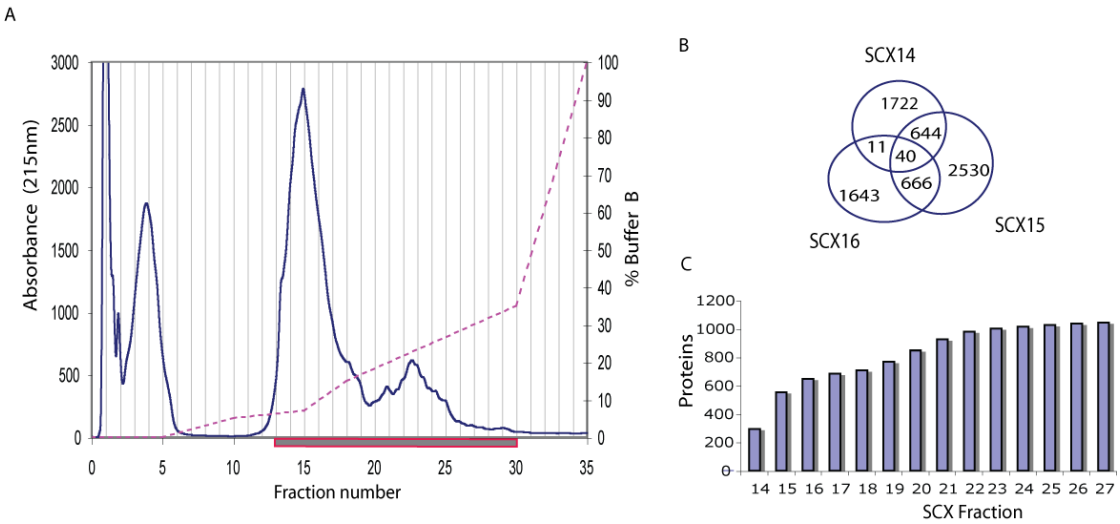

Figure 2.2 Figure 2.2 SCX fractionation of peptides proteolyzed from yeast cell lysate.
(A) The gradient (dashed line) and UV absorption (solid line) profiles of peptide fractionation by SCX (see experimental procedures for further details). One-minute fractions ranging from 13 to 27 (underlined) were subjected to LC-MS/MS analysis.
B) The Venn diagram of the peptide identified for three SCX fractions 14 to 16, showing the overlaps of peptides identified in these SCX fractions.
C) The cumulative number of proteins identified with more SCX fractions analyzed by LC-MS/MS. After SCX fraction 22, analysis of additional SCX fractions does not greatly increase the number of proteins identified.

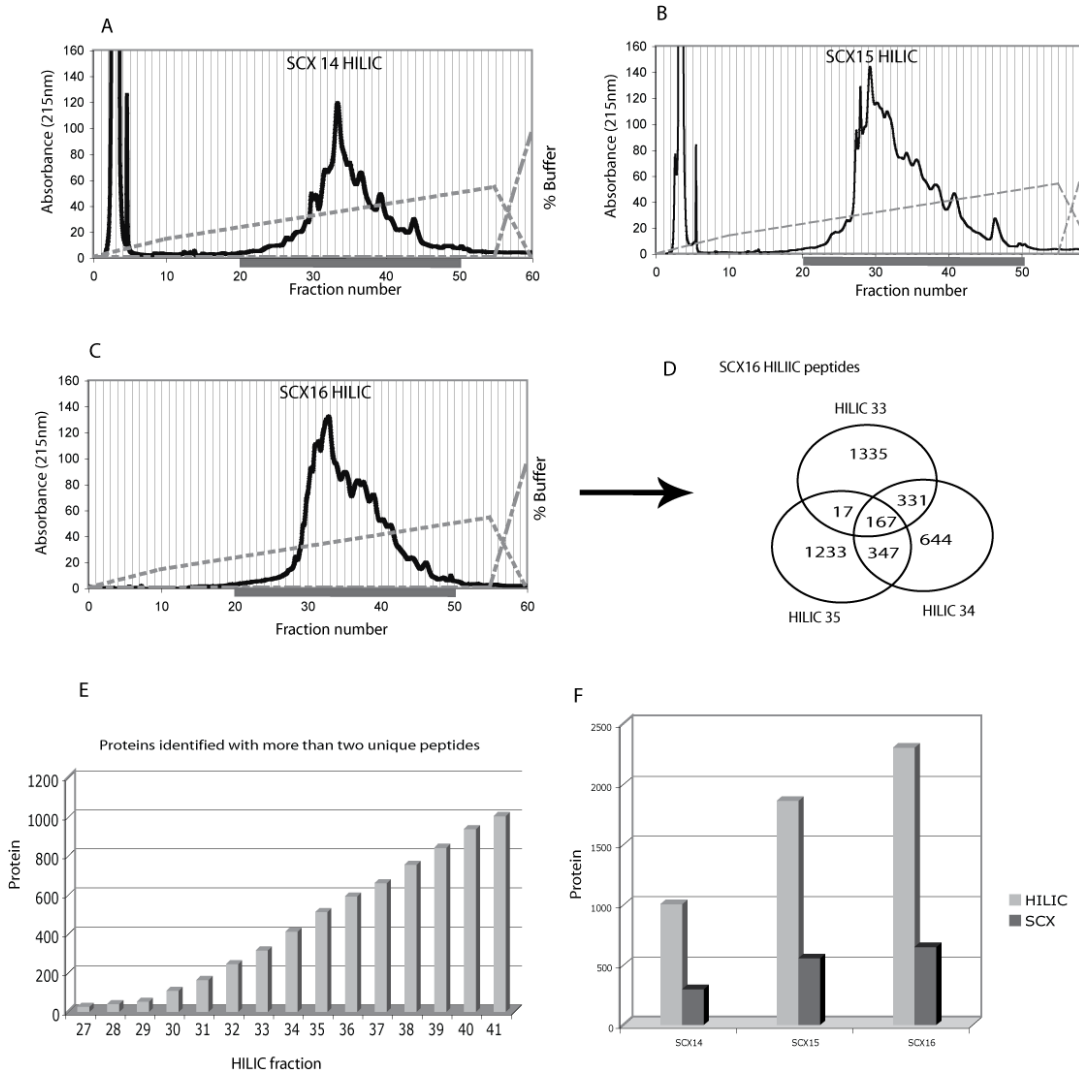

Figure 2.3 Inclusion of HILIC further increases the coverage of proteome.
-C) The gradient (dashed line) and UV absorption (solid line) profiles of peptides separated by HILIC (see experimental procedures for details). Three samples (fraction 14, 15 and 16) from SCX column were separated by HILIC. One-minute fractions ranging from 18 min to 50 min (underlined) were collected and analyzed by LC-MS/MS.
D) The Venn diagram of the overlaps of peptides identified in three adjacent HILIC fractions from the separation of SCX fraction 16.
E) The cumulative number of identified proteins by each additional HILIC fractionation of the SCX fraction 14.
F) Summary of the number of proteins identification before and after HIILC fractionation for SCX fractions 14, 15, and 16. Considerably more proteins were identified with the use of HILIC.

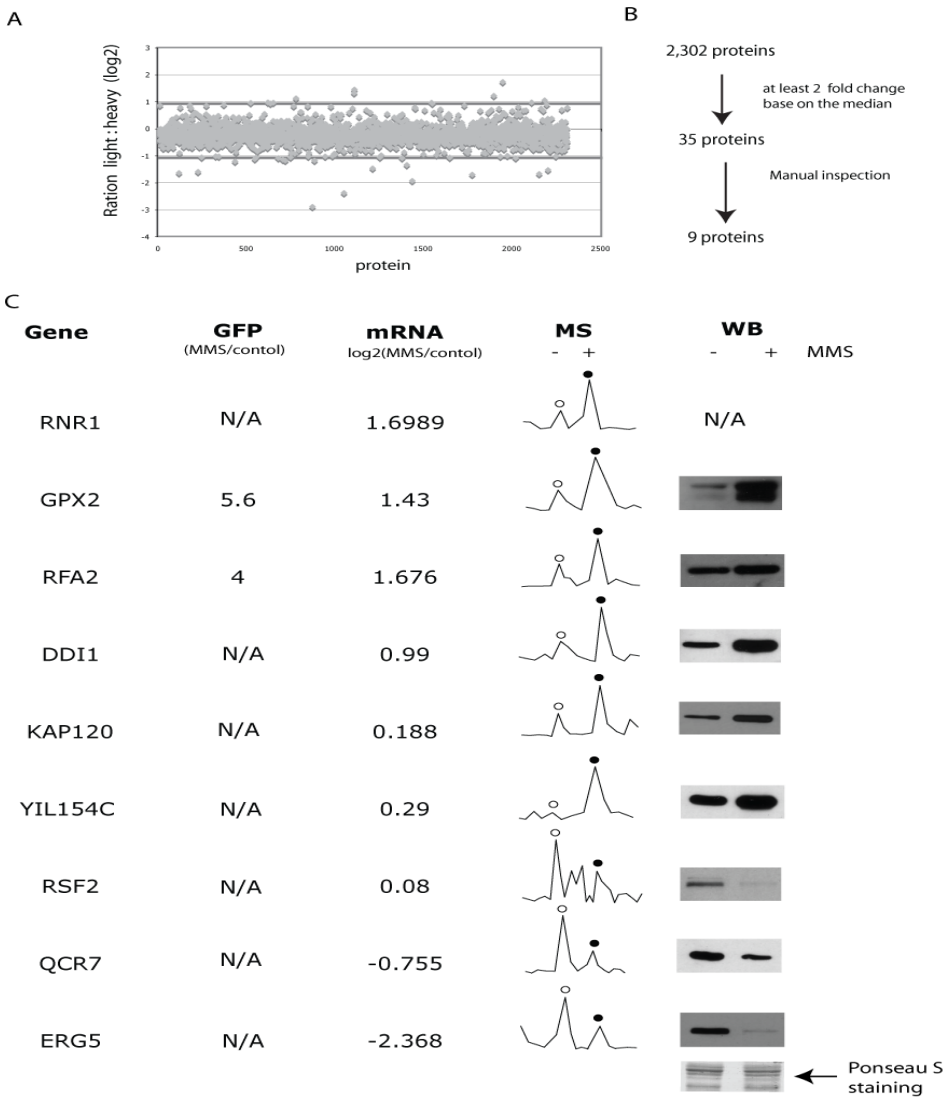

Figure 2.4 Quantification of the relative abundances of 2302 proteins identified.
(A) The median (in Log_2 scale, heavy area/light area) of the quantifications of each identified protein was shown. Most of them are within two-fold between untreated and MMS-treated cells. A threshold of over 2-fold change in the median of the quantification of each protein was applied. This led to a list of 35 candidate proteins for further manual analysis as described in the text, resulting the identification of 9 candidate proteins for further experimental validation.
B) Analysis of MMS-induced protein abundance or gene expression changes using several approaches, including GFP-tagged proteins approach (Lee et al., 2007), mRNA expression analysis (2 hours MMS treatment) (Jelinsky and Samson, 1999), MS data and Western blot results of TAP tagged genes of interest in this study. The empty or solid circle indicates the peptide derived from either untreated or MMS treated cells, respectively. Anti-TAP WB shows the specific TAP tagged protein of interest before and after MMS treatment. A representative Ponseau S staining of the membrane is shown at the bottom to indicate equal sample loading.

Chapter 3: A Multidimensional Chromatography Technology for In-depth Phosphoproteome Analysis

3.1 Abstract

Protein phosphorylation is a post-translational modification widely used to regulate cellular responses. Recent studies showed that global phosphorylation analysis could be used to study signaling pathways and to identify targets of protein kinases in cells. A key objective of global phosphorylation analysis is to obtain an in-depth mapping of low abundance protein phosphorylation in cells, which necessitates the use of suitable separation techniques due to the complexity of the phosphoproteome. Here we developed a multidimensional chromatography technology, combining Immobilized Metal Affinity Chromatography (IMAC), Hydrophilic Interaction Chromatography (HILIC) and Reverse-Phase Liquid Chromatography (RP-LC), for phosphopeptide purification and fractionation. Its application to the yeast S. cerevisiae after DNA damage led to the identification of 8,764 unique phosphopeptides from 2,278 phosphoproteins using tandem mass spectrometry (MS). Analysis of two low abundance proteins, Rad9 and Mrc1, revealed that approximately 50 % of their phosphorylation was identified via

this global phosphorylation analysis. Thus, this technology is suited for in-depth phosphoproteome studies.

3.2 Introduction

Cells are highly responsive to their environment. Protein phosphorylation is a widely used post-translational modification that regulates many biological processes in cells. The phosphoproteome, referring to the phosphorylation profile of cells, undergoes many changes in response to various stimuli. Recent advances in mass spectrometry (MS) have made it possible to identify thousands of phosphopeptides (Beausoleil et al., 2004; Bodenmiller et al., 2007a; Chi et al., 2007; Ficarro et al., 2002; Li et al., 2007). Further, various stable isotope labeling methods were used to quantify changes of protein phosphorylation in cells (Gruhler et al., 2005; Olsen et al., 2006; Smolka et al., 2007). These studies illustrated the potential of phosphoproteomics technology to study phosphorylation-mediated signal transduction processes on a large scale.

The phosphoproteome in cells is highly complex, containing phosphorylation of both high and low abundance proteins on multiple sites. In order to identify phosphorylation of low abundance proteins in a global phosphorylation analysis, it is necessary to have an in-depth mapping of the phosphoproteome. Towards this goal, several methods were developed to purify phosphopeptides prior to their MS analysis, including the use of titanium

dioxide, the immobilized metal affinity chromatography (IMAC), and others (Bodenmiller et al., 2007b; Ficarro et al., 2002; Larsen et al., 2005; Tao et al., 2005). These methods allowed a high-degree enrichment of phosphopeptides for subsequent analysis. To analyze the purified phosphopeptides, the use of reverse phase based high performance liquid chromatography (RP-HPLC) and tandem mass spectrometry (MS/MS) can identify hundreds of phosphopeptides in a single analysis; however, this alone is insufficient for an in-depth mapping of the phosphoproteome.

The Strong Cation Exchange (SCX) chromatography is commonly used to fractionate peptides (Washburn et al., 2001). Recently, SCX column was used to fractionate phosphopeptides and, as a result, many more phosphopeptides were identified (Beausoleil et al., 2004; Gruhler et al., 2005; Li et al., 2007; Smolka et al., 2007). Clearly, fractionation of phosphopeptides is necessary for a higher coverage of the phosphoproteome. However, it was found that some phosphopeptides did not bind to the SCX column and other phosphopeptides were mostly eluted from the SCX column earlier than unphosphorylated peptides ((Peng et al., 2003), also our unpublished observation). These observations indicated a relatively weak binding between the SCX resins and phosphopeptides. To analyze phosphopeptides using MS, RP-HPLC is commonly used to further fractionate the phosphopeptides. It was observed that the presence of salt, regardless whether it is volatile or not, would compromise the binding and separation of many phosphopeptides via

RP-HPLC. This is likely due to that many phosphopeptides are hydrophilic and bind less tightly to RP-HPLC. Unfortunately, desalting of phosphopeptides is typically performed using reverse phase C18-based columns. Therefore, to minimize loss of phosphopeptides, it is important to have fewer desalting steps. With these considerations, we sought to develop an alternative chromatography method for phosphopeptide fractionation. Ideally, such chromatography does not require the use of salt-containing buffers and provides an orthogonal separation compared to RP-HPLC.

Hydrophilic interaction chromatography (HILIC) is a less commonly used method for peptide fractionation, despite being often used to fractionate small metabolites (Yoshida, 2004). Interestingly, a recent study of the separation of unphosphorylated peptides using SCX, HILIC and RP-HPLC indicated that a better orthogonal separation could occur between HILIC and RP-HPLC for unphosphorylated peptides (Boersema et al., 2007; Gilar et al., 2005). The observed orthogonal separation between HILIC and RP-HPLC likely reflects their different mechanisms of separation. While RP-HPLC depends on the interaction with the hydrophobic amino acid side chains, HILIC depends on the interaction with those hydrophilic and possibly charged amino acid residues via hydrogen bonding and ionic interactions. Moreover, because phosphopeptides are generally hydrophilic and charged, one would expect that phosphopeptides should interact more strongly with HILIC than

unphosphorylated peptides. Thus, it should be possible to separate phosphopeptides using HILIC.

Here we investigated the use of HILIC for phosphopeptide separation and found that it provided a largely orthogonal separation of phosphopeptides with RP-HPLC. A multidimensional chromatography technology combining IMAC, HILIC and RP-HPLC in a sequential order was thus developed for the purification and separation of phosphopeptides. This technology was designed to have minimal manual steps and found to provide an in-depth and sensitive mapping of the phosphoproteome of the yeast S. cerevisiae after genotoxic stress.

3.2 Experimental procedures

3.2.1 Preparation of IMAC resin

All commonly used chemicals were obtained from Sigma-Aldrich unless noted otherwise. Resins from three silica-NTA spin columns (Qiagen, Valencia, CA) were added to 50 mL of buffer containing 5 mM EDTA (pH 8.0) and 1 M NaCl and incubated for 1 hour at room temperature under rotation. The resins were then spun down and washed sequentially with 50 mL water, with 50 mL 0.6 % acetic acid, and finally incubated with 50 mL of 100 mM $FeCl_3$ in 0.3 % acetic acid for 1 hour under rotation. The resins were then washed with 50 mL of 0.6 % acetic acid, then with 50 mL of a solution containing 25 % acetonitrile, 0.1 M NaCl and 0.1 % acetic acid, and then two

more times with 50 mL 0.1 % acetic acid. Finally, the resins were resuspended in 0.1 % acetic acid as a 50 % (v/v) slurry and stored at 4° C (Smolka et al., 2005).

3.2.2 Purification of phosphopeptides from whole cell protein extract

Yeast cells (RDK2669: MATα, ura3–52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3–10, ade2Δ1, ade8) were grown in 2 liters of YPD media until an Abs_{600} of 0.7, then treated by 0.05 % methyl methanesulfonate (MMS) for 3 hours and harvested. Protein extract was prepared by grinding 5 g of yeast cell pellet in an ice cold bead beater in 40 mL of lysis buffer containing 50 mM TrisHCl (pH 8.0), 150 mM NaCl, 0.2 % NonidetP-40, 5 mM EDTA, 5 mM NaF, 10 mM ß-glycerolphosphate, and 1 mM phenylmethysufonyl. Cell debris was removed by centrifugation at 30,000 G for 30 min. SDS was added to a final concentration of 1 % along with 5 mM dithiothreitol (DTT) for 5 min at 50 °C. After cooling to room temperature, iodoacetamide was added to a final concentration of 30 mM for 40 min. The proteins were precipitated using three volumes of 50 % v/v ethanol/acetone for 30 min at 4 °C. Precipitated proteins were resuspended with a buffer containing 8 M urea, 100 mM TrisHCl (pH 8.0) and the protein concentration was measured using the Bradford Assay. The sample was then diluted using 50 mM TrisHCl (pH 8.0) and 150 mM NaCl (TBS) so that the final concentration of urea is below 2 M. Six mg of proteins

were then digested by 100 µg of trypsin at 37 °C overnight. The proteolyzed sample was acidified using TFA to a final concentration of 0.2 %, and then spun down for 20 min at 4000 G to remove any insoluble material. The cleared peptide sample was loaded to a 200 mg Sep-Pak C18 column (Waters), washed 2 times with 3 mL of 1 % acetic acid, and then eluted by 600 µL of 80 % acetonitrile and 0.1 % acetic acid and dried. The dried peptides were resuspended in 100 µL of 1 % acetic acid and loaded to a gel loading tip column containing 20 µL of IMAC resin. After loading, the IMAC was washed twice with 20 µL of wash buffer containing 25 % acetonitrile, 100 mM NaCl and 0.1 % acetic acid, and then with 20 µL of 1% acetic acid, and finally eluted by 100 µL of 1 % phosphoric acid and dried under vacuum.

