Mass spectrometry as a tool to dissect the role of chromatin assembly factors in regulating nucleosome assembly

par
Marlène Gharib

Département de Biochimie, Université de Montréal
Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de *Philosophiæ Doctor* (Ph.D.)
en biochimie

Décembre, 2010

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée:

Mass spectrometry as a tool to dissect the role of chromatin assembly factors in regulating nucleosome assembly

Présentée par :
Marlène Gharib

a été évaluée par un jury composé des personnes suivantes :

Dr. Benoit Coulombe, président-rapporteur
Dr. Pierre Thibault, directeur de recherche
Dr. Alain Verreault, co-directeur
Dr. François Robert, membre du jury
Dr. Momtchil Vodenitcharov, examinateur externe
Dr.Éric Milot, Représentant du doyen

Résumé

L'assemblage des nucléosomes est étroitement couplée à la synthèse des histones ainsi qu'à la réplication et la réparation de l'ADN durant la phase S. Ce processus implique un mécanisme de contrôle qui contribue soigneusement et de manière régulée à l'assemblage de l'ADN en chromatine. L'assemblage des nucléosomes durant la synthèse de l'ADN est crucial et contribue ainsi au maintien de la stabilité génomique. Cette thèse décrit la caractérisation par spectrométrie de masse (SM) des protéines jouant un rôle critique dans l'assemblage et le maintien de la structure chromatinienne. Plus précisément, la phosphorylation de deux facteurs d'assemblage des nucléosome, le facteur CAF-1, une chaperone d'histone qui participe à l'assemblage de la chromatine spécifiquement couplée à la réplication de l'ADN, ainsi que le complexe protéique Hir, jouant de plus un rôle important dans la régulation transcriptionelle des gènes d'histones lors de la progression normale du cycle cellulaire et en réponse aux dommages de l'ADN, a été examiné.

La caractérisation des sites de phosphorylation par SM nécéssite la séparation des protéines par éléctrophorèse suivi d'une coloration a l'argent. Dans le chapitre 2, nous demontrons que la coloration à l'argent induit un artéfact de sulfatation. Plus précisément, cet artéfact est causé par un réactif spécifiquement utilisé lors de la coloration. La sulfatation présente de fortes similitudes avec la phosphorylation. Ainsi, l'incrément de masse observé sur les peptides sulfatés et phosphorylés (+80 Da) nécéssite des instruments offrant une haute résolution et haute précision de masse pour différencier ces deux modifications.

Dans les chapitres 3 et 4, nous avons d'abord démontré par SM que Cac1, la plus grande sous-unité du facteur CAF-1, est cible de plusieurs sites de phosphorylation. Fait intéréssant, certains de ces sites contiennent des séquences consensus pour les kinases Cdc7-Dbf4 et CDKs. Ainsi, ces résultats fournissent les premières évidences que CAF-1 est potentiellement régulé par ces deux kinases *in vivo*. La fonction de tous les sites de

phosphorylation identifiés a ensuite été évaluée. Nous avons démontré que la phosphorylation de la Ser-503, un site consensus de la DDK, est essentielle à la répréssion transcriptionelle des gènes au niveau des télomères. Cependant, cette phosphorylation ne semble pas être nécéssaire pour d'autres fonctions connues de CAF-1, indiquant que le blocage de la phsophorylation de Cac1 Ser-503 affecte spécifiquement la fonction de CAF-1 aux structures hétérochromatiques des télomères. Ensuite, nous avons identifiés une intéraction physique entre CAF-1 et Cdc7-Dbf4. Des études *in vitro* ont également demontré que cette kinase phosphoryle spécifiquement Cac1 Ser-503, suggérant un rôle potential pour la kinase Cdc7-Dbf4 dans l'assemblage et la stabilité de la structure hétérochromatique aux télomères. Finalement, les analyses par SM nous ont également permi de montrer que la sous-unité Hpc2 du complexe Hir est phosphorylée sur plusieurs sites consensus des CDKs et de Cdc7-Dbf4. De plus, la quantification par SM d'un site spécifique de phosphorylation de Hpc2, la Ser-330, s'est révélée être fortement induite suite à l'activation du point de contrôle de réplication (le "checkpoint") suite au dommage a l'ADN. Nous montrons que la Ser-330 de Hpc2 est phopshorylée par les kinases de point de contrôle de manière Mec1/Tel1- et Rad53-dépendante. Nos données préliminaires suggèrent ainsi que la capacité du complex Hir de réguler la répréssion transcriptionelle des gènes d'histones lors de la progression du cycle cellulaire normal et en réponse au dommage de l'ADN est médiée par la phosphorylation de Hpc2 par ces deux kinases.

Enfin, ces deux études mettent en évidence l'importance de la spectrométrie de masse dans la caractérisation des sites de phosphorylation des protéines, nous permettant ainsi de comprendre plus précisement les mécanismes de régulation de l'assemblage de la chromatine et de la synthèse des histones.