3.2.3 Purification of Rad9 and Mrc1 using a pulldown approach

A TAF tag containing 2 copies of protein A, TEV cleavage site, 6 histidines and 3 copies of the flag peptide was inserted at the c-terminus of either the Rad9 or Mrc1 protein using G418 selection marker and RDK2669 yeast background as previously described (Chen et al., 2007a). Rad9-TAF and Mrc1-TAF cells were grown in 2 liters of YPD media until an Abs_{600} of 0.7, then treated by 0.05 % methyl methanesulfonate (MMS) for 3 hours and harvested. Protein extract was prepared the same way as above and 500 mg of proteins were obtained. Protein extract was incubated with 0.1 mL of human immunoglobulin G (IgG)-Sepharose resin (GE Healthcare, Piscataway, NJ) for

4 hours in the cold room under rotation. The IgG resin was then washed with 10 mL of lysis buffer and then with 5 mL of TBS. The proteins were eluted from the IgG resin using 800 µL 5 % acetic acid and 1 M urea. The eluted proteins were dried under reduced pressure and then 1 M Tris base was used to neutralize the sample to pH 8. After the sample volume was adjusted to 100 µL, it was reduced by DTT and alkylated by iodoacetamide as described above. The sample was then diluted with TBS to a final concentration of 2M urea, and then digested with 2 µg of trypsin at 37 °C overnight. The resulting peptides were desalted using a 50 mg Sep-Pak C18 column (Waters, Milford, MA) and dried under vacuum. The dried peptides were resuspended in 100 µL of 1 % acetic acid and loaded to a tip-column containing 10 µL of IMAC resins. After loading, the IMAC resin was washed twice with 10 µL of wash buffer containing 25 % acetonitrile, 100 mM NaCl and 0.1 % acetic acid, then washed with 20 µl of water, and finally eluted by 100 µL of 3 % ammonium hydroxide and dried under vacuum.

3.2.4 The use of HILIC for phosphopeptide separation

A TSK gel Amide-80 column (2 mm x 15 cm 5 µm; Tosoh, Grove City, OH) was used for the HILIC experiment. Three buffers were used for the gradient: Buffer A: 90 % acetonitrile and 0.005 %TFA; Buffer B: 0.005 % TFA; Buffer C: 0.1 % phosphoric acid and 0.005 % TFA. Phosphopeptides eluted from IMAC were resuspended in 30 µl of 90 % acetonitrile with 0.1 % formic

acid and then injected into the HILIC Amide-80 column via a 100 µL loop with a flow rate of 150 µL/min. One and a half minute fractions from the HILIC were collected and dried under reduced pressure. The gradient used is shown in Figure 1B (100 % buffer A at time = 0 min, 11 % buffer B at 5 min, 29 % buffer B at 20 min, 95 % buffer C at 45 min, 95 % buffer C at 50 min and, finally, 100 % buffer A at 55 min).

3.2.5 Mass Spectrometry

Experiments were performed using the 1100 QuadPump HPLC system (Agilent, Santa Clara, CA), the Ultimate 3000 autosampler (Dionex, Sunnyvale, CA), and the LTQ tandem mass spectrometer (Thermo Fischer Scientific, San Jose, CA). Each HILIC fraction was transferred to a silanized glass insert (National Scientific, Rockwood, TN) dried under reduced pressure, and then resuspended in 10 µL 0.1 % TFA. Four microliters were loaded using the autosampler via a 5 µL sample loop directly to an in-house packed 125 µm (inner diameter) x 20 cm microcapillary RP-HPLC column, packed with 3 µm C18 resin (Magic beads; Michrom Bioresources, Auburn, CA). For RP-HPLC-MS/MS analysis, Buffer I consisted of 0.1 % formic acid and 2 % acetonitrile. Buffer II consisted of 0.1 % formic acid and 80 % acetonitrile. A 120 min gradient from 15 % to 35 % Buffer II was used. Xcalibur 2.2 software (Thermo Fischer Scientific, San Jose, CA) was used for the data acquisition, and mass spectrometer was set to perform one full MS scan followed by 6

consecutive MS/MS scans according to the ion intensities detected in the full MS scan. The minimal threshold for the dependant scans was set to 6500 counts, and a dynamic exclusion list was used with the following settings: repeat count of 1, repeat duration of 2 seconds, exclusion list size of 150, exclusion duration of 60 seconds, and exclusion mass width of 0.2 % relative to the reference mass.

3.2.6 Data analysis using SEQUEST and InsPecT

To search tandem mass spectra, a composite database was generated using both the yeast protein database (downloaded from www.yeastgenome.org on Jan 12, 2007) and its reverse protein database. The use of a decoy database allows an estimate of the false-discovery rate (Elias and Gygi, 2007). Parameters for the search were: parent mass tolerance of 3.0 Da, +80.0 Da variable modification of STY due to phosphorylation, and a maximum of 2 modifications per peptide. For the use of SEQUEST (version 3.4 beta 2), a Sorcerer system (SageN, San Jose, CA) was used and a semi-tryptic restriction was applied to the search. For the use of InsPecT (version July 12, 2007) (Tanner et al., 2005), no tryptic restriction was used in the search. After the search, the raw results of SEQUEST and InsPecT were ranked according to their provided p-value and filtered to a 1 % false discovery rate, as measured by hits to the decoy sequence. The identified phosphopeptides were then filtered to remove redundancy due to different

charge states, oxidation, repeated identification, or any possible ambiguity in phosphorylation site assignment, so that only the number of phosphopeptides with unique amino acid sequences was reported in Figure 3. This provides a simple, though underestimated, measure of the number of phosphopeptides identified in this study.

3.2.7 Phosphate Localization Score (PLscore)

To measure confidence in the assignment of phosphorylation site localization, we implemented a metric similar to the Ascore introduced by Beausoleil (Beausoleil et al., 2006), with minor changes. Binomial probabilities for observing peaks supporting a specific phosphate site localization over another were calculated similarly to the Ascore, except that a single peak density was used, as opposed to the multiple iterative peak densities in the Ascore. With this simplification, we compute a peptide score using the default peak density of Inspect, 12 peaks per 100 m/z units. This results in slightly lower scores than in Ascore. The algorithm was applied to both Inspect and SEQUEST results. This script, named PhosphateLocalization.py, is freely available from our web server (http://peptide.ucsd.edu) as part of the Inspect package.

3.3 Results

3.3.1 Experimental Strategy and Rationale

The objective was to develop a multi-dimensional liquid chromatography technology for global phosphorylation analysis. First, RP-HPLC is the standard tool to fractionate peptides immediately prior to their analysis by a tandem MS. Any additional chromatography should provide a separation that is orthogonal to RP-HPLC. Second, for a sensitive detection of phosphopeptides, the use of such chromatography should avoid the use of any desalting step that may lead to sample loss. Third, for a rapid phosphoproteome analysis, manual operations should be minimized. Based on these considerations, a technology was developed for the purification, separation and analysis of phosphopeptides using MS (see Fig. 3.1A).

As shown in Figure 3.1A, 6 mg peptides derived from proteolyzed protein extract were applied to IMAC. One percent phosphoric acid was used to elute phosphopeptides. The eluted phosphopeptides were dried under vacuum, then resuspended in a buffer containing 90 % acetonitrile and 0.1 % formic acid and loaded to a HILIC column. The phosphopeptides were then eluted from HILIC using an increasing concentration of aqueous buffers. The eluent from HILIC was dried under vacuum, resuspended in 0.1 % TFA and then loaded to RP-HPLC-MS/MS via an autosampler. TFA was found to be

important for loading of phosphopeptides to the RP-HPLC because it helps the binding of phosphopeptides to the C18 column.

As shown in Figure 3.1B, a dual-gradient program was used for the separation of phosphopeptides by HILIC using volatile buffers without any salt. In the first region of separation (grey dashed line in Fig. 3.1B), a gradient of increasing concentration of water with a constant concentration of 0.005 % trifluoroacetic acid (TFA) was used. In the second region of separation (red dashed line in Fig. 3.1B), a gradient of increasing concentration of phosphoric acid and water was introduced. The elution profiles of phosphorylated (IMAC-bound, dark blue line) and the unphosphorylated peptides (IMAC-flowthrough, light blue line) were examined. As shown in Figure 3.1B, unphosphorylated peptides generally eluted earlier than the phosphorylated peptides, indicating that phosphopeptides bind more tightly to the HILIC column than unphosphorylated peptides. The partial overlap of the phosphorylated and unphosphorylated peptides indicated that HILIC is not sufficient for phosphopeptide purification; instead, the use of IMAC or other affinity methods for phosphopeptide purification is essential. After IMAC purification, a low level of unphosphorylated peptides was often found in the earlier HILIC fractions, which are no longer considered for later analysis (see asterisks in Fig. 3.1B). This UV profile of the phosphopeptides on the HILIC is reproducible from multiple experiments (results not shown).

3.3.2 The use of HILIC to fractionate phosphopeptides

To see whether HILIC could provide a high-resolution separation of phosphopeptides that is orthogonal to RP-HPLC, the phosphopeptides in HILIC fractions 14-16 were examined. As shown in Figure 3.2A, numerous peptide ions are present throughout the RP-HPLC gradient for these fractions, underscoring the need for fractionation of phosphopeptides purified from proteolyzed cell lysate. Interestingly, each HILIC fraction has a different elution profile on the RP-HPLC and peptides are present throughout the RP-HPLC gradient, indicating an orthogonal separation of phosphopeptides between HILIC and RP-HPLC. The phosphopeptides in HILIC fractions 14-16 were then identified using SEQUEST and analyzed for potential overlaps between the fractions. As shown in Figure 3.2B, there is an approximate 20 % overlap in the identified phosphopeptides between two adjacent HILIC fractions. Interestingly, a small overlap (~ 5 %) was found between the non-neighboring HILIC fractions 14 and 16. Similar results were also observed for other HILIC fractions (not shown). These observations showed that HILIC provides a high-resolution separation of phosphopeptides that is largely orthogonal to RP-HPLC.

3.3.3 An in-depth mapping of the phosphoproteome of budding yeast after DNA damage

Next, SEQUEST and InsPecT were used to identify phosphopeptides present in 20 HILIC fractions (see Fig. 3.3A) (Tanner et al., 2005; Yates et al., 1995). A decoy protein database of the budding yeast Saccharomyces cerevisiae was included in the database search to evaluate the false-discovery rate (Elias and Gygi, 2007). With a false discovery rate of 1 %, a total of 8764 unique phosphopeptides (irrespective of their phosphorylation site assignment) from 2278 proteins were identified using both SEQUEST and InsPecT (see Fig. 3.3A). While SEQUEST was able to identify 6419 unique phosphopeptides, InsPecT identified 7681 phosphopeptides from the same dataset. Further, 5336 unique phosphopeptides were found by both InsPecT and SEQUEST (see Fig. 3.3B). Thus, SEQUEST and InsPecT are partially complementary for phosphopeptide identification. This is likely because the scoring mechanisms for SEQUEST and InsPecT are quite different (Tanner et al., 2005).

Of all the top-matched phosphorylation sites found, serine phosphorylation is most common and it constitutes approximately 83 % of all phosphorylation found, while threonine phosphorylation accounts for the rest. Further, the number of doubly phosphorylated peptides is much fewer than singly phosphorylated peptides (756 vs 8008). This is likely because multiple phosphorylated peptides are harder to be identified with high confidence using

LTQ-MS due to its low mass resolution and the use of collision-induced dissociation. As commonly observed, tandem mass spectra of phosphopeptides obtained via collision-induced dissociation are often of lower quality due to the neutral loss of phosphoric acid, which compromises the assignment of the precise phosphorylation site within a peptide. Only the number of unique phosphopeptides regardless of the phosphorylation site assignment was summarized here (see Figs. 3.3A and 3.3B). Despite being an underestimation, the large number of unique phosphopeptides identified here showed the potential of this multidimensional chromatography technology for large-scale phosphoproteomic studies.

A key question is how well phosphorylation of low abundance proteins was identified using this approach. To address this question, we chose to further examine phosphorylation of Rad9 and Mrc1, two low abundance proteins with estimated 400 and 721 copies per cell, respectively (Ghaemmaghami et al., 2003). Rad9 and Mrc1 are two key adaptor proteins in the yeast DNA damage and replication checkpoints and they are known to be phosphorylated in response to MMS treatment (Alcasabas et al., 2001; Emili, 1998). However, their phosphorylation sites have not been mapped previously. To identify the most phosphopeptides of Rad9 and Mrc1 using MS, we purified each protein from 500 mg of protein extract derived from MMS-treated Rad9-TAF and Mrc1-TAF cells using a pulldown approach (Smolka et al., 2005). Phosphopeptides of purified Rad9 and Mrc1 were further purified using IMAC

and identified. After database search using SEQUEST and InsPecT as described in the data analysis section, the identified phosphopeptides of Rad9 and Mrc1 were subjected to manual inspection. To this end, all the identified phosphopeptides of Rad9 and Mrc1 are required to be doubly tryptic and all significant fragment ions in their MS/MS spectra should be correctly assigned. The results are summarized in Figure 3.3C. While the pulldown approach identified 38 phosphopeptides from Rad9, 19 of them were also identified by the HILIC method. This corresponds to approximately 50 % coverage of the phosphorylation of Rad9. Next, the phosphopeptides of Rad9 identified here were compared to those identified using the SCX based method (Smolka et al., 2007). Despite that almost 8-fold more starting material (50 mg) was used previously (Smolka et al., 2007), more phosphopeptides of Rad9 were found using the HILIC-based method. Similar results were obtained with Mrc1 (see Fig. 3.3D). Thus, this IMAC-HILIC-RP-HPLC based technology is more sensitive, requires less manual efforts, and allows a higher coverage of the phosphorylation of low abundant Rad9 and Mrc1 in cells, compared to our previous study (Smolka et al., 2007).

3.3.4 Phosphorylation of SQ/TQ sites of budding yeast after DNA damage

As shown previously, phosphorylation of SQ/TQ sites is a relatively rare event in cells (Smolka et al., 2007). Mec1 and Tel1 are responsible for about

26 % of the observed SQ/TQ phosphorylation in cells treated with MMS. In the present study, phosphorylated SQ/TQ sites were found in 355 proteins, which is approximately 4 times the number of proteins identify by the SCX method (Smolka et al., 2007) (see Fig. 3.4A). Figure 3.4B summarizes the known and potential targets of Mec1/ Tel1 that are involved in the DNA damage checkpoint, DNA repair, and DNA replication. Several proteins that were known to undergo DNA damage induced phosphorylation, including Slx4 (Flott and Rouse, 2005), Rtt107 (Roberts et al., 2006), Mrc1 (Alcasabas et al., 2001), H2A (Downs et al., 2000) and others were identified here (see Fig. 3.4B, color in blue), most of them were not described previously (see Fig. 3.4B, colored in green). Further experiments are in progress to see whether they undergo Mec1 and/or Tel1 dependent phosphorylation *in vivo*. It should also be noted that additional targets of Mec1 and Tel1 are likely present in the database of SQ/TQ phosphorylated peptides.

3.4 Discussion

With the enormously complex phosphoproteome unveiled in several recent studies, a key objective of global phosphorylation analysis is to obtain an in-depth mapping of protein phosphorylation in cells. This is needed to identify and characterize many lower abundant and regulatory phosphorylation events in cells. To this end, a suitable multi-dimensional chromatography technology should be used to separate phosphopeptides into simpler samples

for mass spectrometric analysis. Ideally, such technology should be simple, automatable, and allow a sensitive detection of phosphopeptides.

In this study, we described the development of a multi-dimensional chromatography method based on a combination of IMAC, HILIC and RP-HPLC to purify and fractionate phosphopeptides. Several features of this technology are demonstrated. First, HILIC was found to be largely orthogonal to RP-HPLC for phosphopeptide separation. Little overlap of phosphopeptides was found between non-neighboring HILIC fractions. As a result, a higher coverage of the phosphoproteome was obtained, compared to our previous study (Smolka et al., 2007). During the reviewing of this manuscript, a recent report by McNulty et al showed that phosphopeptides bind to HILIC more tightly than unphosphorylated peptides and HILIC could be used to fractionate phosphopeptides (McNulty and Annan, 2008), which is in agreement with our observation (see Fig. 3.1B). Second, a salt-free buffer system was developed for the use of HILIC to separate phosphopeptides. As a result, the only sample-handling step used here is sample drying and resuspension, which should not cause any sample loss. This is likely the main reason why more phosphopeptides were identified with less starting material, compared to our previous study (Smolka et al., 2007). Third, because the elution buffer of HILIC contains organic solvent (acetonitrile), an offline strategy must be used between HILIC and RP-HPLC. Following sample drying and resuspension, the use of an autosampler permits the analysis of many samples continuously

without manual intervention. Finally, this technology should be compatible with various stable isotope-labeling methods for quantitative analysis of protein phosphorylation (Ong et al., 2002; Ross et al., 2004; Smolka et al., 2005).

Comparison of the phosphopeptides of Rad9 and Mrc1 detected by this multidimensional chromatography technology and the conventional pulldown method showed that approximately 50 % of the phosphorylation of Rad9 and Mrc1 was obtained from approximately 1 % of the starting material used in the multidimensional chromatography technology. Compared to the SCX based method, this HILIC based technology appears to be more sensitive and allows a higher coverage of phosphoproteome and importantly the ease of use. Further, this study allowed us to generate a more complete database of the SQ/TQ phosphorylation motifs, which is expected to contain the potential substrates of Mec1 and Tel1. We suggest that many of these newly identified proteins, with SQ/TQ phosphorylation and known roles in DNA replication and repair, could be directly phosphorylated by Mec1 and/or Tel1.

Recently, it was reported that the use of hybrid instruments, that combine the high resolution of a Fourier transform ion cyclotron resonance detector (FT-ICR) with the high rate of MS/MS acquisition of a linear ion trap, could lead to a 3-fold increase in the number of identified phosphopeptides as compared to the LTQ-MS instrument used here (Bakalarski et al., 2007). By combining the multidimensional chromatography method described here with

high mass resolution MS instruments, we anticipate that an even higher coverage of the phosphoproteome could be achieved.