Mots-clés: CAF-1, protéines Hir, spectrométrie de masse, coloration à l'argent, sulfatation, phosphorylation, assemblage de la chromatine, réplication de l'ADN, répréssion télomérique, régulation des gènes d'histones, dommage à l'ADN.

Abstract

Nucleosome assembly entails a controlled mechanism that is tightly coupled to DNA and histone synthesis during DNA replication and repair in S-phase. Importantly, this contributes to the prompt and carefully orchestrated assembly of newly replicated DNA into chromatin, which is essential for the maintenance of genomic integrity. This thesis describes the mass spectrometric characterization of proteins critical in the regulation of nucleosome assembly behind the replication fork and chromatin structure. More specifically, the phosphorylation of Chromatin Assembly Factor 1 (CAF-1), a nucleosome assembly factor that uniquely functions during replication-coupled *de novo* nucleosome assembly in S-phase and the Hir protein complex, a second nucleosome assembly factor that also contributes to the transcriptional regulation of histone genes during normal cell cycle progression and in response to DNA damage, was examined.

We first demonstrated that characterization of protein phosphorylation by mass spectrometry (MS), which often relies on the separation of proteins by gel electrophoresis followed by silver staining for visualization, should be given careful considerations. In chapter 2, we report a potential pitfall in the interpretation of phosphorylation modifications due to the artifactual sulfation of serine, threonine and tyrosine residues caused by a specific reagent used during silver staining. Sulfation and phosphorylation both impart an 80 Da addition of these residues making them distinguishable only with MS systems offering high resolution and high mass accuracy capabilities.

Chapter 3 and 4 present the MS characterization of *in vivo* phosphorylation occurring on CAF-1 and Hir proteins, respectively. We first demonstrated that Cac1, the largest subunit of CAF-1, is phosphorylated on several novel residues containing the consensus sequences recognized by either Cdc7-Dbf4 (DDK) or cyclin-dependent kinases (CDKs). These results have provided the first evidence that CAF-1 is regulated by these two kinases *in vivo*. The function of all identified Cac1 phosphorylation sites was then assessed. *In vivo* phenotypic studies showed that the specific phosphorylation of Ser-503, a

Cac1 DDK-like site identified in our study, is essential for heterochromatin-mediated telomeric silencing. Cac1-Ser-503 did not appear to be required for other known functions of CAF-1, including DNA damage resistance and mitotic chromosome segregation, indicating that blocking Cac1 phosphorylation on Ser-503 sepcifically cripples CAF-1 function at telomeres. Next, biochemical purifications identified a physical interaction between CAF-1 and Cdc7-Dbf4. Consistent with this physical interaction data, *in vitro* kinase assay studies showed that Cdc7-Dbf4 specifically phosphorylates Cac1 Ser-503 thereby uncovering a novel role for Cdc7-Dbf4 in heterochromatin assembly and/or stability that is potentially mediated through CAF-1. Finally, MS analysis also showed that the Hpc2 subunit of the Hir protein complex is phosphorylated on several CDK- and DDK-like consensus sites. Furthermore, MS quantification of a specific phosphorylation site, Hpc2 Ser-330, was shown to be highly induced following the activation of the DNA damage checkpoint in response to DNA damage. We show that Hpc2 Ser-330 is phopshorylated by checkpoint kinases in a Mec1/Tel1- and Rad53-dependent manner. Our preliminary data suggest that the ability of the Hir protein complex to regulate the transcriptional repression of histone genes during normal cell cycle progression and in response to DNA damage is mediated through the regulated phosphorylation of Hpc2 by these kinases.

Finally, these two studies highlight the importance of mass spectrometry in characterizing protein phosphorylation events, which has yielded novel insights into the regulation of chromatin assembly by CAF-1 and histone synthesis mediated by Hir proteins.

Keywords: CAF-1, Hir proteins, mass spectrometry, silver staining, sulfation, phosphorylation, chromatin assembly, DNA replication, telomeric silencing, histone gene regulation, DNA damage.

Table of contents

List of tables

Liste of figures

List of abbreviations

1D	one dimensional
2D	two dimensional
3`hExo	3`exoribonuclease
5-FOA	5-fluoroorotic acid
6-4 PP	6-4 photoproduct
ABF1	autonomously replicating sequence binding factor 1
AC	alternating current
ACN	acetonitrile
ACS	ARS consensus sequence
Ade	adenine
ADP	adenosine diphosphate
Ag_2S	silver sulfide
$AgNO_3$	silver nitrate
AP	apyrimidinic/apurinic
ARS	autonomously replicating sequence
Asf1	Anti-silencing factor 1
ATM	ataxia telangiectasia mutated
ATP	adenosine 5`-triphosphate
ATR	ataxia telangiectasa rad3-related
$Ba(OH)_2$	barium hydroxide
BER	base excision repair
bp	base pair
Cac	chromatin assembly complex
$CaCl_2$	calcium chloride
CAF-1	chromatin assembly factor 1
CBB	coomassie brilliant blue
CBP	calmodulin binding peptide
CCR5	C-C chemokine receptor 5
Cdc	cell division cycle
CDK	cyclin dependent kinase
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
Chk1/Chk2	checkpoint kinase 1/2
CID	collision induced dissociation
Clr4	cryptic loci regulator 4
CPD	cyclobutane pyrimidine dimers

Da	daltons
Dbf4	dumbbell former 4
DDA	data-dependent acquisition
DDK	Dbf4-dependent kinase
DHB	2,5-dihydroxybenzoic acid
DNA	deoxyribonucleic acid
DNMT1	DNA (5-cytosine) methyltrasnferase 1
DSB	double strand break
DTT	dithiothreitol
ECD	electron capture dissociation
EDT	ethanedithiol
EDTA	ethylene diamine tetra-acetic acid
EMS	enhanced MS
EPI	enhanced product ion
ERK1	extracellular signal-related kinase 1
ESI	electrospray ionization
ETD	electron transfer dissociation
FA	formic acid
FAB	fast atom bombardment
FACT	facilitates chromatin transcription
FEN-1	flap endonuclease 1
FT-ICR	fourier transform ion cyclotron
GCN5	general control non-derepressible 5
GCR	gross chromosomal rearrangement
GST	gluthatione-S-transferase
H_3PO_4	phosphoric acid
HAT	histone acetyltransferase
HDAC	histone deacetylase
HEPES	N-(2-hydroxyethyl) piperazine-N`-(2-ethanosulfonic acid)
HIR	histone gene regulation
His	histidine
HML	homothallic mating type locus L
HMR	homothallic mating type locus R
HP1	heterochromatin protein 1
Hpc2	histone promoter control 2
HPLC	high performance liquid chromatography
HPO_3	phosphate
HU	hydroxyurea

IDA	iminodiacetate
IDA	information-dependent acquisition
IDCL	interdomain connecting loop
IDL	insertion/deletion
IEF	isoelectric focusing
IgG	immunoglobulin G
IMAC	immobilized metal affinity chromatography
IP	immunoprecipitation
IPTG	isopropyl-β-D-thiogalactopyranoside
$K_3F_2(CN)_6$	potassium hexacyanoferrate
Kb	kilobase
KCl	Potassium chloride
KOH	potassium hydroxide
LB	lysogeny broth
LC	liquid chromatography
Leu	leucine
LIT	linear ion trap
m/z	mass-to-charge
MCM	minichromosome maintenance
MeOH	methanol
MES	2-(N-morpholino)ethanesulfonic acid
$MgCl_2$	magnesium chloride
MMR	mismatch repair
MMS	methyl metanosulfonate
MOAC	metal oxide affinity chromatography
MRM	multiple reaction monitoring
mRNA	messenger RNA
MS	mass spectrometry
MSA	multistage activation
MSH2/3/6	mutS homologue 2/3/6
Na_2CO_3	sodium carbonate
$Na_2S_2O_3$	sodium thiosulfate
NaCl	sodium chloride
NAD	nicotinamide adenine dinucleotide
NaOH	sodium hydroxide
NCBI	national center for biotechnology information
NER	nucleotide excision repair
NH_4HCO_3	ammonium bicarbonate

NH_4OH	ammonium hydroxide
NP-40	nonidet P-40
nt	nucleotide
NTA	nitrilotriacetate
NuRD	nucleosome remodelling histone deacetylase complex
NuRF	nucleosome remodelling factor
OD	optical density
ORC	origin replication complex
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PD	plasma desorption
pI	isoelectric point
PIP	PCNA interacting peptide
PIPES	1,4-piperazine diethane sulfonic acid
PMSF	phenylmethylsulfonyl fluoride
Pol	DNA polymerase
ppm	parts per million
PRC2	polycomb repressor complex 2
pre-IC	pre-initiation complex
pre-RC	pre-replication complex
Q-q-LIT	quadrupole linear ion trap
Q-q-Q	triple quadrupole
Q-TOF	quadrupole time of flight
Rb	retinoblastoma
rDNA	ribosomal DNA
RF	radiofrequency
RFC	replication factor C
RNA	ribonucleic acid
Rnase H1	RNA nuclease H1
RP	reverse phase
RPA	replication protein A
RT	retention time
S/N	signal-to-noise
SAHF	senescence associated heterochromatin foci
Sas2	something about silencing 2
SCX	strong cation exchange
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel