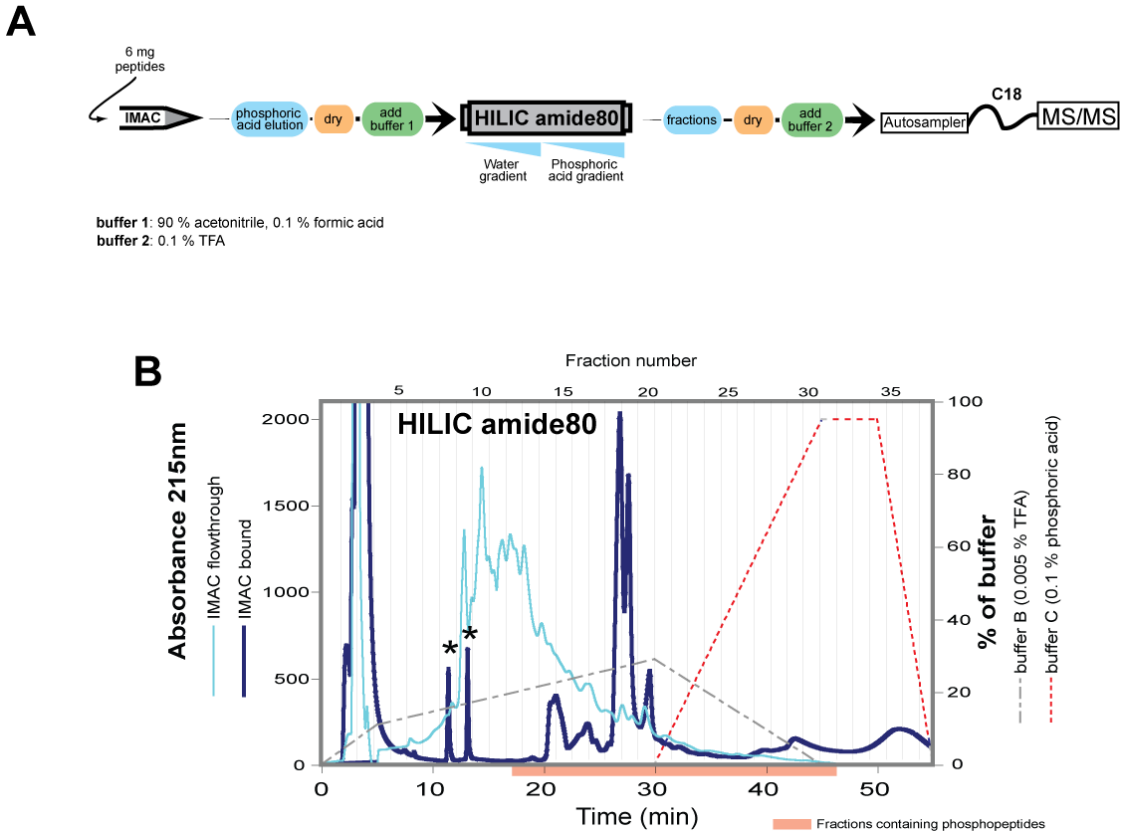

Figure 3.1 The use of HILIC for phosphopeptide
A) Schematics of a multi-dimensional chromatography technology used to purify, separate and analyze the phosphopeptides purified from proteolyzed cell lysate. B) The gradients used in HILIC (dashed lines) and the UV absorbance of the peptides in the bound fraction (dark blue line) and flowthrough (light blue line) of IMAC, separated by the HILIC. There is a partial overlap between the unphosphorylated and phosphorylated peptides. The HILIC fractions from 13-32 containing mostly phosphopeptides, colored in red, were analyzed by RP-HPLC-MS/MS. The peak appears before 10 min is due to injection and contains few peptides. Unphosphorylated peptides were mostly found in the fractions
indicated by asterisks.

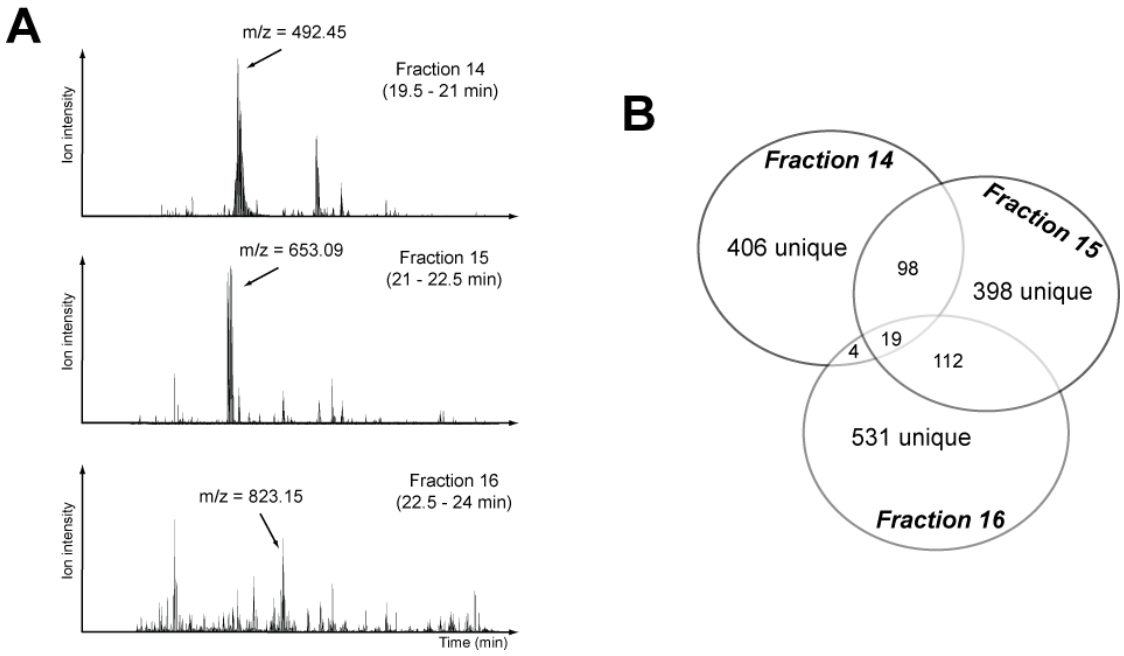

Figure 3.2 Comparison of the RP-HPLC profile and phosphopeptide identification from three adjacent HILIC fractions.
A) Ion abundances of the peptides detected by RP-HPLC-MS for HILIC fractions 14-16. B) Numerous peptide ions appear throughout RP-HPLC in each fraction, which are different from one another. B) Venn diagram indicates the overlaps of the identified phosphopeptides in HILIC fractions 14-16. While there is approximately 20 % overlap between adjacent fractions, few overlaps (~ 5 %) were found between fractions 14 and 16.

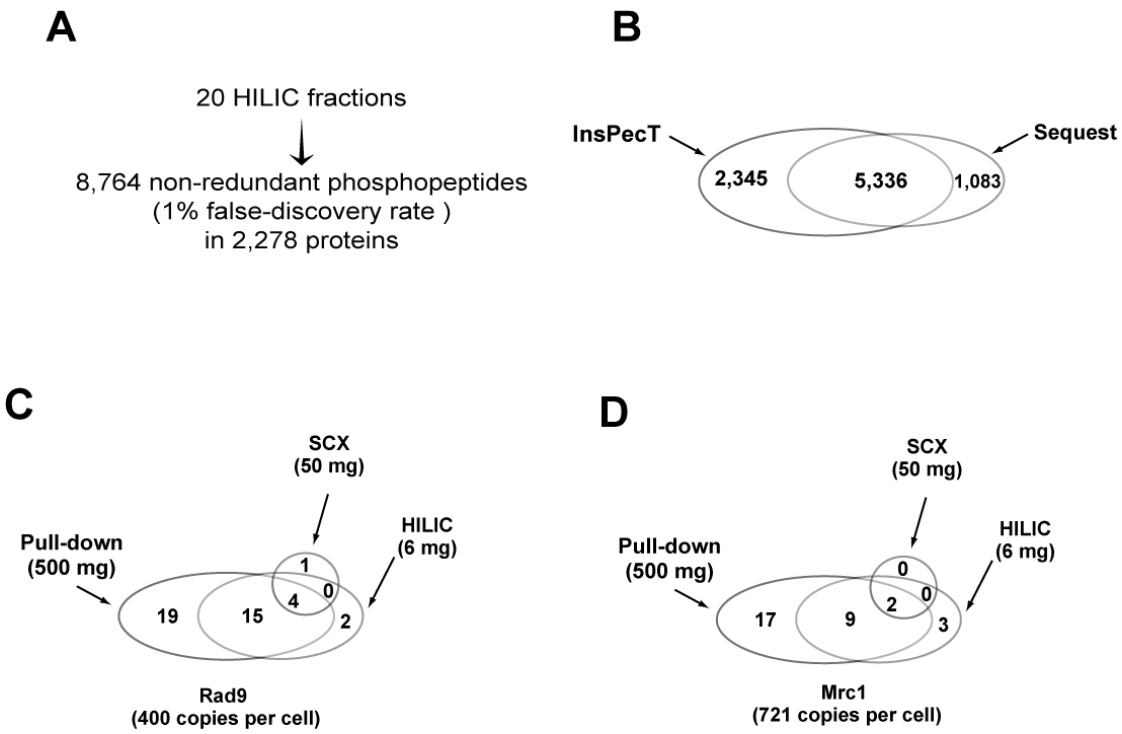

Figure 3.2 Summary of the phosphopeptides and phosphoproteins identified from yeast.
A) Phosphopeptides were purified using 6 mg of total proteins as a starting material. After separation using HILIC, 20 fractions were analyzed by RP-HPLC-MS/MS and searched using SEQUEST and InsPecT. This led to the identification of 8,764 unique phosphopeptides from 2,278 proteins. Possible redundancy due to charge state, oxidative state or even possible ambiguity on phosphorylation site assignment was removed to calculate the number of unique phosphopeptides reported here. The false discovery rate is less than 1 % as judged by the target-decoy strategy. B) Venn diagram of the phosphopeptides identified using InsPecT and SEQUEST, indicating these two search tools are partially complementary to each other. C) Venn diagram of the number of phosphopeptides of Rad9 identified using the HILIC, SCX and pulldown approaches. D) Venn diagram of the number of phosphopeptides of Mrc1 identified using the HILIC, SCX and pulldown approaches. The amount of starting material used in each case is indicated. In this case, manual examination of each phosphopeptide of Rad9 and Mrc1 was performed.

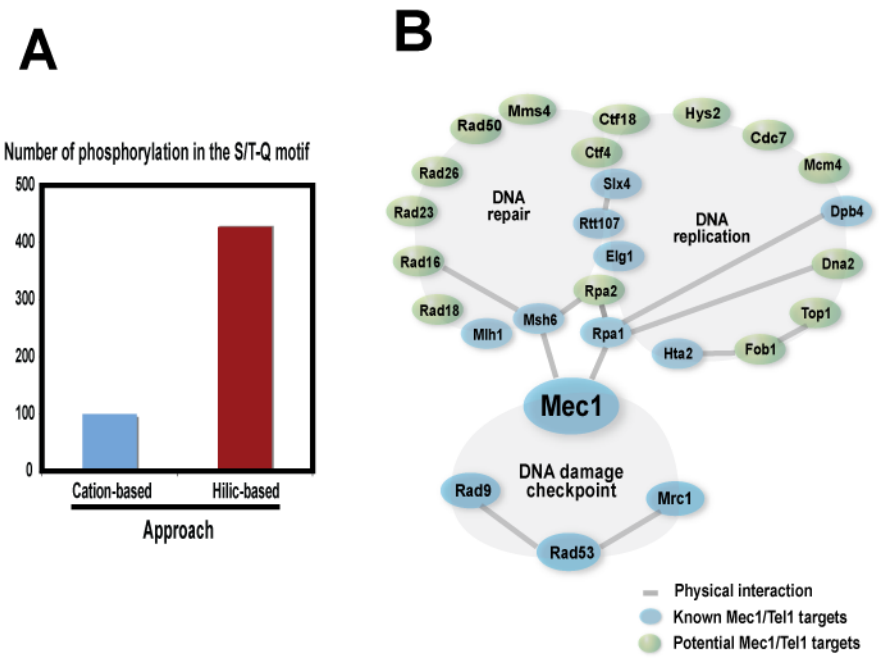

Figure 3.4 ***of Analysis of the SQ/TQ phosphorylation in the budding yeast***
(A) Comparison of the number of SQ/TQ phosphorylated peptides using either the SCX based or the HILIC based technology. About 4-fold more SQ/TQ phosphorylated peptides were found here. B) A summary of the SQ/TQ phosphorylated proteins with known roles in DNA replication, repair and DNA damage checkpoint. Those known Mec1/Tel1 targets are colored in blue, while the newly identified potential targets are colored in green.

Chapters 3, are reprints of the material as it appears in Mol Cell Proteomics, 2008, Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H. The dissertation author was the primary researcher and one of the primary author of this paper.

Chapter 4: Identification of novel *In vivo* substrates of the yeast DNA damage checkpoint

4.1 Abstract

The DNA damage checkpoint, consisting of an evolutionarily conserved protein kinase cascade, controls the DNA damage response in eukaryotes. Knowledge of the *in vivo* substrates of the checkpoint kinases is essential towards understanding their functions. Here we used quantitative mass spectrometry to identify 53 new and 34 previously known targets of Mec1/Tel1, Rad53 and Dun1 in *Saccharomyces cerevisiae*. Analysis of Replication Protein A (RPA) associated proteins reveals extensive physical interactions between RPA-associated proteins and Mec1/Tel1 specific targets. Among them, multiple subunits of the chromatin remodeling complexes including ISW1, ISW2, INO80, SWR1, RSC and SWI/SNF are identified and appear to undergo DNA damage induced phosphorylation by Mec1 and Tel1. Taken together, this study greatly expands the existing knowledge of the targets of DNA damage checkpoint kinases and provides insights into the role of RPA associated chromatins in mediating Mec1 and Tel1 substrate phosphorylation *in vivo*.

4.2 Introduction

Cells are highly responsive to their environment, especially DNA damaging agents. Damaged DNA in cells is rapidly sensed and turned into signals by the DNA damage checkpoint to control many processes, including cell cycle progression, DNA replication and repair, and gene transcription (Harrison and Haber, 2006). The DNA damage checkpoint consists of several evolutionarily conserved protein kinases (Kolodner et al., 2002; Nyberg et al., 2002). Understanding the function of the DNA damage checkpoint requires knowledge of their *in vivo* substrates. While the regulation of DNA damage checkpoint kinases has been studied extensively, the knowledge of their *in vivo* substrates is limited. This can be attributed to the lack of suitable technology to detect low abundant proteins and phosphorylation events in cells. With the use of stable isotope labeling, the advancement of high mass resolution mass spectrometry (MS), and the recent development of analytical and computational tools by many laboratories (Aebersold and Mann, 2003; Albuquerque et al., 2008; Beausoleil et al., 2006; Bodenmiller et al., 2007b; Olsen et al., 2006; Ong et al., 2002), changes in low abundant and regulatory phosphorylation in cells are increasingly detected. Combined with the use of genetics, *in vivo* kinase substrates have been identified using a quantitative MS approach (Smolka et al., 2007).

In the yeast *Saccharomyces cerevisiae*, Mec1 and Tel1, homologs of the mammalian ATR and ATM kinase respectively, function at the top of the

signal transduction cascade in the DNA damage checkpoint (Harrison and Haber, 2006; Kolodner et al., 2002; Nyberg et al., 2002). Mec1 is primarily responsible for the activation of downstream checkpoint kinases including Rad53 (Sanchez et al., 1996; Sun et al., 1996), while Tel1 has a more prominent role in regulating telomere length (Greenwell et al., 1995). Interestingly, deletion of both *MEC1* and *TEL1* leads to a synergistic increase in gross chromosomal rearrangements, indicating their redundant role in genome maintenance (Craven et al., 2002; Myung et al., 2001b). Mec1 is recruited to the site of DNA damage via Replication Protein A (RPA) that coats single-stranded DNA. Tel1 on the other hand is recruited by the Mre11-Rad50-Xrs2 complex (MRX), which recognizes DNA double stranded breaks (DSBs) (Falck et al., 2005; Nakada et al., 2005; You et al., 2005; Zou and Elledge, 2003). Importantly, DNA DSBs in cells undergo 5' to 3' resection to generate 3' single strand DNA overhangs, which recruit RPA and Mec1. Several proteins including MRX, Sae2, Dna2, Exo1 and Sgs1 participate in the DNA DSB resectioning process (Mimitou and Symington, 2008; Zhu et al., 2008). Many of them have been shown to act as substrates of Mec1 and Tel1 (Brush et al., 1996; Cartagena-Lirola et al., 2006; Mallory and Petes, 2000; Usui et al., 2001). Moreover, a number of Mec1 and Tel1 substrates have been identified, including histone H2A, Rtt107, Slx4, Cdc13, Ies4, Rad9, Mrc1 and Rad53 (Downs et al., 2000; Emili, 1998; Morrison et al., 2007; Osborn and Elledge,

2003; Roberts et al., 2006; Tanaka and Russell, 2001; Tseng et al., 2006), likely more remain to be identified (Smolka et al., 2007).

Amongst the most studied Mec1 and Tel1 substrates is the Rad53 kinase, a homolog of the mammalian Chk2 kinase (Sanchez et al., 1999; Sanchez et al., 1996; Sun et al., 1996). Rad53 helps to maintain stalled DNA replication forks (Desany et al., 1998; Lopes et al., 2001), which may be mediated by its substrates Exo1, Dbf4, and possibly others (Morin et al., 2008; Segurado and Diffley, 2008). The Dun1 kinase is phosphorylated and activated directly by Rad53 (Bashkirov et al., 2003; Chen et al., 2007b; Lee et al., 2003). Following its activation, Dun1 phosphorylates Sml1, a ribonucleotide reductase (RNR) inhibitor, leading to Sml1 degradation and increased dNTP levels in cells (Zhao et al., 1998; Zhao and Rothstein, 2002). Both Rad53 and Dun1 appear to control the phosphorylation of Rfx1/Crt1 (Huang et al., 1998), which regulates the transcriptional induction of RNR genes. Acting in parallel to Rad53, the Chk1 kinase also acts downstream of Mec1 and Tel1 (Sanchez et al., 1999), although it appears to have a more limited role compared to Rad53 in *S. cerevisiae*.

Despite extensive studies, the current knowledge of the substrates of DNA damage checkpoint kinases in yeast appears to be far from complete. Our previous study identified a number of new substrates of Mec1, Tel1 and Rad53 in yeast, yet several known substrates of these kinases were not identified (Smolka et al., 2007). Furthermore, recent MS analyses of ATM and

ATR substrates in mammalian cells have identified several hundred potential substrates (Matsuoka et al., 2007; Stokes et al., 2007). To obtain a more in-depth and comprehensive knowledge of the substrates of DNA damage checkpoint kinases in yeast, we applied a recently developed multidimensional chromatography technology (Albuquerque et al., 2008), stable isotope labeling via amino acids in culture (SILAC) (Ong et al., 2002) and high-resolution MS to determine substrates of Mec1/Tel1, Rad53 and Dun1 in cells following DNA damage treatment by DNA alkylating agent methyl methane sulfonate (MMS).

4.3 Experimental procedures

Yeast strains used and SILAC growth conditions: Standard yeast cell culture and genetic methods were used. A list of the yeast strains used is shown in Table 4.2.

4.3.1 Experimental method for a large-scale quantitative MS screen

SILAC media consisting of CSM-Lys-Arg with 2% glucose is supplemented with 30 mg/liter of lysine (either isotopically light or heavy), 20 mg/liter arginine (either isotopically light or heavy), and 80 mg/ml of proline. Heavy isotope containing L-arginine·HCl($^{13}C_6$ 98%, $^{15}N_4$ 98%) and L-lysine·HCl($^{13}C_6$ 98%, $^{15}N_2$ 98%) were purchased from Sigma Aldrich. For DNA damage treatment, WT and kinase-null cells were grown in the SILAC media to an OD_{600nm} of 0.3 and then treated with 0.05% methyl methane sulfonate (MMS) for 2 hours. Cells were then harvested and lysed using the glass bead

beating method (Albuquerque et al., 2008). For each large-scale MS screen, equal amounts of proteins extracted from WT and each kinase-null cells were combined. 5 mg of the combined protein extracts were used to purify phosphopeptides using an immobilized metal affinity column (IMAC) as previously described (Albuquerque et al., 2008; Smolka et al., 2007). Fractionation of phosphopeptides using hydrophilic interaction column (HILIC) was described previously with the following two modifications (Albuquerque et al., 2008). First, purified phosphopeptides were dried and then resuspended in 60% acetonitrile in water. The sample was then loaded into a HILIC column. Second, 0.04% trifluoroacetic acid was used in Buffer C instead of 1% phosphoric acid. After elution, each HILIC fraction was dried under vacuum, resuspended in 0.1% trifluoroacetic acid and loaded via a DIONEX autosampler to the RP-LC and Orbitrap LTQ mass spectrometer. From each the large-scale MS screen, 24 HILIC fractions were analyzed by LC-MS/MS. Each MS screen was repeated twice to ensure that similar results were obtained.