Ser	serine
Sir	silent information regulator
SLBP	stem-loop binding protein
snRNP	small nuclear ribonucleoprotein
SSB	single strand break
ssDNA	single stranded DNA
SUV39H1	suppressor of variegation 3-9 homologue 1
SV40	simian virus 40
T Ag	T antigen
TAP	tandem affinity purificatin
TCEP	tris (2-carboxyethyl) phosphine
TEL	telomere
TFA	trifluoroacetic acid
Thr	threonine
TiO_2	titanium dioxide
TLS	translesion synthesis
TOF	time of flight
TPST1/2	tyrosylprotein sulfotransferase 1/2
Tyr	tyrosine
URA	uracyl
UV	ultraviolet
XIC	extracted ion chromatogram
XPG	xeroderma pigmentosum G
YPD	yeast peptone dextrose

Aknowledgements

First of all, I would like to thank my research director, Dr. Pierre Thibault for giving me the opportunity to work in his lab and gain experience in the field of mass spectrometry as well as for his continual guidance, support and advice throughtout my Ph.D. I would also like to express my gratitude to my co-director, Dr. Alain Verreault for giving me the opportunity to work on exciting projects, for always having time for me, his advice and constant support, motivation through all of his helpful discussions and for providing me with a scientifically motivating environment. I will always be greatful for the opportunities they have given me to advance in their respective fields.

I would also like to thank Dr. Benoit Coulombe and Dr. Martine Raymond for being on my internal thesis committee and for their guidance of my research project throughout many committee meetings.

I am grateful to all past and present members of the Thibault and Verreault lab, for providing me help whenever needed, scientific discussions and constant moral support. I would also like to acknowledge Dr. Eric Bonneil for his help and advice in mass spectrometry and Eun-Hye Lee for her assistance in molecular biology. Special thanks go to Louiza, Gaelle and Eun-Hye for their friendship.

I would also like to thank my parents and family who always encouraged me to pursue my goals in life. Special thanks go to my husband who had to experience first hand the life of a graduate student. He tolerated the long nights of research and encouraged me whenever I needed encouragement which promoted me to keep going through difficult times. I will always be grateful.

Finally, I would also like to thank MDS Sciex and the Faculté des Études Supérieures de l'Université de Montréal for their scholarships.

1. Introduction

1.1. Introduction to chromatin structure and organization

The organization of DNA into chromatin fibres is a truly elegant mechanism that eukaryotic cells have evolved to regulate genome function. Although once widely thought to be mainly required for storing 2 meters of genomic DNA in a confined nucleus of just 10 μm in diameter, chromatin is also a fundamental regulator of many DNA-templated processes. In fact, changes in chromatin structure are at the basis of important functions of DNA such as gene transcription, DNA replication, DNA repair and silencing.

Histones are highly conserved proteins that play a central role in modulating the structure and the dynamics of chromatin. During S-phase of the cell cycle these highly conserved proteins associate with DNA to form the first level of packaging of the genetic material. Consequently, cells that replicate their DNA are also faced with the challenge of accurately duplicating their chromatin structures in an efficient and organized manner. This complex regulation is achieved partly through the combined actions of histone post-translational modifications (PTMs) and histone chaperones, whose function is to accurately escort histones with specific modification patterns from their point of synthesis to their delivery sites on newly replicated DNA.

This thesis will describe chromatin structure and the complexity of its maintenance during DNA replication. More specifically, the first part of the introductory chapter will focus on our current understanding of chromatin assembly and will highlight the key components that couple histone synthesis to their deposition on newly replicated DNA during S-phase.

1.1.1. The nucleosome

In the 1970s, the first insights into the nucleosomal organization of DNA were provided. A combination of physical and molecular biology studies revealed the bead-like structure of the nucleosome and suggested that it is mainly composed of DNA that is wrapped around proteins (Olins, 2003). Subsequent elucidation of the crystal structure of the nucleosome confirmed the molecular details of this nucleoprotein complex (Luger, 1997).

The nucleosome is the fundamental structural and repetitive unit of chromatin. Each nucleosome is composed of about 165 to 200 base pairs of DNA wrapped around an octameric histone protein core. This DNA length, which also comprises the linker region that separates adjacent nucleosomes, varies between different eukaryotic organisms and cell types of the same organism. The actual nucleosomal core consists of 147 base pairs of DNA wrapped in approximately 1.7 circular turns (~ 86 bases pairs per turn) around an octamer composed of two copies of each of the four core histones (Figure 1.1). The histone octamer, created by the direct association of a tetramer containing two copies each of histones H3 and H4 with two histones H2A and H2B dimers, is shaped like a cylindrical disc with a diameter of 13 nm and a height of 11 nm. Each of the histone proteins contain a globular domain known as the histone fold, composed of three alpha helices which are connected by two loops. Pairs of core histones use their globular domains to associate with each other through a handshake-like interaction surface. In addition, each histone contains unstructured N- and/or C-terminal globular domain extensions, comprising between 20-35 residues that are rich in basic amino acids and protrude outside of the nucleosomal surface. These exposed tails are critical for the folding of nucleosomes into higher-organization levels and the structure of chromatin as a whole. (Chakravarthy, 2005; Khorasanizadeh, 2004)

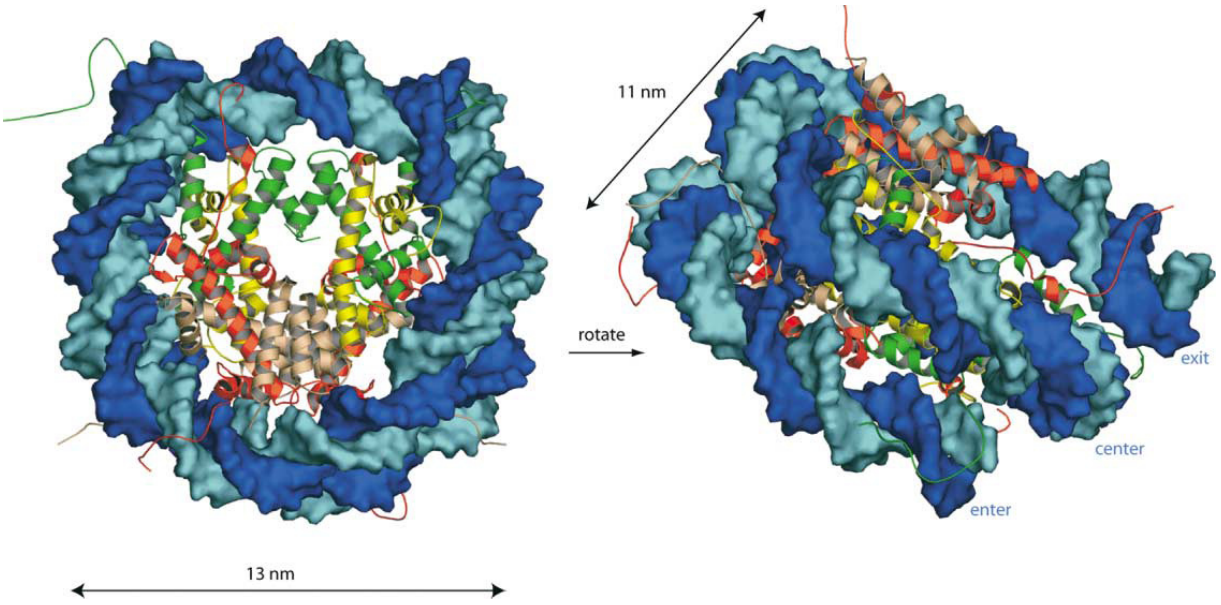


Figure 1.1 The crystal structure of the yeast nucleosome core particle. The DNA molecule is wrapped around the histone octamer about 1.7 times to form an overall disk-like structure. The two strands of the DNA helix are coloured in dark and light blue. Histones are coloured as followed: H3: green, H4: yellow, H2A: red, H2B: pink (Khorasanizadeh, 2004)

1.1.2. Histone post-translational modifications

Years of research have documented a variety of post-translational modifications that are mainly clustered on the tails of histone proteins. The tails protrude outside of the nucleosome core particle where they are readily accessible to histone-modifying enzymes. These modifications include acetylation, methylation, phosphorylation, ubiquitylation and poly (ADP-ribosylation). It has been proposed that these covalent modifications constitute a « histone code » that is deciphered by various cellular machineries to influence chromatin structure and modulate a variety of nuclear processes (Jenuwein, 2001; Strahl, 2000).

One of the best characterized histone modifications is the acetylation of lysine residues (Kurdistani, 2003) which is associated with many functions. For example, mutational studies have provided evidence for the involvement of histone acetylation in transcription. Acetylation of lysine residues reduces the net positive charge of histone tails thereby weakening their affinity to DNA molecules. As a result, the genetic material is more accessible to transcription factors. Furthermore, acetylation can also regulate DNA replication and repair by specifically recruiting various replication and repair proteins that have modules that interact with acetylated residues. Such domains are termed bromodomains.

Experimental evidence also supports the notion that specific acetylation patterns can regulate the deposition of histones onto DNA and the formation of the nucleosome itself. In proliferating cells, the bulk of histone synthesis occurs during S-phase of the cell cycle, when new molecules of histones H2A, H2B, H3 and H4, the four main core histones, are deposited onto replicating DNA by histone chaperones to form nucleosomes. Immediately after their synthesis, histones H3 and H4 are deposited onto DNA in a pre-acetylated form. Newly synthesized histone H4 molecules were shown by pulse-labelling to be acetylated on 2 specific residues, lysines 5 and 12 (Chicoine, 1986; Sobel, 1995). This pattern of acetylation on new H4 molecules differs from those found in mature chromatin, in which H4 can be acetylated on one, two or three lysine residues, namely lysines 5, 8, 12 and 16. Even though acetylation of lysines 5 and 12 in new H4 molecules is evolutionary conserved, its function in chromatin assembly remains poorly understood. In yeast, acetylation of H4 lysines 5 and 12 is not essential for cell viability or nucleosome assembly. However, nucleosome assembly is severely hampered when H4 is simultaneously mutated at three residues, lysines 5, 8 and 12 in combination with mutations that prevent acetylation of the H3 N-terminal tail. This data suggests that a functional redundancy exists between H3 and H4 N-terminal tail acetylation sites (Ma, 1998). Finally, it has been proposed that acetylation of yeast histone H4 at lysine 91, within its core domain, plays an important role in the stability of the histone octamer and the proper