4.3.2 Analyzing phosphorylation of purified protein complexes by LC-MS/MS

TAP tagged cells were grown in 0.5 liters of SILAC media containing either light or heavy isotopically labeled lysine/arginine until OD_{600nm} of 0.3, then one of them (labeled by heavy Lys/Arg) was treated with 0.05% MMS for

2 hours. Cells were then harvested and lysed using glass bead beating method (Albuquerque et al., 2008). Typically 50 mg of total protein extract was incubated with 25 ml of IgG beads (Amersham Bioscience) for 4 hours in the cold room. Then the IgG beads were washed 3 times by 1 ml of TBS buffer. A Tev protease solution consisting of 0.5 mg of Tev protease in 200 ml TBS buffer with 1 mM DTT was added and the sample was incubated in the cold room for overnight. Next, 5-10 % of the Tev-eluted sample was analyzed by Silver staining. The remaining samples were combined, mixed with 50 ml of 8 M Urea with 0.1M Tris pH 8.0, and then processed for MS analysis as previously described (Albuquerque et al., 2008). Initially, 10% of the peptides were analyzed directly by MS. The remaining 90% of the sample was further purified using IMAC and then analyzed by MS as described previously (Albuquerque et al., 2008).

4.3.4 A method of data analysis for large-scale MS screens

Raw MS/MS spectra were searched by SEQUEST on a Sorcerer system using a composite yeast protein database, consisting of both the normal yeast protein sequences and their reversed protein sequences as a decoy to estimate false discovery rate in the search results (Beausoleil et al., 2006). The following parameters were used in the database search: semi-tryptic requirement, a mass accuracy of 10 ppm for the precursor ions, differential modification of 8.0142 Daltons for lysine and 10.00827 Daltons for

arginine, differential modification of 79.966331 Daltons for phosphorylation of serine, threonine and tyrosine, a static mass modification of 57.021465 Daltons for alkylated cysteine residues. XPRESS was used to quantify all the identified peptides using a mass tolerance of 10 ppm (Smolka et al., 2007). The search results were then filtered using the following criteria: a probability score over 0.8, Sprank scores of less than 5, mass tolerance of no more than 10 ppm, a requirement for valid XPRESS ratios, phosphorylated peptides, and the identification of the either heavy lysine or arginine labeled peptides. The use of these criteria typically yields over 10,000 phosphopeptides from all three large-scale MS screening experiments with approximately 0.02 % false discovery rate based on the number of hits in the decoy database (Beausoleil et al., 2006).

For any phosphopeptide to be considered as dependent on Mec1/Tel1, Rad53 or Dun1, a threshold of a 4-fold more abundant in WT than the corresponding kinase-null mutant was used. Further, to be considered as a Dun1-specific target, the phosphopeptide of interest must be dependent on Dun1, Rad53 and Mec1/Tel1. To be considered as a Rad53-specific target, the phosphopeptide must be Rad53-dependent, Mec1/Tel1-dependent but not Dun1-dependent. To be considered as a direct Mec1/Tel1 target, each phosphopeptide of interest must be Mec1/Tel1 dependent but not Rad53 and Dun1 dependent, and contain a phosphorylated SQ/TQ site. The use of these requirements led to the classification of specific targets of Mec1/Tel1, Rad53

and Dun1 for further manual analysis. For manual examination, the MS/MS spectra were examined to identify ions with the loss of phosphate, which is characteristic of Ser/Thr phosphorylated peptides, and to confirm that most of the major ions were properly assigned and the assignment of the phosphate group to the specific site was correct. For quantification using XPRESS (Smolka et al., 2007), a Gaussian-like elution profile was confirmed for the heavy isotope labeled peptide with at least a signal/noise ratio over 5. Most of the quantification errors were due to a failure to properly locate the elution profile of a peptide of lower intensities. Thus, manual inspection was used to eliminate such cases. Because a stringent set of criteria was used (see above), manual inspection confirmed most of the identified and quantified phosphopeptides. Moreover, most of the kinase-specific phosphopeptides were found only in WT cells and not the corresponding kinase-null cells, thus revealing a highly stringent kinase-dependency of kinase-specific phosphopeptides.

4.4 Results

4.4.1 Identifying targets of DNA damage checkpoint kinases by large-scale quantitative MS screens

Three independent large-scale MS screens were performed to identify specific targets of Mec1/Tel1, Rad53 and Dun1 (See Fig. 1A for experimental strategy). Using results from the Mec1/Tel1 screen as an example, the ratios

of the integrated intensity of heavy over light isotopically labeled phosphopeptides are plotted (see Fig. 1B). As shown in Figure 1B, the median of the relative abundance ratios is close to 1, and most phosphopeptides show a relative abundance ratio within 4-fold. Compared to the *mec1*Δ *tel1*Δ cells, significantly more phosphopeptides are 4-fold more abundant in WT cells than those that are 4-fold less abundant in WT cells. Therefore, the deletion of *MEC1* and *TEL1* causes a preferential loss of some phosphorylation. Based on this observation, any phosphopeptide that is 4-fold more abundant in WT than in *mec1*Δ *tel1*Δ cells is considered Mec1/Tel1 dependent. Similar results are also found for the Rad53 and Dun1 screens (not shown), thus a 4-fold cutoff is used to determine Rad53 and Dun1 dependent phosphorylation in the large-scale MS screens using *rad53*Δ and *dun1*Δ mutants.

Over 14,000 unique phosphopeptides from 2186 proteins are found in the Mec1/Tel1 screen. Using a 4-fold ratio as a cutoff, 238 Mec1/Tel1-dependent phosphopeptides are found (see Fig. 1C). Using the same criterion, 177 and 77 phosphopeptides are found to be Rad53-dependent and Dun1-dependent, respectively. Because Mec1 and Tel1 control the activity of Rad53, which in turn controls the activity of Dun1, we next examine the overlap among the kinase-dependent phosphopeptides. 17 phosphopeptides are dependent on Mec1/Tel1, Rad53 and Dun1; thus they are considered candidates for being Dun1-specific targets. 42 phosphopeptides are dependent on both Mec1/Tel1 and Rad53, yet independent of Dun1; thus they are considered

candidate Rad53 targets. Finally, 169 phosphopeptides are Mec1/Tel1 dependent and independent of both Rad53 and Dun1. Among them, 87 phosphopeptides contain the expected consensus phosphorylation site, SQ/TQ, of Mec1 and Tel1 (Kim et al., 1999). Therefore they are considered to be Mec1/Tel1-specific. These kinase-specific phosphopeptides were subjected to further manual examination (see Methods for details), leading to the identification of 75, 33, and 13 phosphopeptides specific to Mec1/Tel1, Rad53 and Dun1, respectively (see Fig. 1C). They are derived from 58, 24 and 12 potential substrates of Mec1/Tel1, Rad53 and Dun1, respectively (see Table1). Due to the large scope of these large-scale MS screens, some phosphopeptides might be found in one MS screen but not the others. Here we chose to consider only those phosphopeptides that satisfy the criteria above, which may lead to the loss of some potential kinase targets.

4.4.2 Functional classification and physical interactions of the kinase-specific targets

Among the 87 kinase-specific targets, 15 of them are known targets from previous studies, 19 of them were also identified as targets in our prior study (Smolka et al., 2007), and 53 targets are identified here for the first time (see Fig. 2A). Thus this study greatly expands the existing knowledge of the substrates of Mec1/Tel1, Rad53 and Dun1. Moreover, most known substrates of these kinases are found here, which not only validates our approach, but

also illustrates the scope and depth of the present study. However, we cannot exclude the possibility that variations in cell cycle stages and/or prolonged MMS treatment might alter protein abundance of some of the targets. For example, DNA damage induced expression of Rnr3 and Rad54 are known (Cole and Mortimer, 1989; Zhou and Elledge, 1993), which may contribute to their elevated phosphorylation in WT cells. Nevertheless, in most other cases, either there is no previous report of DNA damage induced protein abundance changes, or our own studies show that their protein abundances are not affected by MMS treatment (see below). Furthermore, most WT and kinase-null cells are expected to be in the S phase following MMS treatment. Possible variations of cell cycle stages between asynchronized WT and kinase-null cells are likely insufficient to account for the complete loss of phosphorylation on most kinase-specific targets in the corresponding kinase-null mutants. Therefore, most of these findings here are likely direct kinase substrates, although further studies are needed to establish each newly identified substrate as direct and functional.

Functional classification shows that most Mec1 and Tel1 specific targets function in nuclear processes, including mRNA transcription, DNA damage checkpoint, DNA replication and repair, chromatin remodeling and cell cycle control (see Table 4.1 and Fig. 4.2B). Among them are ten previously known substrates including Hta1, Rad9, Mrc1, Rfa1, Rfa2, Rtt107, Slx4, Ies4, Rad53 and Mec1. Eleven proteins were also previously identified

(Smolka et al., 2007), including Cbf1, Dpb4, Msh6, Spt7, Cbf5, Prp19, Ysh1, Hsp12, Hpr1, Sum1 and Asg1. It should be noted that the method used previously differs greatly from the one used here, which may explain the differences in the findings (Smolka et al., 2007). In the end, 37 new candidate substrates of Mec1 and Tel1 are identified here. Among Rad53 specific targets, Dun1, Mrc1, Nup1, Nup60, Dbf4, Rfx1 and Rph1 are known from previous studies. In addition to nuclear transport proteins (Smolka et al., 2007), proteins involved in chromatin remodeling and DNA replication/repair are among the newly identified Rad53 specific targets (see Fig. 2B and Table 1). Dun1-specific targets are classified as having a partial or complete dependence on Dun1. A partial Dun1 dependence is detected for Ecm21, Rnr3, Hpc2, and Sec3, suggesting that these proteins may be indirect rather than direct substrates of Dun1. A complete Dun1 dependence is found for Gle1, Npl3, Nup159, Mrp8, Rco1, Rlp7 and Mlp1 (see Tables 4.1), suggesting that these proteins may be potential Dun1 substrates. Because the consensus phosphorylation sites for Rad53 or Dun1 are not well defined (Chen et al., 2007b; Smolka et al., 2007), the establishment of *bona fide* substrates of Rad53 and Dun1 requires further studies.

Protein-protein interactions often mediate substrate phosphorylation by protein kinases *in vivo*. A summary of the known physical interactions between the kinase-specific targets identified here is shown (see Fig 4.2C). Among Mec1 and Tel1 specific targets, Taf9, Taf10 and Taf12 are components of the

TFIID complex, which interact with Spt7 in the SAGA complex and the Paf1 complex including Cdc73 and Hpr1. Dad1 and Dad3 are components of the Dam1 complex involved in chromosome segregation. Rco1, Eaf3, and Sin3 are subunits of the histone deacetylase Rpd3S complex. Dpb4 and Isw2 are involved in chromatin remodeling, along with several other chromatin remodeling complexes (see Table 4.1). Thus there appears to be extensive physical associations between Mec1/Tel1 specific targets. Interestingly, multiple checkpoint kinases can also phosphorylate the same protein complex. Rpd3S is a histone deacetylase complex whose Rco1 subunit shows Dun1-specific phosphorylation, while Sin3 and Eaf3 subunits undergo Mec1/Tel1 dependent phosphorylation. Moreover, Mec1/Tel1 and Rad53 appear to phosphorylate different residues of Mrc1, Rtt107 and Mcd1. Thus the large-scale quantitative MS screens have uncovered a complex phosphorylation regulation of these targets.

4.4.3 Mec1/Tel1 substrates may transiently associate with RPA via chromatin

Mec1 is recruited to the site of DNA damage via its interaction with RPA coated single stranded DNA (Nakada et al., 2005; Zou and Elledge, 2003), which may mediate the phosphorylation of Mec1 and/or Tel1 substrates *in vivo*. To test this hypothesis, a quantitative MS approach was used to identify RPA-associated proteins (see Fig. 3A). Silver staining of the one-step purified

RPA sample reveals two specific bands corresponding to Rfa1 and Rfa2 (see Fig. 4.3B). No other specific bands could be detected by Silver staining, suggesting that other RPA-associated proteins are less abundant and likely to interact with RPA in a more transient manner. To identify the lower abundant RPA-associated proteins, the remainder of the sample was analyzed directly using quantitative MS. Three examples of MS data of a representative peptide from Mec1, Taf14, and Isw2 are shown (see Fig. 4.3C). In each case, it is identified as exclusively labeled by heavy Lys/Arg and is thus from the RPA purification, suggesting that it is specifically associated with RPA. The same criterion is used to determine other RPA associated proteins (see Fig. 4.3D). Among the RPA-specific associated proteins, subunits of histones were found along with subunits of chromatin remodelers including Isw1, Isw2, Ies3, Ino80, Taf14, Rtt102 and Rsc8 (Clapier and Cairns, 2009). These findings suggest that RPA may associate with these proteins via DNA/nucleosome structures with DNA possibly bridging their interactions. As expected, known direct RPA binding proteins including Mec1, Dna2 and Rad52 are also detected as RPA-specific. Interestingly, There are extensive physical interactions between the RPA-specific associated proteins identified here (indicated by open circles in Fig. 4.3D) and the Mec1/Tel1 specific targets identified in the large-scale MS screens (indicated by stars). Moreover, Isw2 and Msh6 are both RPA-associated and identified above as candidates for being Mec1/Tel1 substrates

(dark circles). Thus, RPA-associated chromatin may contribute to Mec1/Tel1 substrate phosphorylation *in vivo*.

4.4.4 RPA, TFIID-SAGA, and chromatin remodeling complexes undergo MMS induced phosphorylation of SQ/TQ sites

To further examine the role of RPA in mediating the phosphorylation of Mec1 and Tel1 substrates, we chose to purify RPA, TFIID-SAGA and chromatin remodeling complexes including ISW1, ISW2, INO80, SWR1, SNF/SWI and RSC from both untreated and MMS-treated cells. Using SILAC and quantitative MS, MMS-induced phosphorylation of the above protein complexes were identified and quantified. To address whether there are any MMS-induced changes in protein abundance, the IMAC-unbound fractions containing the unphosphorylated peptides were also analyzed. Unless noted otherwise, no MMS-induced change in the abundance of these proteins was found (results not shown), thus the observed changes in phosphorylation are not due to protein abundance changes. Since Mec1 and Tel1 prefer to phosphorylate SQ/TQ sites, MMS-induced SQ/TQ phosphorylation of these proteins are shown, together with results from the above large-scale MS screens (see Fig. 4.4).

The analysis of purified RPA identifies the MMS-induced SQ/TQ phosphorylation of all three subunits of RPA as well as its associated proteins including Mec1, Msh6, Ies1 and Dna2 (see Fig. 4.4A). The analysis of purified

TFIID-SAGA identifies the same MMS-induced SQ/TQ phosphorylation sites of Taf9, Taf10, Taf12 and Hsp26 as found in large-scale MS screens, in addition to a new MMS-induced SQ/TQ phosphorylation of Sgf73 (see Fig. 4.4B). Among the chromatin remodelers, analysis of the ISW1 complex reveals two MMS-induced SQ/TQ phosphorylation sites of the Ioc2 subunit. In addition to the known phosphorylation of histone H2A on S129, a new MMS-induced TQ phosphorylation of H2B (T129) is identified (see Fig. 4.4C). For the ISW2 complex, the same SQ/TQ phosphorylation sites on Isw2, Dpb4 and Rfa1 are found to be both MMS-induced and Mec1/Tel1-specific (see Fig. 4.4D). Thus, the ISWI family of chromatin remodeling complexes appears to be phosphorylated by Mec1 and Tel1.

The INO80 and SWR1 complexes are known to be involved in the DNA DSB response and recruited to the site of DNA damage (Morrison and Shen, 2009). Ies4, a subunit of the INO80 complex, was previously found to be a target of Mec1 and Tel1 (Morrison et al., 2007). Our analysis of the INO80 complex identifies MMS-induced SQ/TQ phosphorylation of Ino80 and Ies1, in addition to Ies4 (see Fig. 4.4E). Furthermore, several MMS-induced SQ/TQ phosphorylation sites are found in proteins that may be transiently associated with INO80, including Bdf1, Rfa1, Leo1, Msh6, Hta1, Prp19, Dpb4 and Spn1. SWR1, another INO80 family chromatin remodeling complex in yeast, was also purified and analyzed (Clapier and Cairns, 2009). Analysis of SWR1 identifies the same SQ/TQ phosphorylation of Bdf1 and a new MMS-induced

SQ/TQ phosphorylation of Swr1 (see Fig. 4.4F). RSC is an essential chromatin remodeling complex in yeast (Clapier and Cairns, 2009). Previous studies have implicated its role in DNA DSB response (Chai et al., 2005; Liang et al., 2007; Morrison et al., 2004; Shim et al., 2005). MMS-induced SQ/TQ phosphorylation on five subunits of RSC, including Sth1, Rsc2, Rsc4, Rsc30 and Arp7, are identified from purified RSC (see Fig. 4.4G). Again, MMS-induced SQ/TQ phosphorylation of RSC-associated proteins including Rfa1, Dpb4, Hta1, Hsp26, Leo1 and Yta7 are found. Analysis of purified SWI/SNF identifies MMS-induced SQ/TQ phosphorylation of Swi3, as well as Arp7, Prp19 and H2A (see Fig. 4.4H). Taken together, both INO80 and RSC families of chromatin remodelers appear to be targeted by Mec1 and Tel1, and they show extensive interactions with histones, RPA and each other.

In summary, the results from the large-scale MS screens and the analysis of purified protein complexes show that RPA, TFIID-SAGA, histones and multiple chromatin remodeling complexes form a large protein interaction network targeted by Mec1 and Tel1. Although using purified protein complexes will give more information about their phosphorylation, a significant percentage of these proteins are already identified from the large-scale MS screens; illustrating the scope and depth of these screens (see Fig. 4.4).

4.4.5 Involvement of chromatin remodeling complexes in the DNA damage response

To further examine the role of RPA-associated proteins, in particular the chromatin remodeling complexes, in the DNA damage response, we examined the DNA damage sensitivities of null mutants of non-essential subunits of these complexes. DNA damage hypersensitivities have been reported previously for mutations to various subunits of RSC and INO80 (Chai et al., 2005; Morrison et al., 2004). A total of 35 deletion mutants are examined. Among them, deletion of *TAF14, ADA1*, *SPT7, SPT20, ARP5, ARP8, SWI3, SNF5* and *SNF6* causes strong sensitivities to both MMS and hydroxyurea (see Fig. 4.5). On the other hand, mutants with deletion of *ISW1* or *ISW2* have relatively minor DNA damage sensitivities under the same conditions. Ada1, Spt7 and Spt20 are subunits of the SAGA complex. Thus it is not surprising that similar patterns of DNA damage hypersensitivity are observed for mutants devoid of these proteins. Arp5 and Arp8 are subunits of the INO80 complex and their mutation also causes similar DNA damage hypersensitivities. Lastly, Swi3, Snf5 and Snf6 are subunits of the SWI/SNF complex and their mutations lead to strong DNA damage hypersensitivities. These observations are consistent with the idea that these chromatin remodelers are involved in the DNA damage response.

4.5 Discussion

The DNA damage checkpoint regulates cellular responses to DNA damage (Harrison and Haber, 2006; Kolodner et al., 2002; Nyberg et al., 2002), yet its molecular basis is far from being understood. Therefore it is necessary to identify and characterize substrates of the checkpoint kinases. In the present study, multiple large-scale and quantitative MS screens are used to determine kinase-substrate relationships in the yeast DNA damage checkpoint. This leads to the identification of 53 new targets in addition to 34 previously known substrates of these checkpoint kinases; thus greatly expanding the present knowledge of their substrates. Quantitative analysis of RPA associated proteins uncovers a large protein-protein interaction network involving RPA, histones, TFIID-SAGA and multiple chromatin remodeling complexes. In particular, the analysis of chromatin remodeling complexes identifies them as a new family of potential Mec1 and Tel1 substrates. These results strongly support that RPA-associated chromatin may mediate the phosphorylation of Mec1 and Tel1 substrates *in vivo*. The identification of multiple chromatin remodelers as potential Mec1 and Tel1 substrates here also raises the possibility that they could contribute to the genome maintenance function of Mec1 and Tel1 (Myung et al., 2001b; Putnam et al., 2009). Functional analysis of the phosphorylation of chromatin remodelers will take a more focused scope than the present study, which nevertheless

provides the initial evidence for the extensive links between Mec1 and Tel1 and this family of proteins.