formation of chromatin structure (Ye, 2005). For histone H3, the acetylation pattern varies among different organisms. In budding yeast, new H3 molecules are preferentially acetylated on lysine 9 (Kuo, 1996) while in *Tetrahymina*, H3 is acetylated on lysines 9 and 14. In *Drosophila melanogaster*, lysines 14 and 27 are the preferred sites of acetylation (Sobel, 1995).

One recent and interesting finding is the highly abundant acetylation of newly synthesized H3 molecules on lysine 56, with levels that specifically peak during S-phase (Masumoto, 2005). In yeast, H3 lysine 56 acetylation plays a unique role in nucleosome assembly during DNA replication. Importantly, H3K56Ac acts through a mechanism that is non-redundant with other H3 and H4 acetylation sites and promotes replication-coupled nucleosome assembly by increasing the binding affinity of H3 molecules to histone chaperones such as CAF-1 (Li, 2008; Masumoto, 2005). Moreover, yeast cells that are defective in H3K56Ac are sensitive to DNA damaging agents that interfere with DNA replication (Driscoll, 2007; Han, 2007a; Han, 2007b; Masumoto, 2005). However, compared to the budding yeast *Saccharomyces cerevisiae*, H3K56Ac is much less abundant in human cells and whether H3K56Ac is involved in chromatin assembly in a mechanism that is similar to yeast remains an open question (Garcia, 2007).

Histone acetylation is performed by enzymes known as histone acetyltransferases (HATs), which catalyze the addition of an acetyl group from acetyl coenzyme A (acetyl-CoA) to the ε-amino group of lysine residues. Histone deacetylases (HDACs) catalyze the reverse reaction, namely the removal of acetyl groups from lysine residues. HATs can be divided into two groups: nuclear A-type HATs and cytoplasmic B-type HATs. A-type HAT enzymes have been implicated in gene transcription and DNA repair, whereas B-type HATs are involved in catalyzing the acetylation of newly synthesized histones (Brownell, 1996). One of the first identified B-type HAT is the evolutionary conserved Hat1, which forms a complex with Hat2 and catalyzes the acetylation of H4 on lysines 5 and 12 (Ai, 2004; Ruiz-García, 1998). The HAT enzyme responsible for the deposition-related

acetylation of H3 molecules in *Saccharomyces cerevisiae* is unclear, although it has been suggested that a new GCN5-dependent complex termed HATB3.1, which shows a strong preference for free histone H3, might be responsible for lysine 9 acetylation (Sklenar, 2004). Finally, several groups have shown that Rtt109 is a member of a novel HAT family that acetylates H3K56 in a mechanism that requires both Asf1 and Vps75 chaperones (Driscoll, 2007; Fillingham, 2008; Han, 2007a; Schneider, 2006; Tsubota, 2007). In flies and humans, recent studies have shown that the Rtt109 homologue CBP/300 promotes H3K56 acetylation providing supporting evidence that acetylation of H3K56 by CBP/p300 promotes assembly of newly synthesized histones into chromatin in higher eukaryotes as well (Das, 2009; Tang, 2008).

1.1.3. Euchromatin and heterochromatin

During most of the cell cycle, chromatin from eukaryotic genomes is packaged into two varieties of structures named euchromatin and heterochromatin. Each type of chromatin structure is formed by the presence of distinct chromosomal proteins and histone post-translational modifications. Euchromatin is considered an open structure with loosely spaced nucleosomes and constitutes active genes that are proficient for transcription and replicate early during S-phase. On the other hand, heterochromatin is a highly condensed and relatively inaccessible structure that comprises late-replicating genes and contains either transcriptionally silenced genes or genomic regions that are devoid of genes (Richards, 2002). In the eukaryotic genome, several regions are packaged into specialized heterochromatin structures that contribute to various epigenetic phenomena. One classic example is the inactivation of the X-chromosome in female mammals, where one of the two X-chromosomes is packaged into silent chromatin while the active X remains transcriptionally active despite being exposed to the same nuclear environment (Richards, 2002). In yeast, regions near telomeres, mating type loci and ribosomal DNA repeats present many features of heterochromatin. For instance euchromatic reporter genes that are normally expressed in cells become strongly silenced when inserted within or near these

heterochromatic regions. This difference in gene expression reflects differences in chromatin structure (Rusche, 2003). Studies in fission yeast and metazoans have also established a role for heterochromatin in chromosome segregation, as mutations in factors that interfere with heterochromatin formation cause defects in cohesion and segregation of sister chromatids (Bernard, 2001).

In the budding yeast *Saccharomyces cerevisiae*, the establishment of silent chromatin at the silent mating type loci, known as *HML*α and *HMR***a**, is determined by *cis*-acting elements called silencers and *trans*-acting proteins complexes such as Rap1, Abf1 and the Origin Recognition Complex (ORC). These bind to silencer DNA and help recruit Silent Information Regulator (Sir) proteins (Rusche, 2003). Sir2, Sir3 and Sir4 are essential for efficient silencing at both the *HML*α and *HMR***a** mating type loci and sub-telomeric regions, whereas Sir1 is specifically involved in silencing the mating type loci (Aparicio, 1991). While numerous histone acetylation sites are present in active euchromatin regions, nucleosomes in heterochromatin regions are hypoacetylated (relative to euchromatin) at all lysine residues of histones H3 and H4 and it has been demonstrated that Sir proteins are preferentially recruited to hypoacetylated histones (Carmen, 2002; Suka, 2001). Of particular importance for gene silencing is the Sir2-mediated deacetylation of histone H4 on lysine 16. Studies have suggested that deacetylation of H4K16 by Sir2, an evolutionary conserved NAD^+-dependent histone deacetylase (HDAC), provides high-affinity binding sites for additional Sir proteins in adjacent nucleosomes and allows the spreading of the Sir complex along chromatin to form silent heterochromatin (Imai, 2000; Rusche, 2003; Tanny, 1999). Sas2 (Something about silencing 2), a histone acetyltransferase (HAT), counteracts the HDAC activity of Sir2 by acetylating H4K16. This creates a boundary that prevents heterochromatin from spreading into adjacent euchromatin regions (Kimura, 2002; Suka, 2002). In recent years, histone methylation has also emerged as an important mark that is associated with functionally specialized chromatin structures (Grewal, 2003). For instance, in fission yeast, flies and mammals, H3K9 methylation is associated with heterochromatin formation, whereas H3K4 and H3K79 di-methylation are excluded from

heterochromatin regions and are linked with transcriptionally active domains (Sims, 2003). Methylation of H4K9 by Clr4 in fission yeast and its orthologue SUV39H1 in flies and mammals directly binds to key structural components of heterochromatin known as HP1 proteins (Swi6 in fission yeast) (Bannister, 2001; Lachner, 2001; Nakayama, 2001; Rea, 2000). Once bound to K9-methylated H3 molecules, these proteins recruit histone modifying enzymes that create additional docking sites for Swi6/HP1 molecules, thus allowing the spreading of repressive chromatin (Hall, 2002) in a manner analogous to the spreading of Sir proteins in *Saccharomyces cerevisiae* heterochromatn. It has also been proposed that acetylation of H3K9 counteracts methylation of this site (the two modifications are mutually exclusive) and blocks the spreading of heterochromatin structures to adjacent regions (Litt, 2001). One striking observation is that, despite the divergence of the various proteins that are involved in heterochromatin formation and gene silencing from one organism to another, the molecular events that mediate the establishment and spreading of heterochromatin structures are highly conserved. In each case, the covalent modification of histones is necessary for recruitment of enzymes and structural proteins that polymerize and spread in a directional manner to form heterochromatin domains (Grewal, 2003; Moazed, 2001).

Once the silenced state of chromatin is established, the specific patterns of histone modifications must be preserved, such that silenced chromatin structures are faithfully maintained during DNA replication in S-phase and inherited after each cell cycle division, thereby ensuring perpetuation of the integrity and stability of heterochromatin structures (Grewal, 2002; McNairn, 2003). During DNA replication, the transient disruption of pre-existing heterochromatin structures raises the important question of how specialized chromatin states are reformed after replication. Several observations strongly suggest that some factors that function in DNA synthesis are also involved in epigenetic inheritance. For example, hypomorphic mutations in several different replication proteins reduce heterochromatin-mediated gene silencing in *Saccharomyces cerevisiae* and higher eukaryotes without affecting DNA replication processes (McNairn, 2003; Wallace, 2005).

One particularly important protein in which specific mutations affect heterochromatin-mediated silencing is the DNA polymerase processivity factor PCNA (Sharp, 2001; Zhang, 2000). Furthermore, studies have also demonstrated that chromatin assembly factors whose function is specifically coupled to DNA replication contribute to the maintenance and inheritance of silenced chromatin. Consistent with this, mutations in genes coding for chromatin assembly factors such as CAF-1 cause defects in gene silencing in sub-telomeric regions and the silent mating type loci (Enomoto, 1998; Monson, 1997). Interestingly, CAF-1 is targeted to replication forks during S-phase by direcly binding to PCNA (Shibahara, 1999). Precisely how these two proteins participate in the process of chromatin formation remains poorly understood. The following sections provide a detailed description of the process of DNA replication, followed by an overview of chromatin assembly events that occur behind the replication fork and how these processes work in conjunction to maintain chromatin structure in proliferating cells.