Global analysis of the targets of the DNA damage checkpoint kinases here reveals several interesting features. First, Mec1/Tel1, Rad53 and Dun1 have functionally distinct targets *in vivo*. While most of the targets of Mec1 and Tel1 appear to participate in major DNA and RNA metabolisms such as mRNA transcription, DNA replication and repair, many targets of Rad53 and Dun1 function in nuclear transport. Second, there exists a complex phosphorylation regulation of some of the targets. For example, phosphorylation of Mrc1, Mcd1, Rfx1, Hpc2 and Rtt107 appears to be controlled by multiple kinases. In addition, phosphorylation of multiple subunits of the same protein complex is observed for TFIID-SAGA, INO80, RSC and others. The complex phosphorylation regulation identified here should help to direct functional studies of these proteins in future studies. Third, the identification of RPA associated proteins suggests that RPA-associated chromatins may mediate Mec1 and Tel1 signaling *in vivo*. This notion is reinforced by the observation that phosphorylation of Rfa1 and H2A is often detected in multiple protein complexes including RPA itself, INO80, ISW2 and RSC complexes (see Fig. 4.4). The extensive physical associations between RPA and chromatin remodeling complexes, which may be mediated by DNA structures, are also consistent with the known recruitment of INO80, RSC and SWR1 to the site of DNA damage (Chai et al., 2005; Liang et al., 2007; Morrison et al., 2004; Shim

et al., 2005). On the other hand, not all Mec1/Tel1 substrates may associate with RPA and may be phosphorylated in an RPA-independent manner, which remains to be determined. Finally, aside from chromatin remodeling complexes, Mec1 and Tel1 also control the phosphorylation of proteins involved in DNA replication, DNA repair and chromosome segregation (see Table 4.1). In most cases, their phosphorylation by Mec1 and Tel1 has not been previously described. The findings here thus open many new opportunities for further mechanistic and functional studies.

The DNA alkylating agent, MMS, used here likely causes DNA damage throughout the genome and may broadly induce phosphorylation of many targets of the DNA damage checkpoint. While the use of MMS treatment in this study allows us to identify a broad spectrum of potential substrates of the checkpoint kinases, we speculate that distinct substrates may be phosphorylated in a DNA damage-specific and temporal manner. Understanding how each substrate is recruited to a specific DNA damage in the genome and determining the function of its phosphorylation in the DNA damage response represents an important future direction.

Table 4.1 Functional classification and summary of the substrates of Mec1/Tel1, Rad53 and Dun1 identified in the large-scale quantitative MS screens.

Functions	Mec1/Tel1	Rad53	Dun1
DNA Damage Checkpoint	Rfa1, Rfa2, Psy4, Mec1, Mrc1, Rad9, Rad53	Mrc1, Dun1, Rfx1, Rad53	Dun1, Rnr3
RNA PolII Transcription	Bdf1, Taf9, Taf10, Taf12, Spt7, Cdc73, Hpr1, Dst1, Toa2, Spn1		
Histone Modification	Sin3, Eaf3, Hta1	Rph1, Hhf1,	Rco1
Chromatin Remodeling	Swi3, Ioc2, Isw2, Dpb4, Ies4	Hpc2, Esc1	Hpc2
mRNA processing	Prp19, Cbf5, Pwp2, Ysh1, Abd1, Bcd1	Enp1	Npl3
Cell Cycle Regulation	Sap185, Ndd1, Bir1, Dad1, Dad3, Cbf1, Mcd1	Mcd1, Ycg1, Cep3	
DNA Replication & Repair	Pol31, Msh6, Rtt107, Rad23, Rad26, Slx4, Sir4, Sum1	Rad54, Rtt107, Rad27, Dbf4, Pol1	
Nuclear Transport	Nup60	Nup1, Nup2, Nup60	Gle1, Nup159, Mlp1
Stress Response	Dak1, Hsp26, Hsp12, Zpr1, Asg1, Nma111	Plm2	
Others	Lsc2, Spe4, Acs2, Met12	Pho81, Fun30, Nug1	Ecm21, Sec3, Mrp8

Table 4.2 Yeast strains used in this study

Name	Genotype	
HZY201	*MAT**a**/MATα, ura3-52/ura3-52, leu2Δ1/leu2Δ1, trp1Δ63/trp1Δ63, his3Δ200/his3Δ200, lys2ΔBgl/lys2ΔBgl, hom3-10/hom3-10, ade2Δ1/ade2Δ1, ade8/ade8, arg4Δ, sml1Δ::TRP1, mec1Δ::HIS3, tel1Δ::URA3*	*
HZY249	*MAT**a**/MATα, ura3-52/ura3-52, leu2Δ1/leu2Δ1, trp1Δ63/trp1Δ63, his3Δ200/his3Δ200, lys2ΔBgl/lys2ΔBgl, hom3-10/hom3-10, ade2Δ1/ade2Δ1, ade8/ade8, arg4Δ, sml1Δ::TRP1, rad53Δ::HIS3, dun1Δ::URA3*	*
HZY400	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, RFA1-TAP::Kan*	**
HZY401	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, TAF9-TAP::Kan*	
HZY278	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, ISW1-TAP::Kan*	
HZY408	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, ISW2-TAP::Kan*	
HZY420	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, IES1-TAP::Kan*	
HZY277	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, SWR1-TAP::Kan*	
HZY428	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, STH1-TAP::Kan*	
HZY427	*MAT**a**, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys2ΔBgl, hom3-10, ade2Δ1, ade8, arg4Δ, sml1Δ::TRP1, SWI2-TAP::Kan*	

*: All strains were derived from SCY003. Haploid strains used for screening Mec1/Tel1, Rad53, and Dun1 substrates in Figures 1 and 2 were generated from spores of HZY201 and HZY249.

**: TAP: 6xHis-3xFLAG-Tev-ProteinA as described in Chen, S-H. *et al, JBC* 282, 986-995 (2007)

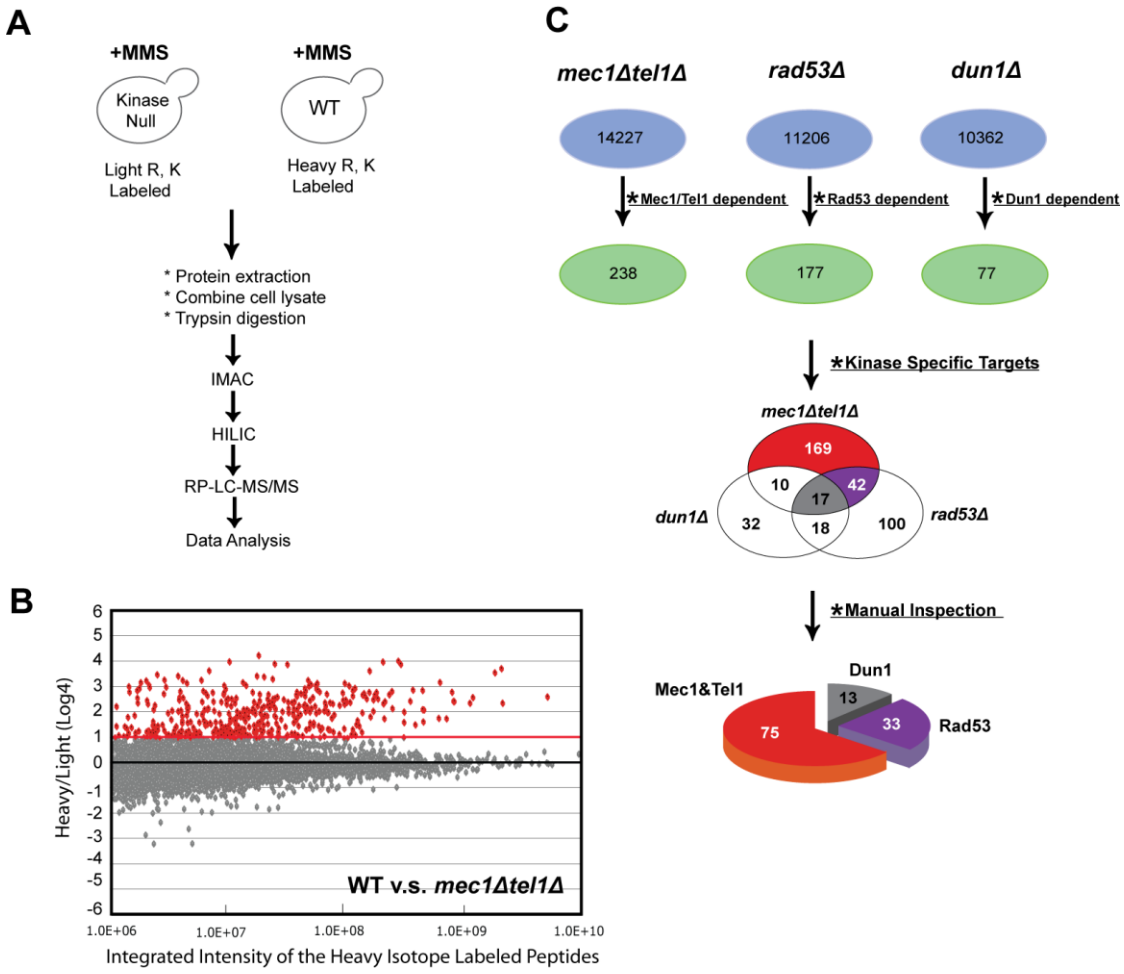

Figure 4.1 Quantitative MS identifies the targets of DNA damage checkpoint kinases
A) Strategy used in identification of kinase-specific phosphopeptides (see Experimental Methods for details). B) Abundance ratio as a function of the integrated intensity for each phosphopeptide identified in the Mec1/Tel1 screen. A Log10 scale is used for the integrated ion intensity of each heavy Lys/Arg labeled phosphopeptide, while a log4 scale is used for the corresponding abundance ratio of heavy over light Lys/Arg labeled phosphopeptide. C) Steps toward identification of kinase-specific targets. A 4-fold cutoff was used to determine kinase-dependent phosphopeptides from the Mec1/Tel1, Rad53 and Dun1 screens, whose overlaps are analyzed using Venn diagrams (see text for further details).

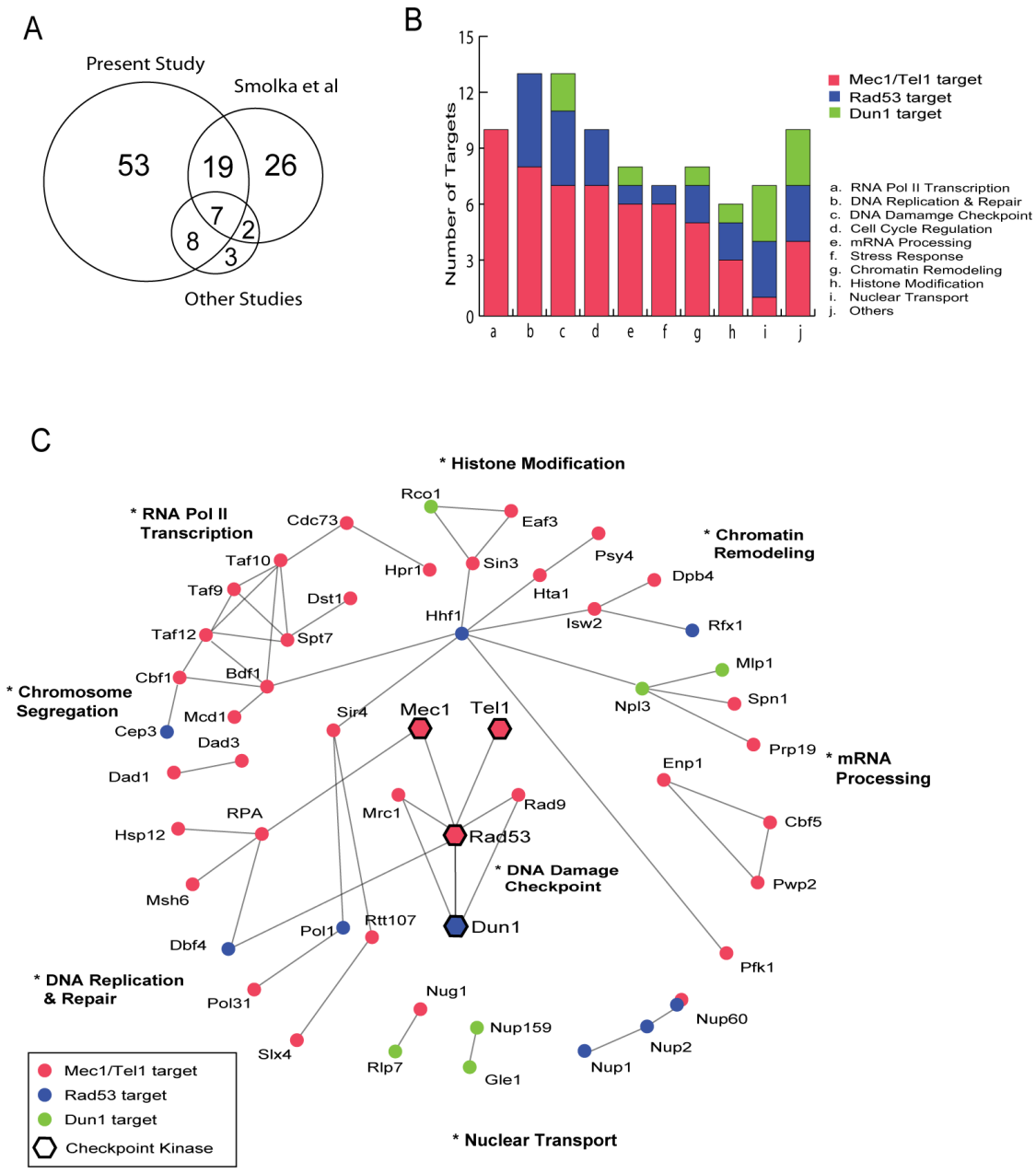

Figure 4.2 Summary of the results from large-scale MS screens
Comparison between the present study, a previous study by Smolka *et al* (Smolka et al., 2007), and all other studies. B) Functional classification of the kinase-specific targets identified in our present study (also see Table 1). C) A partial summary of Mec1/Tel1-specific targets are colored in red, Rad53-specific targets are colored in green, and Dun1-specific targets are colored in blue. Solid lines indicate the known physical interactions between these targets.

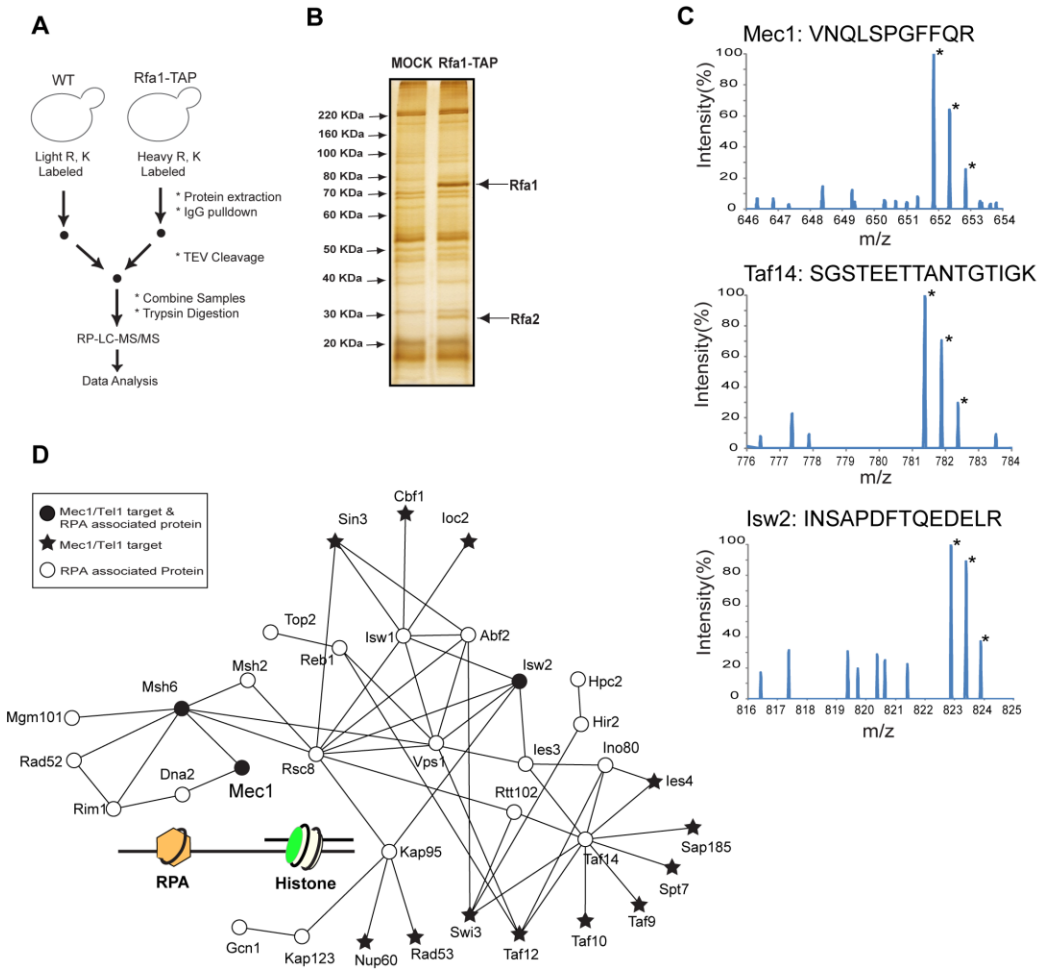

Figure 4.3 Quantitative analysis of RPA-specific associated proteins
A) Strategy used to identify specific associated proteins of RPA using Rfa1-TAP cells and the SILAC method. B) Silver staining of 5% of the one-step affinity purified RPA compared to mock purified sample. Arrows indicate specific bands corresponding to Rfa1 and Rfa2. C) Examples of MS data of a representative peptide derived from each RPA-specific binding protein. In each case, the corresponding light-isotope containing peptide is below the detection limit, while the heavy Lys/Arg containing peptide is detected. Asterisks indicate the natural occurrences of C^{13}/N^{15} in the peptide, as revealed by high resolution MS. D) Summary of RPA-associated proteins (indicated by filled and open circles, along with Mec1/Tel1 substrates identified form the large-scale MS screens (indicated by a star). Solid lines indicate known physical interactions between them. Filled circles indicate the proteins found as both RPA-specific and Mec1/Tel1 substrates. Open circles indicate those only found as RPA-specific associated proteins.

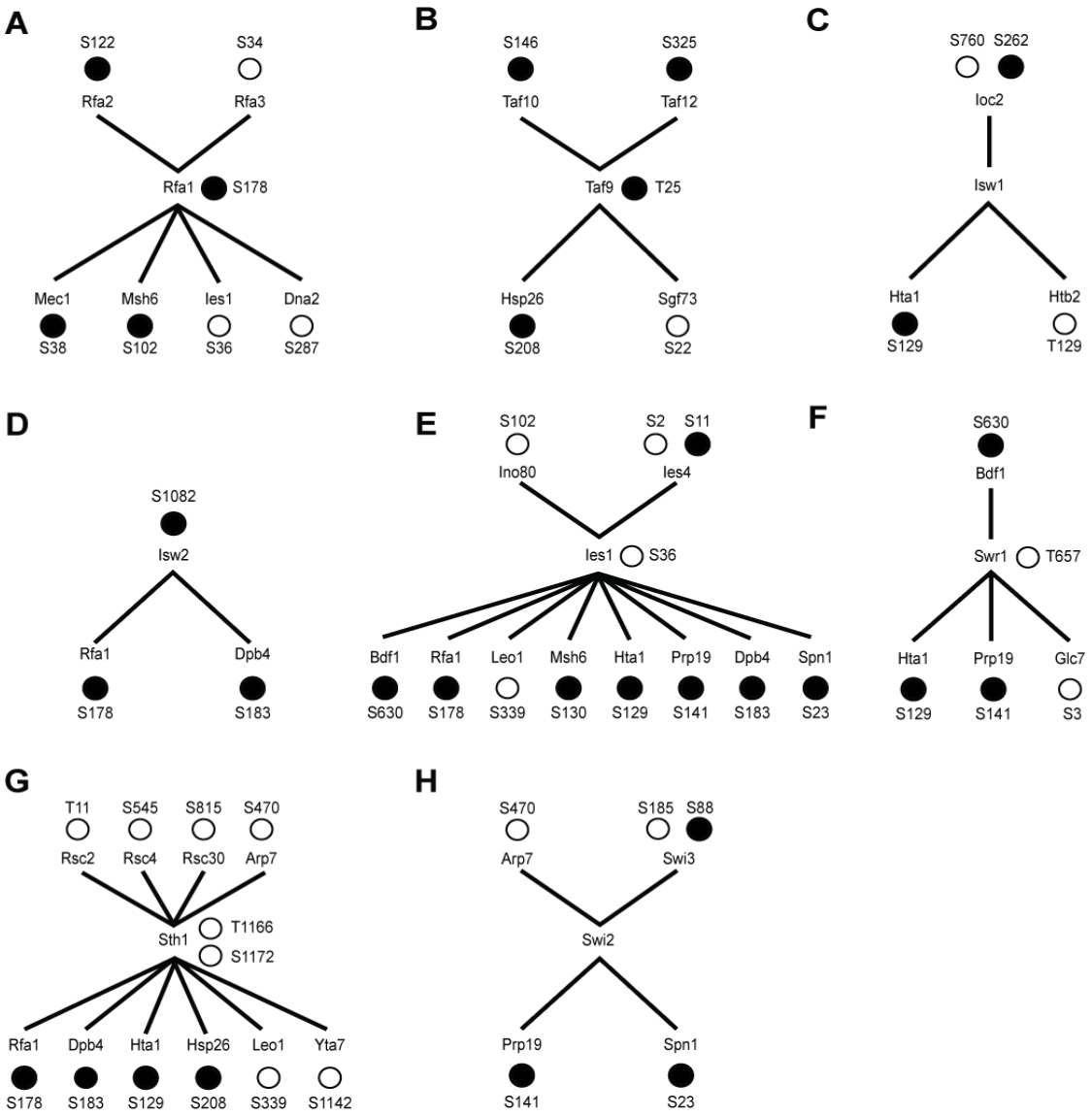

Figure 4.4 Characterization of MMS-induced SQ/TQ phosphorylation of RPA, TFIID-SAGA, ISW1, ISW2, SWI/SNF, RSC, SWR1 and INO80 complexes

Only SQ/TQ phosphorylation sites of each protein complex and its associated proteins are shown. The number indicates the position of phosphorylated serine or threonine in each protein. Solid lines indicate known physical associations between them. Filled circles indicate the phosphorylation site being detected as both Mec1/Tel1-specific from the large-scale MS screens and MMS-induced from the protein complex analysis, while empty circles indicates that this phosphorylation is only found to be MMS-induced from protein complex analysis. For protein complex purification, TAP tagged yeast strains are used as following: A) Rfa1-TAP. B) Taf9-TAP. C) Isw1-TAP. D) Isw2-TAP. E) Ies1-TAP. F) Swr1-TAP. G) Sth1-TAP. H) Swi2-TAP.

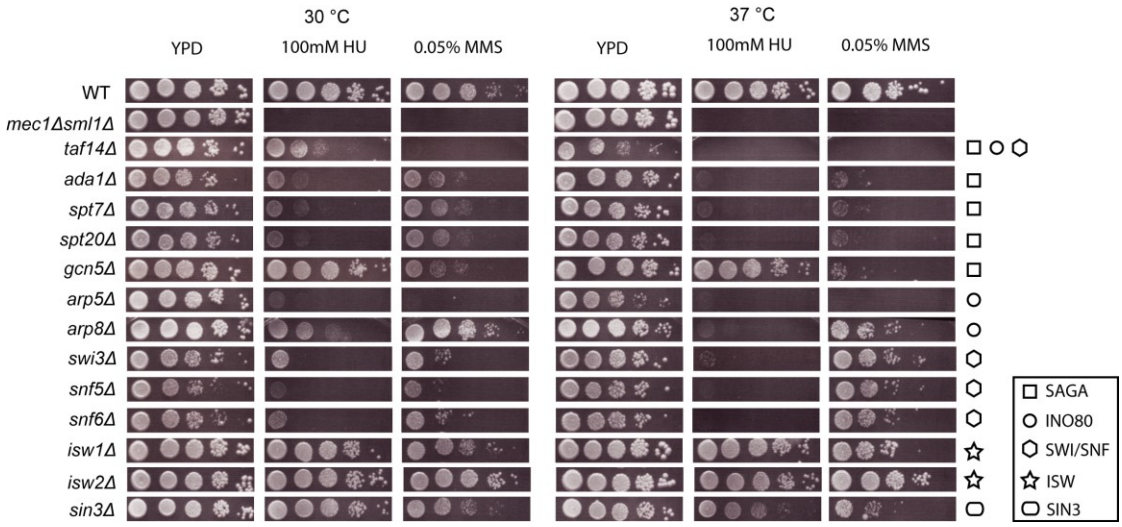

Figure 4.5 DNA damage sensitivities of a subset of null mutants of various subunits of chromatin remodeling complexes
A serial dilution of 5 fold was used to spot cells on the plates, which were then incubated at the indicated temperatures. (Sensitivity assays were performed by Sheng-Hong Chen.)

Chapter 4, is accepted and it will appear in J Biol Chem, 2010,Sheng-hong Chen, Claudio P. Albuquerque, Jason Liang, Raymond T. Suhandynata, and Huilin Zhou. The dissertation author was one of the primary researcher and author of this paper.

Chapter 5: Assigning the kinase substrate dependency for the DNA damage kinases, Mec1 and Tel1

5.1 Summary

Several substrates of Mec1 and Tel1, DNA damage checkpoint kinases, were indentified in Chapter 4. The majority of them were found to be chromatin associated. In this study, we used quantitative proteomics combined with sub-cellular fractionation to assign the dependency of Mec1 and Tel1 substrates in order to elucidate their overlapping and distinguishing roles in genome maintenance including: preventing gross chromosomal rearrangements, maintaining telomere length, and controlling DNA repair. Thirty four substrates were found to be either Mec1 or Tel1 dependent. Of those thirty four proteins, sixteen were Mec1 dependent and eighteen were Tel1 dependent. While the phosphorylation status of the majority of Mec/Tel1 substrates did not change in any single kinase deletion, strikingly, single kinase deletions of either abolishes the Mec1/Tel1 phosphorylation of Asg1, Cyc8, Hsp26, and Hsp12. Eight functional groups were identified in this screening, and specifically RNA transcription and histone modifiers were found to be primarily Tel1 dependent, while DNA damage checkpoint and cell cycle regulators were Mec1 dependent.

5.2 Introduction

Gross chromosomal rearrangements (GCRs) observed in *Saccharomyces cerevisiae* are similar to those observed in many tumor cells (Padilla-Nash et al., 1999). Genetic studies in yeast have revealed that many genes participate in the suppression of GCRs (Kolodner et al., 2002; Myung et al., 2001a; Myung et al., 2001b; Putnam et al., 2009). Among them, deletion of both MEC1 and TEL1, which encode two protein kinases in the DNA damage checkpoint (Nyberg et al., 2002), cause a 12,857-fold increase in GCRs, indicating a key function of their role in suppressing GCRs (Myung et al., 2001b). Remarkably, deletion of either MEC1 or TEL1 individually causes only a modest increase of GCRs; 194-fold for the *mec1*Δ mutant and no increase for the *tel1*Δ mutant. Thus, the function of these kinases in the suppression of GCRs appears to be highly redundant. Aside from their role in suppressing GCRs, Mec1 is known to have a major role in the activation of the Rad53 kinase; a response to DNA damage and replication stress (Paulovich and Hartwell, 1995; Sanchez et al., 1996; Sun et al., 1996). In contrast, Tel1 has a relatively limited role in Rad53 activation. Instead, Tel1 has a more important role, relative to Mec1, in regulating telomere length (Greenwell et al., 1995). Deletion of both MEC1 and TEL1 causes even shorter and unstable telomeres; leading to telomere fusions and additional chromosomal arrangements (Craven et al., 2002). Therefore, these kinases appear to have a redundant role in telomere maintenance. Importantly, the functions of Mec1

and Tel1 are not shared by their downstream kinases Rad53, Dun1 and Chk1; thus, these processes are likely attributed to the activities of Mec1 and Tel1 directly (Mallory and Petes, 2000; Myung et al., 2001b).

The role of Mec1 and Tel1 in genome maintenance appears to be evolutionarily conserved, as mutations to Rad3, the ortholog of Mec1 in fission yeast, lead to chromosomal fusion and circular chromosomes (Naito et al., 1998). ATR mutations in humans, the Mec1 homolog, have been implicated in the Seckel syndrome (O'Driscoll et al., 2003). Inherited mutation of the ATM kinase, the Tel1 homolog, is known to cause ataxia telangiectasia; a cancer prone syndrome (Savitsky et al., 1995). Taken together, this family of kinases has a critical and evolutionarily conserved role in preventing GCRs and telomere dysfunctions. Understanding their functions require knowledge of not only their *in vivo* substrates, but also how their substrates are recruited to the site of damage. A hallmark of the DNA damage response is the formation of foci by many DNA damage checkpoint signaling and repair proteins, including the early and temporal recruitment of Mec1 and Tel1 in yeast (Lisby et al., 2004). The Mec1-Ddc2 kinase complex interacts with Replication Protein-A (RPA), which binds to single stranded DNA (ssDNA) (Zou and Elledge, 2003). The Tel1 kinase interacts with the Mre11-Rad50-Xrs2 (MRX) complex (Falck et al., 2005; You et al., 2005), which recognizes DNA double stranded breaks (DSBs). It is in this manner that Mec1 and Tel1 are recruited to the site of DNA damage. A conserved mode of recruitment of ATR and ATM is also found in

higher eukaryotes (Falck et al., 2005). Following their recruitment to the site of DNA damage, Mec1 and Tel1 may use phosphorylation as a signal to orchestrate the assembly of other proteins, leading to a robust activation of the DNA damage checkpoint, efficient DNA repair, and proper coordination among different processes. In addition to the recruitment of Tel1, MRX also participates in the resection of DNA DSBs into ssDNA, leading to RPA binding and initiation of the homologous recombination (HR) pathway (Paull and Gellert, 1998; Trujillo et al., 1998). Additionally, MRX contributes to the nonhomologous end-joining (NHEJ) pathway via the Ku proteins, which bind to DNA DSBs (Zhang et al., 2007). Thus MRX may regulate the balance between the recruitment of Ku and RPA for different DNA repair pathways.

Interestingly, many Mec1 and Tel1 substrates are chromatin associated and are recruited to the site of DNA damage, including a conserved phosphorylation of histone subunit H2A (Downs et al., 2000). We previously identified chromatin remodelers as a major family of Mec1 and Tel1 substrates. Moreover, RPA was found to mediate the phosphorylation of Mec1 and Tel1 substrates, perhaps via its association to chromatin with DNA bridging the kinase-substrate interactions. These observations are consistent with the idea that chromatin may serve as a platform to mediate substrate recognition for Mec1 and Tel1. Understanding how Mec1 and Tel1 may control the phosphorylation of chromatin associated proteins, such as DNA damage signaling and repair proteins, is of considerable interest. In this study, we

applied quantitative mass spectrometry (MS) to quantify DNA damage induced phosphorylation of chromatin-associated proteins on a proteome-wide scale to determine the roles of Mec1 and Tel1. The results show that Mec1 and Tel1 have both highly redundant and distinct roles in regulating the phosphorylation of their substrates.

5.3 Experimental procedures

5.3.1 Western blot

50 ml of yeast cells where grown in light SILAC media until OD_{600nm} reaches 0.3. Then experimental cells were treated with 0.05% MMS for 2h while control cells were allowed to grow untreated for 2h. Cells were then harvested and lysed using the glass bead beating method previously described (Smolka et al., 2007), with the following modifications. The cell debris after high speed centrifugation was resuspended in 400ul of a buffer containing 50mM Tris pH8.0, 8 M Urea, and 10mM DTT shaken for 30 min at 37C, and then subjected to high speed centrifugation. The lysate and the urea extractions were then combined. Western blot analyses of proteins were detected by using anti-flag (Sigma) or anti H2A S129Q (abcam) according to the manufacture instruction.

5.3.2 Chromatin Extraction

1L of the WT or kinase-null cells were grown in SILAC media and treated by MMS as described in Chapter 4. Cell pellets were combined and chromatin associated proteins were enriched as previously described (Liang and Stillman, 1997), with the following modifications. After the Sucrose gradient the chromatin pellet was extracted twice by resuspendeding it in 4ml of 50mM Tris pH8.0, 8 M Urea, and 10mM DTT and shaken for 30 min at 37°C.

5.4 Results and Discussion

5.4.1 A quantitative assay for the roles of Mec1 and Tel1 in their substrate phosphorylation.

Since most Mec1 and Tel1 substrates are known to be chromatin-associated proteins, we sought to enrich them by purifying crude chromatin. The experimental strategy is shown in Figure 5.1A. WT and kinase-null cells, *mec1*Δ, *tel1*Δ and *mec1*Δ *tel1*Δ were first grown in either light or heavy Lys/Arg containing media and then treated by methyl methane sulfonate (MMS) to induce DNA damage (see Fig. 5.1A). To examine how MMS treatment may change their phosphorylation, we compare an MMS treated vs. untreated sample using a similar method (not shown) where the MMS treated cells were grown with heavy media. Cells were then combined and the chromatin was

purified. The purified chromatin typically contained about 10-times less protein when compared to whole cell lysate (see Fig. 5.1B). We next determined the relative abundance of heavy and light isotope labeled phosphopeptides of proteins in the crude chromatin using quantitative MS (Albuquerque et al., 2008). Most of the phosphorylated peptides identified in these experiments have an abundance ratio of 1 (results not shown).

Since we are interested in the phosphorylation of Mec1 and Tel1 substrates we examined the phosphorylation of S129 of H2A, which has the consensus, SQ/TQ, site of Mec1 and Tel1 and is a known target of these kinases. As shown in Figure 5.2C, S129 phosphorylation of H2A was identified by tandem MS analysis. The intensities of this phosphopeptide of H2A from untreated and MMS treated WT cells show a 3-fold increase in abundance with a signal to noise ratio over 10,000 as detected by MS. Similar increases in the abundance of phosphorylated H2A was also detected using quantitative Western blot analysis (see Fig. 5.1D). Next, we examined how this phosphopeptide of H2A is affected by the deletion of MEC1 and TEL1. Compared to WT cells, its intensity is virtually unaffected in the *mec1*Δ mutant, partially reduced in the *tel1*Δ mutant, and completely lost in the *mec1*Δ *tel1*Δ mutant, as shown by anti-phospho-H2A Western blot (see Fig. 5.1E) and quantitative MS results (see Fig. 5.1F). The quantitative MS results, based on stable isotope labeling, are generally more accurate than conventional Western blot analysis and provide information on the site of phosphorylation

(Ong et al., 2002). Thus, this quantitative proteomic method will illustrate the role of Mec1 and Tel1 in the phosphorylation of *mec1Δ/tel1Δ* substrates in a site-specific manner.

5.4.2 The role of Mec1 and Tel1 on the phosphorylation of their substrates

Next, we examine the phosphorylation of the checkpoint and sensor proteins, since their dependence on Mec1 & Tel1 are better understood. As expected, Mec1 preferentially phosphorylated Rfa1 and 2, subunits of RPA, and Rad53 (see figure 5.2). RPA is known to physically interact with Mec1, and it has been suggested to play a role in mediating how Mec1 targets its substrates. Additionally, the activation of the DNA damage checkpoint is primarily Mec1 dependent. One would expect Rad53 and Mrc1 phosphorylation to be Mec1 dependent, but Mrc1 phosphorylation was unaffected by the individual kinase deletions. However, the activity of Rad53 was found to be completely dependent on Mec1 as seen by its auto-phosphorylation (see figure 5.2) (Chen and Zhou, 2009). Tel1 on the other hand preferentially phosphorylates Rad50, a subunit of MRX. MRX is composed of three subunits Xrs2, Mre11, and Rad50. Tel1 physically interacts with MRX, and Xrs2 is required for the recruitment of Tel1 to a DSB (Nakada et al., 2003).

Next, we analyzed the phosphorylation of the chromatin remodeling complexes, a major group of protein complexes that were identified in our prior study. The Isw1, Isw2, SWI/SNF, RSC, and the Ino80 complex exhibit MMS inducible phosphorylation and are substrates of Mec1 and Tel1 (See Fig. 5.3). Isw1 and Isw2 were found to be Mec1 substrates, while the Ino80 and SWI/SNF were found to be Tel1 substrates. In the RSC complex, Rsc1 and 4 were Tel1 dependent substrates while Rsc2 was dependent on Mec1. Interestingly, the Mec1 dependent phosphorylation of Rsc2 increases in the absence of Tel1, suggesting that Tel1 has an inhibitory role in the Mec1 phosphorylation of Rsc2. The ISW1 complex can be subdivided into two complexes, a and b (Mellor and Morillon, 2004). Isw1b is a potential Mec1 substrate, since the phosphorylation of Ioc2 was found to be dependent on Mec1. Isw1b is involved in coordinating transcription elongation with termination. The phosphorylation of Isw1a, however, was found to be unaffected by either of the single kinase deletions. INO80 has a role in gene transcription regulation, DNA repair, and replication(Shimada et al., 2008). It has been shown that deleting the subunits of the INO80 causes yeast to be sensitive to DNA damaging agents, and that the complex is recruited to the site of damage. The functional roles of these two kinases in the phosphorylation of chromatin remodeling factors still remain unclear. However, the synergistic role that the chromatin remodeling complexes have with these

kinases, in preventing GCRs as well as regulating telomere maintenance, may validate the results of this screening.

5.4.1 Classification of the kinases substrates

Fifty-nine substrates of Mec1/Tel1 were identified in this screen and can be grouped into three sets. The first set of substrates is either Mec1 or Tel1 dependent, where sixteen were found to be Mec1 dependent and eighteen Tel1 dependent. The second set, are substrates that show the redundant role of these kinases, where the individual kinase deletions did not affect their phosphorylation states (see Fig. 5.4). Surprisingly, a set of substrates were dependent on both Mec1 and Tel1, as the single deletion of either kinase abolishes the phosphorylation of these substrates. It is unlikely that the phosphorylation of one kinase regulates the other. One can speculate that in this set, phosphorylation of a protein by one kinase is mediating the recruitment of an additional substrate to be phosphorylated by the other kinase. Mec1 and Tel1 substrates can be further grouped by function (see table 5.1), and systematic analysis of these substrates can help us understand their role in preventing GCRs, telomere maintenance, and DNA repair.

In summary, we have shown that these kinases have a set of distinguishing and overlapping substrates. Currently, we cannot assign any direct function to these phosphorylation sites since proper assays and phosphomutants of the identified substrates are required. More importantly,

only a few phosphorylations by the two upstream kinases have been show to be functionally relevant. It has been shown that mutation of the SQ/TQ, Mec1 and Tel1 consensus sites, on Rad9 and Rad53 affect the activation of the DNA Damage checkpoint.

5.5 Conclusion

The combination of HIILC with quantitative MS has enabled us to identify *in vivo* kinase substrates from whole cell lysate and can be applied to any signal transduction pathway. By combining these with a time-course experiment, the phosphorylation kinetics can be obtained *in vivo*, and the temporal recruitment of proteins to the site of action can be further understood. Since the kinase needs to be physically interacting with its substrate in order to phosphorylate it, this can indicate the recruitment of substrates to the damage site. For the future screenings of the DNA Damage checkpoint in yeast, the types of damage and the state of the cell cycle must be accounted for.

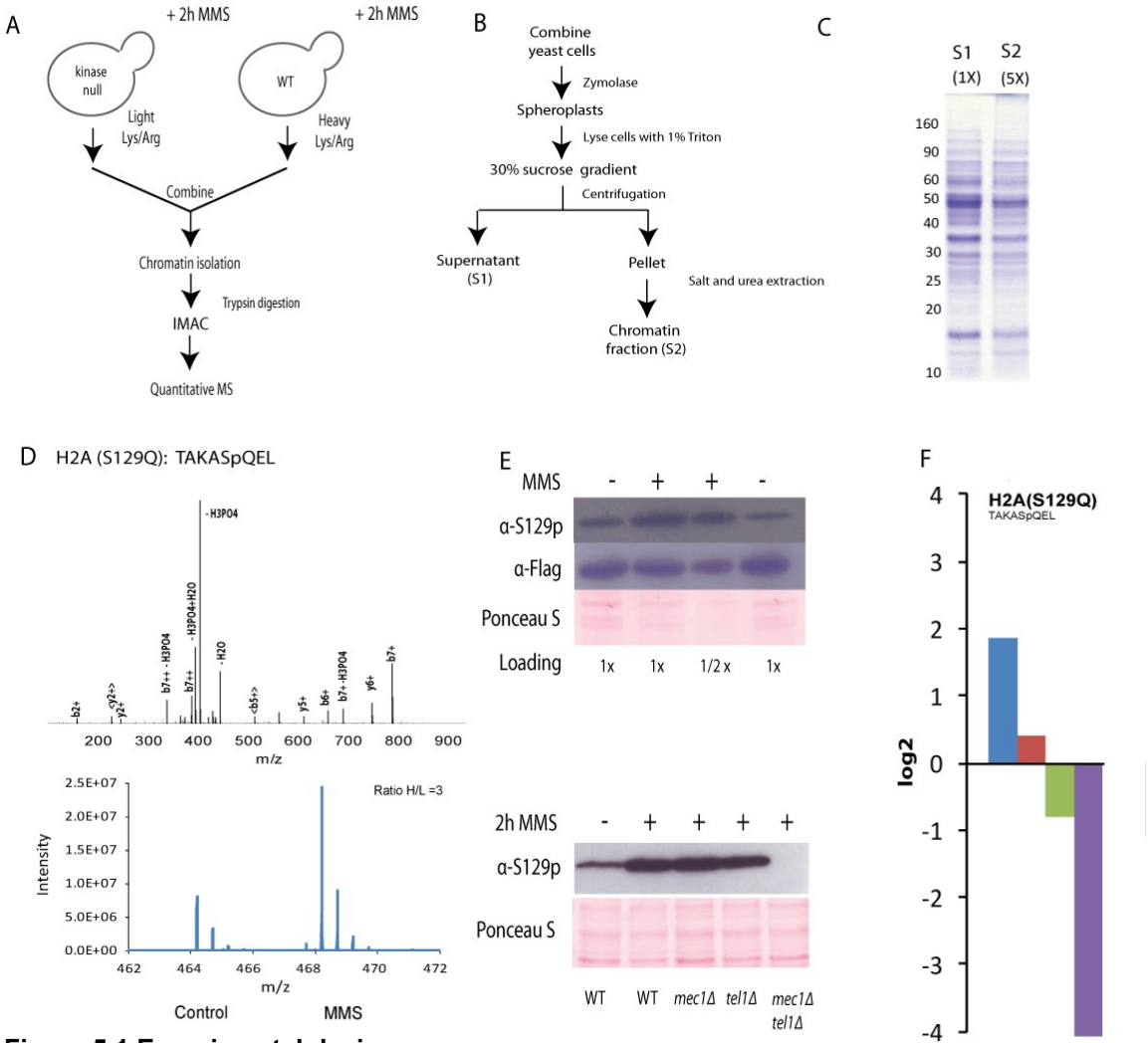

Figure 5.1 Experimental design

A) Experimental strategy to identify chromatin associated Mec1 and Tel1 substrates. B) Strategy for enrichment of chromatin associated proteins. S1 is the supernatant after the sucrose gradient, and S2 is the chromatin fraction. C) Coomassie staining of fraction S1 and S2 from the sub-cellular fractionation, with five times more loading of the S2 fraction relative to S1. D) MS data showing the identification and quantification of H2A S129 phophorylation after 2h MMS. The top panel shows the CID that identifies the TAKASpQEL peptide where "p" indicates the location of a phosphate. The bottom panel shows that the signal of TAKASpQEL is increased by 3 fold (intensity of the heavy/light peptide) after MMS treatment, and the signal to noise ratio is over 10,000 (The average intensity for the noise level of a full scan is around 1E4). E) The inducibility of the S129 phosphorylation of H2A by MMS and its dependency on *mec1*Δ, *tel1*Δ, and both *mec1*Δ*/tel1*Δ. F) Graphical representation of the MMS inducible S129 phosphorylation of H2A obtained by quantitative proteomics, and its dependency on *mec1*Δ, *tel1*Δ, and *mec1*Δ*/tel1*Δ. In blue, the integrated intensity ratio of MMS treated vs. untreated cells was plot using a log2 scale. In addition, the integrated intensity ratio of the peptide in kinase null vs. WT cells was plotted using a log2 scale. Red, green, and purple bars corresponding to *mec1*Δ, *tel1*Δ, and both *mec1*Δ*/tel1*Δ, respectively.

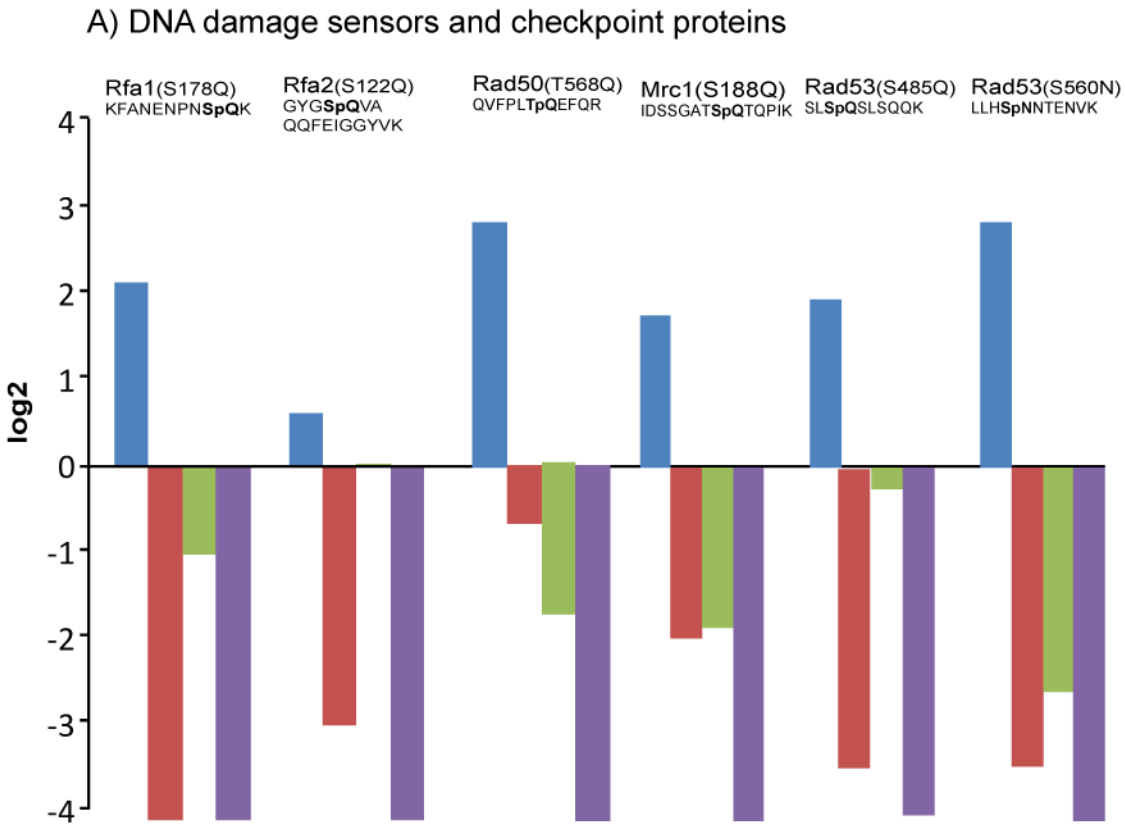

Figure 5.2 The role of Mec1 and Tel1 on the phosphorylation of DNA Damage Sensor and Checkpoint Proteins.
In blue, the integrated intensity ratio of MMS treated vs. untreated cells was plot using a log2 scale. In addition, the integrated intensity ratio of the peptide in kinase null vs. WT cells was plotted using a log2 scale. Red, green, and purple bars corresponding to *mec1*Δ, *tel1*Δ, and m*ec1*Δ*/tel1*Δ, respectively. Phosphopeptides with SpQ/TpQ phosphorylation, where"p" indicates the location of a phosphate, are Mec1/Tel1 dependent. The S560N phosphorylation on Rad53 is a known auto-phosphorylation site.

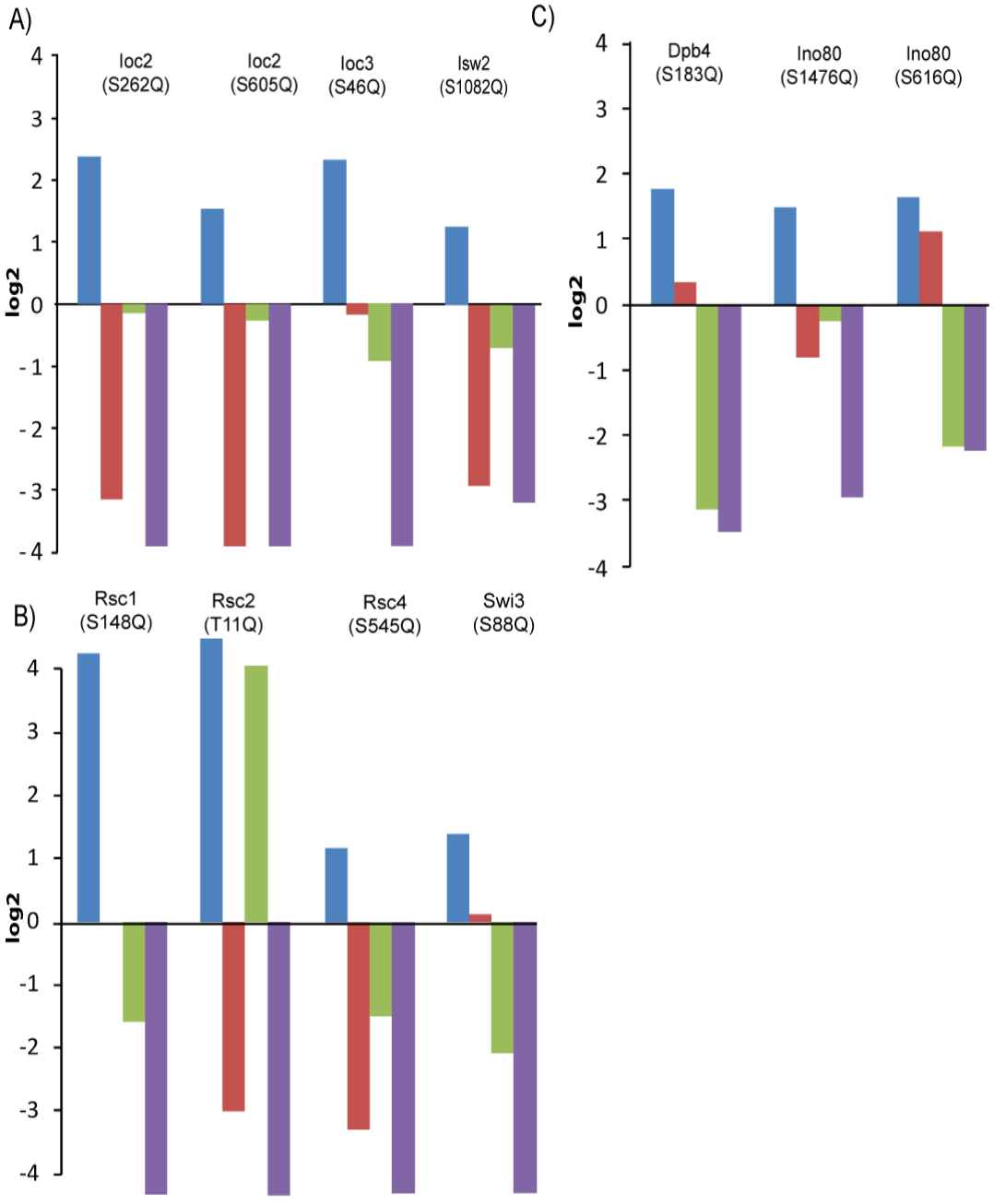

Figure 5.3 The role of Mec1 and Tel1 on the phosphorylation of chromatin remodeler complexes

In blue, the integrated intensity ratio of MMS treated vs. untreated cells was plotted using a log2 scale. In addition, the integrated intensity ratio of the peptide in kinase null vs. WT cells was also plotted using a log2 scale. Red, green, and purple bars corresponding *mec1*Δ, *tel1*Δ, and both *mec1*Δ/*tel1*Δ, respectively.

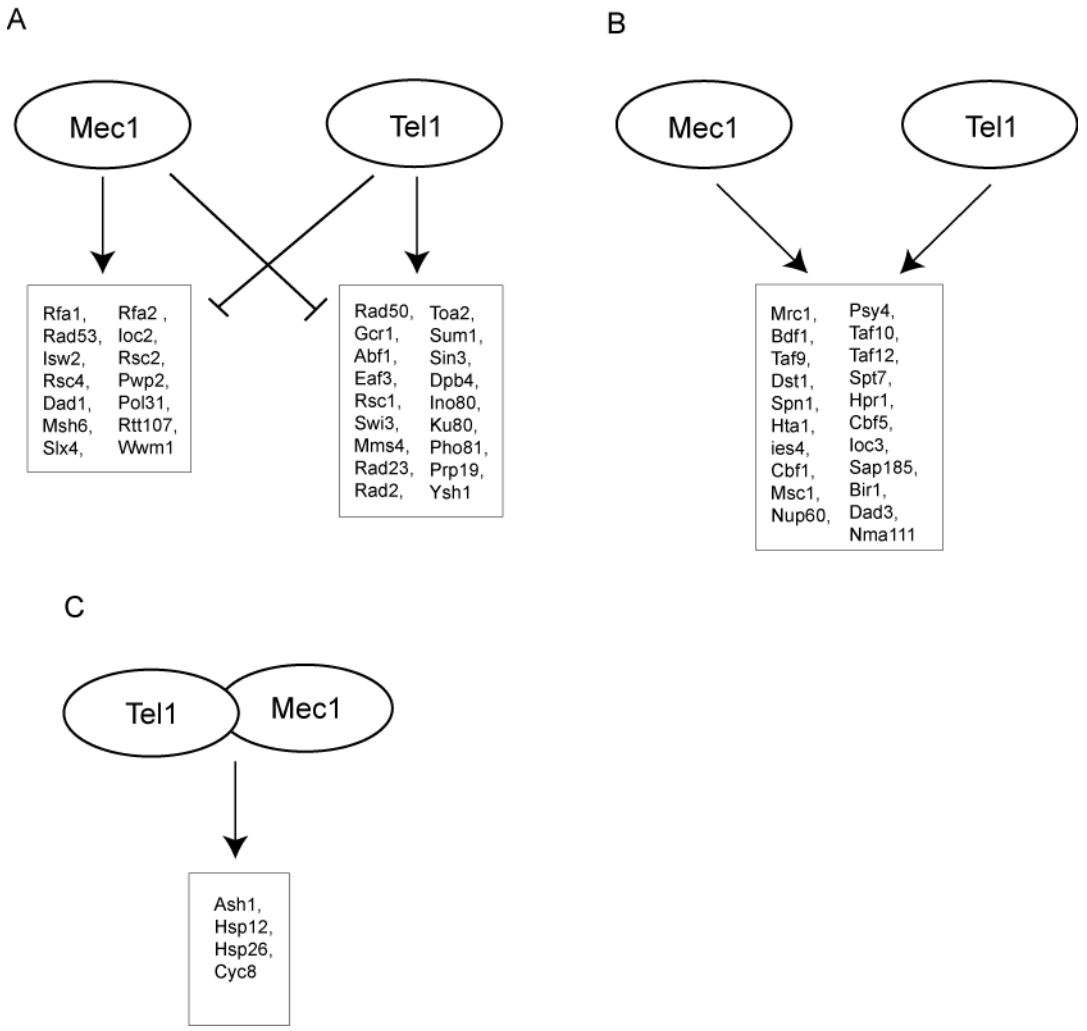

Figure 5.4 Substrate dependency of Mec1 and Tel1.
A) Mec1 or Tel1 dependent substrate B) Substrates that are not affected by the individual kinase deletion. C) Substrates that are dependent on both kinases.

Table 5.1 Functional Classification of Substrates of Mec1 or Tel1 dependent phosphorylation.

Functions	Mec1	Tel1	Dependant on both Mec1 and Tel1	Independent on the individual deletion
DNA Damage Sensors	Rfa1, Rfa2, Mec1	Rad50		
DNA Damage Checkpoint	Rad53			Mrc1,Psy4
RNA Transcription		Toa2, Gcr1, Abf1, Sum1	Cyc8	Bdf1, Taf9, Taf10, Taf12, Spt17, Hpr1, Dst1,Spn1
Histone Modification		Eaf3, Sin3		Hta1
Chromatin Remodeling	Ioc2, Isw2, Rsc2, Rsc4	Rsc1, Swi3, Dpb4, Ino80		Ioc3, Ies4
mRNA processing	Pwp2	Ysh1, Prp19		Cbf5
Cell Cycle Regulation	Dad1			Sap185, Bir1, Dad3,Cbf1,Mcd1
DNA Replication & Repair	Pol31, Msh6, Rtt107, Slx4, Top1	Mms4, Rad23, Rad2, Ku80, Pho81		
Others	Wwm1		Asg1, Hsp12, Hsp26	Nma111, Nup60

References

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198-207.

Albuquerque, C.P., Smolka, M.B., Payne, S.H., Bafna, V., Eng, J., and Zhou, H. (2008). A multidimensional chromatography technology for in-depth phosphoproteome analysis. Mol Cell Proteomics *7*, 1389-1396.

Alcasabas, A.A., Osborn, A.J., Bachant, J., Hu, F., Werler, P.J., Bousset, K., Furuya, K., Diffley, J.F., Carr, A.M., and Elledge, S.J. (2001). Mrc1 transduces signals of DNA replication stress to activate Rad53. Nat Cell Biol *3*, 958-965.

Bakalarski, C.E., Haas, W., Dephoure, N.E., and Gygi, S.P. (2007). The effects of mass accuracy, data acquisition speed, and search algorithm choice on peptide identification rates in phosphoproteomics. Anal Bioanal Chem *389*, 1409-1419.

Bashkirov, V.I., Bashkirova, E.V., Haghnazari, E., and Heyer, W.D. (2003). Direct kinase-to-kinase signaling mediated by the FHA phosphoprotein recognition domain of the Dun1 DNA damage checkpoint kinase. Mol Cell Biol *23*, 1441-1452.

Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C., and Gygi, S.P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A *101*, 12130-12135.

Beausoleil, S.A., Villen, J., Gerber, S.A., Rush, J., and Gygi, S.P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol *24*, 1285-1292.

Bodenmiller, B., Malmstrom, J., Gerrits, B., Campbell, D., Lam, H., Schmidt, A., Rinner, O., Mueller, L.N., Shannon, P.T., Pedrioli, P.G., *et al.* (2007a). PhosphoPep--a phosphoproteome resource for systems biology research in Drosophila Kc167 cells. Molecular systems biology *3*, 139.

Bodenmiller, B., Mueller, L.N., Mueller, M., Domon, B., and Aebersold, R. (2007b). Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat Methods *4*, 231-237.

Boersema, P.J., Divecha, N., Heck, A.J., and Mohammed, S. (2007). Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. J Proteome Res *6*, 937-946.

Bohm, S., Frishman, D., and Mewes, H.W. (1997). Variations of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. Nucleic Acids Res *25*, 2464-2469.

Brush, G.S., Morrow, D.M., Hieter, P., and Kelly, T.J. (1996). The ATM homologue MEC1 is required for phosphorylation of replication protein A in yeast. Proc Natl Acad Sci U S A *93*, 15075-15080.

Cartagena-Lirola, H., Guerini, I., Viscardi, V., Lucchini, G., and Longhese, M.P. (2006). Budding Yeast Sae2 is an In Vivo Target of the Mec1 and Tel1 Checkpoint Kinases During Meiosis. Cell Cycle *5*, 1549-1559.

Chai, B., Huang, J., Cairns, B.R., and Laurent, B.C. (2005). Distinct roles for the RSC and Swi/Snf ATP-dependent chromatin remodelers in DNA double-strand break repair. Genes Dev *19*, 1656-1661.

Chen, S., Smolka, M.B., and Zhou, H. (2007a). Mechanism of Dun1 activation by Rad53 phosphorylation in Saccharomyces cerevisiae. J Biol Chem *282*, 986-995.

Chen, S.H., Smolka, M.B., and Zhou, H. (2007b). Mechanism of Dun1 activation by Rad53 phosphorylation in Saccharomyces cerevisiae. J Biol Chem *282*, 986-995.

Chen, S.H., and Zhou, H. (2009). Reconstitution of Rad53 activation by Mec1 through adaptor protein Mrc1. J Biol Chem *284*, 18593-18604.

Chi, A., Huttenhower, C., Geer, L.Y., Coon, J.J., Syka, J.E., Bai, D.L., Shabanowitz, J., Burke, D.J., Troyanskaya, O.G., and Hunt, D.F. (2007).

Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. Proc Natl Acad Sci U S A *104*, 2193-2198.

Clapier, C.R., and Cairns, B.R. (2009). The Biology of Chromatin Remodeling Complexes. Annu Rev Biochem.

Cole, G.M., and Mortimer, R.K. (1989). Failure to induce a DNA repair gene, RAD54, in Saccharomyces cerevisiae does not affect DNA repair or recombination phenotypes. Mol Cell Biol *9*, 3314-3322.

Craven, R.J., Greenwell, P.W., Dominska, M., and Petes, T.D. (2002). Regulation of genome stability by TEL1 and MEC1, yeast homologs of the mammalian ATM and ATR genes. Genetics *161*, 493-507.

de Godoy, L.M., Olsen, J.V., de Souza, G.A., Li, G., Mortensen, P., and Mann, M. (2006). Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. Genome Biol *7*, R50.

de Lichtenberg, U., Wernersson, R., Jensen, T.S., Nielsen, H.B., Fausboll, A., Schmidt, P., Hansen, F.B., Knudsen, S., and Brunak, S. (2005). New weakly expressed cell cycle-regulated genes in yeast. Yeast *22*, 1191-1201.

Desany, B.A., Alcasabas, A.A., Bachant, J.B., and Elledge, S.J. (1998). Recovery from DNA replicational stress is the essential function of the S-phase checkpoint pathway. Genes Dev *12*, 2956-2970.

Downs, J.A., Lowndes, N.F., and Jackson, S.P. (2000). A role for Saccharomyces cerevisiae histone H2A in DNA repair. Nature *408*, 1001-1004.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods *4*, 207-214.

Elledge, S.J., and Davis, R.W. (1989). DNA damage induction of ribonucleotide reductase. Mol Cell Biol *9*, 4932-4940.

Emili, A. (1998). MEC1-dependent phosphorylation of Rad9p in response to DNA damage. Mol Cell *2*, 183-189.

Falck, J., Coates, J., and Jackson, S.P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. Nature *434*, 605-611.

Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol *20*, 301-305.

Flott, S., and Rouse, J. (2005). Slx4 becomes phosphorylated after DNA damage in a Mec1/Tel1-dependent manner and is required for repair of DNA alkylation damage. Biochem J *391*, 325-333.

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.

Gilar, M., Olivova, P., Daly, A.E., and Gebler, J.C. (2005). Orthogonality of separation in two-dimensional liquid chromatography. Anal Chem *77*, 6426-6434.

Graumann, J., Hubner, N.C., Kim, J.B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008). Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. Mol Cell Proteomics *7*, 672-683.

Greenwell, P.W., Kronmal, S.L., Porter, S.E., Gassenhuber, J., Obermaier, B., and Petes, T.D. (1995). TEL1, a gene involved in controlling telomere length in S. cerevisiae, is homologous to the human ataxia telangiectasia gene. Cell *82*, 823-829.

Gruhler, A., Olsen, J.V., Mohammed, S., Mortensen, P., Faergeman, N.J., Mann, M., and Jensen, O.N. (2005). Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol Cell Proteomics *4*, 310-327.

Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc Natl Acad Sci U S A *97*, 9390-9395.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol *17*, 994-999.

Han, D.K., Eng, J., Zhou, H., and Aebersold, R. (2001). Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol *19*, 946-951.

Harrison, J.C., and Haber, J.E. (2006). Surviving the breakup: the DNA damage checkpoint. Annu Rev Genet *40*, 209-235.

Hemrika, W., and Berden, J.A. (1990). Membrane topography of the subunits of ubiquinol-cytochrome-c oxidoreductase of Saccharomyces cerevisiae. The 14-kDa and the 11-kDa subunits face opposite sides of the mitochondrial inner membrane. Eur J Biochem *192*, 761-765.

Huang, M., Zhou, Z., and Elledge, S.J. (1998). The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. Cell *94*, 595-605.

Hunt, D.F., Henderson, R.A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A.L., Appella, E., and Engelhard, V.H. (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science *255*, 1261-1263.

Inoue, Y., Matsuda, T., Sugiyama, K., Izawa, S., and Kimura, A. (1999). Genetic analysis of glutathione peroxidase in oxidative stress response of Saccharomyces cerevisiae. J Biol Chem *274*, 27002-27009.

Jelinsky, S.A., and Samson, L.D. (1999). Global response of Saccharomyces cerevisiae to an alkylating agent. Proc Natl Acad Sci U S A *96*, 1486-1491.

Kim, S.T., Lim, D.S., Canman, C.E., and Kastan, M.B. (1999). Substrate specificities and identification of putative substrates of ATM kinase family members. J Biol Chem *274*, 37538-37543.

Kolodner, R.D., Putnam, C.D., and Myung, K. (2002). Maintenance of genome stability in Saccharomyces cerevisiae. Science *297*, 552-557.

Kruger, M., Moser, M., Ussar, S., Thievessen, I., Luber, C.A., Forner, F., Schmidt, S., Zanivan, S., Fassler, R., and Mann, M. (2008). SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell *134*, 353-364.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics *4*, 873-886.

Lee, M.W., Kim, B.J., Choi, H.K., Ryu, M.J., Kim, S.B., Kang, K.M., Cho, E.J., Youn, H.D., Huh, W.K., and Kim, S.T. (2007). Global protein expression profiling of budding yeast in response to DNA damage. Yeast *24*, 145-154.

Lee, S.J., Schwartz, M.F., Duong, J.K., and Stern, D.F. (2003). Rad53 phosphorylation site clusters are important for Rad53 regulation and signaling. Mol Cell Biol *23*, 6300-6314.

Li, X., Gerber, S.A., Rudner, A.D., Beausoleil, S.A., Haas, W., Villen, J., Elias, J.E., and Gygi, S.P. (2007). Large-scale phosphorylation analysis of alpha-factor-arrested Saccharomyces cerevisiae. J Proteome Res *6*, 1190-1197.

Liang, B., Qiu, J., Ratnakumar, K., and Laurent, B.C. (2007). RSC functions as an early double-strand-break sensor in the cell's response to DNA damage. Curr Biol *17*, 1432-1437.

Liang, C., and Stillman, B. (1997). Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants. Genes Dev *11*, 3375-3386.

Lisby, M., Barlow, J.H., Burgess, R.C., and Rothstein, R. (2004). Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. Cell *118*, 699-713.

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C.S., and Foiani, M. (2001). The DNA replication checkpoint response stabilizes stalled replication forks. Nature *412*, 557-561.

Mallory, J.C., and Petes, T.D. (2000). Protein kinase activity of Tel1p and Mec1p, two Saccharomyces cerevisiae proteins related to the human ATM protein kinase. Proc Natl Acad Sci U S A *97*, 13749-13754.

Masson, J.Y., and Ramotar, D. (1996). The Saccharomyces cerevisiae IMP2 gene encodes a transcriptional activator that mediates protection against DNA damage caused by bleomycin and other oxidants. Mol Cell Biol *16*, 2091-2100.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y.*, et al.* (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science *316*, 1160-1166.

McNulty, D.E., and Annan, R.S. (2008). Hydrophilic-interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. Mol Cell Proteomics, in press.

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

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature *455*, 770-774.

Morin, I., Ngo, H.P., Greenall, A., Zubko, M.K., Morrice, N., and Lydall, D. (2008). Checkpoint-dependent phosphorylation of Exo1 modulates the DNA damage response. EMBO J *27*, 2400-2410.

Morrison, A.J., Highland, J., Krogan, N.J., Arbel-Eden, A., Greenblatt, J.F., Haber, J.E., and Shen, X. (2004). INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. Cell *119*, 767-775.

Morrison, A.J., Kim, J.A., Person, M.D., Highland, J., Xiao, J., Wehr, T.S., Hensley, S., Bao, Y., Shen, J., Collins, S.R.*, et al.* (2007). Mec1/Tel1 phosphorylation of the INO80 chromatin remodeling complex influences DNA damage checkpoint responses. Cell *130*, 499-511.

Morrison, A.J., and Shen, X. (2009). Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. Nat Rev Mol Cell Biol.

Myung, K., Chen, C., and Kolodner, R.D. (2001a). Multiple pathways cooperate in the suppression of genome instability in Saccharomyces cerevisiae. Nature *411*, 1073-1076.

Myung, K., Datta, A., and Kolodner, R.D. (2001b). Suppression of spontaneous chromosomal rearrangements by S phase checkpoint functions in Saccharomyces cerevisiae. Cell *104*, 397-408.

Naito, T., Matsuura, A., and Ishikawa, F. (1998). Circular chromosome formation in a fission yeast mutant defective in two ATM homologues. Nat Genet *20*, 203-206.

Nakada, D., Hirano, Y., Tanaka, Y., and Sugimoto, K. (2005). Role of the C terminus of Mec1 checkpoint kinase in its localization to sites of DNA damage. Mol Biol Cell *16*, 5227-5235.

Nakada, D., Matsumoto, K., and Sugimoto, K. (2003). ATM-related Tel1 associates with double-strand breaks through an Xrs2-dependent mechanism. Genes Dev *17*, 1957-1962.

Nyberg, K.A., Michelson, R.J., Putnam, C.W., and Weinert, T.A. (2002). Toward maintaining the genome: DNA damage and replication checkpoints. Annu Rev Genet *36*, 617-656.

O'Driscoll, M., Ruiz-Perez, V.L., Woods, C.G., Jeggo, P.A., and Goodship, J.A. (2003). A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet *33*, 497-501.

Oda, Y., Huang, K., Cross, F.R., Cowburn, D., and Chait, B.T. (1999). Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci U S A *96*, 6591-6596.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635-648.

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.

Osborn, A.J., and Elledge, S.J. (2003). Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. Genes Dev *17*, 1755-1767.

Padilla-Nash, H.M., Nash, W.G., Padilla, G.M., Roberson, K.M., Robertson, C.N., Macville, M., Schrock, E., and Ried, T. (1999). Molecular cytogenetic analysis of the bladder carcinoma cell line BK-10 by spectral karyotyping. Genes Chromosomes Cancer *25*, 53-59.

Paull, T.T., and Gellert, M. (1998). The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. Mol Cell *1*, 969-979.

Paulovich, A.G., and Hartwell, L.H. (1995). A checkpoint regulates the rate of progression through S phase in S. cerevisiae in response to DNA damage. Cell *82*, 841-847.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res *2*, 43-50.

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.

Putnam, C.D., Hayes, T.K., and Kolodner, R.D. (2009). Specific pathways prevent duplication-mediated genome rearrangements. Nature *460*, 984-989.

Roberts, T.M., Kobor, M.S., Bastin-Shanower, S.A., Ii, M., Horte, S.A., Gin, J.W., Emili, A., Rine, J., Brill, S.J., and Brown, G.W. (2006). Slx4 regulates DNA damage checkpoint-dependent phosphorylation of the BRCT domain protein Rtt107/Esc4. Mol Biol Cell *17*, 539-548.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S.*, et al.* (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics *3*, 1154-1169.

Sanchez, Y., Bachant, J., Wang, H., Hu, F., Liu, D., Tetzlaff, M., and Elledge, S.J. (1999). Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms. Science *286*, 1166-1171.

Sanchez, Y., Desany, B.A., Jones, W.J., Liu, Q., Wang, B., and Elledge, S.J. (1996). Regulation of RAD53 by the ATM-like kinases MEC1 and TEL1 in yeast cell cycle checkpoint pathways. Science *271*, 357-360.

Savitsky, K., Bar-Shira, A., Gilad, S., Rotman, G., Ziv, Y., Vanagaite, L., Tagle, D.A., Smith, S., Uziel, T., Sfez, S.*, et al.* (1995). A single ataxia telangiectasia gene with a product similar to PI-3 kinase. Science *268*, 1749-1753.

Segurado, M., and Diffley, J.F. (2008). Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks. Genes Dev *22*, 1816-1827.

Shim, E.Y., Ma, J.L., Oum, J.H., Yanez, Y., and Lee, S.E. (2005). The yeast chromatin remodeler RSC complex facilitates end joining repair of DNA double-strand breaks. Mol Cell Biol *25*, 3934-3944.

Shimada, K., Oma, Y., Schleker, T., Kugou, K., Ohta, K., Harata, M., and Gasser, S.M. (2008). Ino80 chromatin remodeling complex promotes recovery of stalled replication forks. Curr Biol *18*, 566-575.

Smolka, M.B., Albuquerque, C.P., Chen, S.H., Schmidt, K.H., Wei, X.X., Kolodner, R.D., and Zhou, H. (2005). Dynamic Changes in Protein-Protein Interaction and Protein Phosphorylation Probed with Amine-reactive Isotope Tag. Mol Cell Proteomics *4*, 1358-1369.

Smolka, M.B., Albuquerque, C.P., Chen, S.H., and Zhou, H. (2007). Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. Proc Natl Acad Sci U S A *104*, 10364-10369.

Smolka, M.B., Zhou, H., Purkayastha, S., and Aebersold, R. (2001). Optimization of the isotope-coded affinity tag-labeling procedure for quantitative proteome analysis. Anal Biochem *297*, 25-31.

Stage-Zimmermann, T., Schmidt, U., and Silver, P.A. (2000). Factors affecting nuclear export of the 60S ribosomal subunit in vivo. Mol Biol Cell *11*, 3777-3789.

Stokes, M.P., Rush, J., Macneill, J., Ren, J.M., Sprott, K., Nardone, J., Yang, V., Beausoleil, S.A., Gygi, S.P., Livingstone, M.*, et al.* (2007). Profiling of UV-induced ATM/ATR signaling pathways. Proc Natl Acad Sci U S A *104*, 19855-19860.

Sun, Z., Fay, D.S., Marini, F., Foiani, M., and Stern, D.F. (1996). Spk1/Rad53 is regulated by Mec1-dependent protein phosphorylation in DNA replication and damage checkpoint pathways. Genes Dev *10*, 395-406.

Tanaka, K., and Russell, P. (2001). Mrc1 channels the DNA replication arrest signal to checkpoint kinase Cds1. Nat Cell Biol *3*, 966-972.

Tanner, S., Shu, H., Frank, A., Wang, L.C., Zandi, E., Mumby, M., Pevzner, P.A., and Bafna, V. (2005). InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. Anal Chem *77*, 4626-4639.

Tao, W.A., Wollscheid, B., O'Brien, R., Eng, J.K., Li, X.J., Bodenmiller, B., Watts, J.D., Hood, L., and Aebersold, R. (2005). Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. Nat Methods *2*, 591-598.

Trujillo, K.M., Yuan, S.S., Lee, E.Y., and Sung, P. (1998). Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. J Biol Chem *273*, 21447-21450.

Tseng, S.F., Lin, J.J., and Teng, S.C. (2006). The telomerase-recruitment domain of the telomere binding protein Cdc13 is regulated by Mec1p/Tel1p-dependent phosphorylation. Nucleic Acids Res *34*, 6327-6336.

Usui, T., Ogawa, H., and Petrini, J.H. (2001). A DNA damage response pathway controlled by Tel1 and the Mre11 complex. Mol Cell *7*, 1255-1266.

Venable, J.D., Wohlschlegel, J., McClatchy, D.B., Park, S.K., and Yates, J.R., 3rd (2007). Relative quantification of stable isotope labeled peptides using a linear ion trap-Orbitrap hybrid mass spectrometer. Anal Chem *79*, 3056-3064.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol *19*, 242-247.

Wei, J., Sun, J., Yu, W., Jones, A., Oeller, P., Keller, M., Woodnutt, G., and Short, J.M. (2005). Global proteome discovery using an online three-dimensional LC-MS/MS. J Proteome Res *4*, 801-808.

Yates, J.R., 3rd, Eng, J.K., McCormack, A.L., and Schieltz, D. (1995). Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal Chem *67*, 1426-1436.

Yoshida, T. (2004). Peptide separation by Hydrophilic-Interaction Chromatography: a review. J Biochem Biophys Methods *60*, 265-280.

You, Z., Chahwan, C., Bailis, J., Hunter, T., and Russell, P. (2005). ATM activation and its recruitment to damaged DNA require binding to the C terminus of Nbs1. Mol Cell Biol *25*, 5363-5379.

Zhai, B., Villen, J., Beausoleil, S.A., Mintseris, J., and Gygi, S.P. (2008). Phosphoproteome analysis of Drosophila melanogaster embryos. J Proteome Res *7*, 1675-1682.

Zhang, Y., Hefferin, M.L., Chen, L., Shim, E.Y., Tseng, H.M., Kwon, Y., Sung, P., Lee, S.E., and Tomkinson, A.E. (2007). Role of Dnl4-Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination. Nat Struct Mol Biol *14*, 639-646.

Zhao, X., Muller, E.G., and Rothstein, R. (1998). A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Mol Cell *2*, 329-340.

Zhao, X., and Rothstein, R. (2002). The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. Proc Natl Acad Sci U S A *99*, 3746-3751.

Zhou, Z., and Elledge, S.J. (1993). DUN1 encodes a protein kinase that controls the DNA damage response in yeast. Cell *75*, 1119-1127.

Zhu, Y., and Xiao, W. (2001). Two alternative cell cycle checkpoint pathways differentially control DNA damage-dependent induction of MAG1 and DDI1 expression in yeast. Mol Genet Genomics *266*, 436-444.

Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell *134*, 981-994.

Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542-1548.