1.2. Genome duplication

The complete, accurate, and processive replication of DNA is vital for the maintenance of cell viability and genomic integrity in all organisms. As a result, eukaryotic cells have evolved molecular regulatory mechanisms to ensure that the duplication of the genetic material occurs with high accuracy and fidelity and, most importantly, takes place only once per cell cycle during S-phase (Bell, 2002; Waga, 1998).

1.2.1. DNA replication

DNA replication in eukaryotes is a complex process that involves the interaction of many evolutionary conserved proteins that act at chromosomal sites known as origins of replication. In *Saccharomyces cerevisiae*, each origin contains *cis*-acting “replicators”, which are specific DNA sequences known as autonomously replicating sequence (ARS)

elements. ARS contain multiple short functional elements, termed A1, B1, B2 and B3. The A1 element, which is highly conserved among all budding yeast origins, comprises an 11 bp ARS consensus sequence (ACS) that is essential for DNA replication. The A1 and B1 elements act together to create DNA binding sites for the origin recognition complex (ORC) and they are required for initiation of DNA replication (Rao, 1995). The remaining B elements are not essential but are collectively required as enhancers of origin efficiency (Rao, 1994; Theis, 1994). ARS are also present in other eukaryotes and humans; their sequences are however less well defined and can extend from 800 to over thousands of base pairs of DNA (Bell, 2002).

In prokaryotes, replication initiates at a single origin of a circular genome. In contrast, initiation of DNA replication in eukaryotes occurs at multiple origins distributed along linear chromosomes. The use of numerous origins of replication is an elegant mechanism that cells have evolved to efficiently replicate the entire eukaryotic genome within the relatively short time span of S-phase (Hyrien, 2003). Despite years of study, few origins have been mapped and functionally dissected in mammals and several other model organisms. The only eukaryote where origins were mapped on a genomic scale is *Saccharomyces cerevisiae*. Using DNA microarray analysis, it was shown that the budding yeast genome, which comprises 16 chromosomes, contains about 332 origins of replication (Raghuraman, 2001; Yabuki, 2002). Replication origins were also mapped by analyzing ORC and MCM2 (a component of the replicative DNA helicase that is recruited to ARS prior to origin firing) binding sites and these studies identified about 529 potential origins of replication (Wyrick, 2001; Xu, 2006).

Replication at multiple origins is controlled in a temporal order and individual origins are fired at different times throughout S-phase. Several observations have suggested a relationship between replication timing and the transcriptional potential of genes and/or the types of chromatin that flank the origins (Goren, 2003; McNairn, 2003). Most transcriptionally active euchromatin genes are replicated in the first half of S-phase whereas

most heterochromatin is replicated late in S-phase. In both yeast and mammalian cells, this timing is established in early G1 phase, where origins are repositioned and modified in specific nuclear compartments (Goren, 2003; Li, 2001; Raghuraman, 1997). This temporal program, which coordinates origin firing within tanscriptionally active and inactive regions, may be important for the propagation of specialized chromatin features to daughter cells (Dillon, 2002; Flickinger, 2001).

1.2.1.1. Replication initiation

DNA replication from individual origins is generally bidirectional and tightly coordinated with cell cycle events. The basic mechanism of initiation of DNA replication in eukaryotic cells occurs in two steps. The origin licensing process, also known as the assembly of the pre-replicative multiprotein complex (pre-RC), is followed by the origin activation process (Figures 1.2, 1.3). Assembly of the pre-RC is initiated by DNA sequence specific binding of the origin recognition complex (ORC) to the replication origin (Bell, 1992). At the end of mitosis and early G1, Cdc6 and Cdt1 are independently recruited to ORC-bound replication origins (Figure 1.2). ORC, Cdc6 and Cdt1 form a complex, and are collectively required for stable association of the minichromosome maintenance 2-7 complex (MCM 2-7) helicase with the origin to form the pre-RC (Speck, 2007; Tanaka, 2002). At the G1/S phase transition and throughout S-phase, the conversion of the pre-RC into the pre-initiation complex (pre-IC) is necessary before DNA synthesis can begin (Figure 1.2). The pre-IC contains additional proteins that activate the DNA helicase activity of the MCM 2-7 complex which then unwinds the two DNA strands as a pre-requisite for origin firing. These include Cdc45, Sld2, Sld3 and Dpb11 as well as S-phase cyclin-dependent kinases (CDKs) and the Dbf4-dependent kinase (DDK or Cdc7-Dbf4) which regulate distinct steps in the activation of replication origins (Zou, 2000).

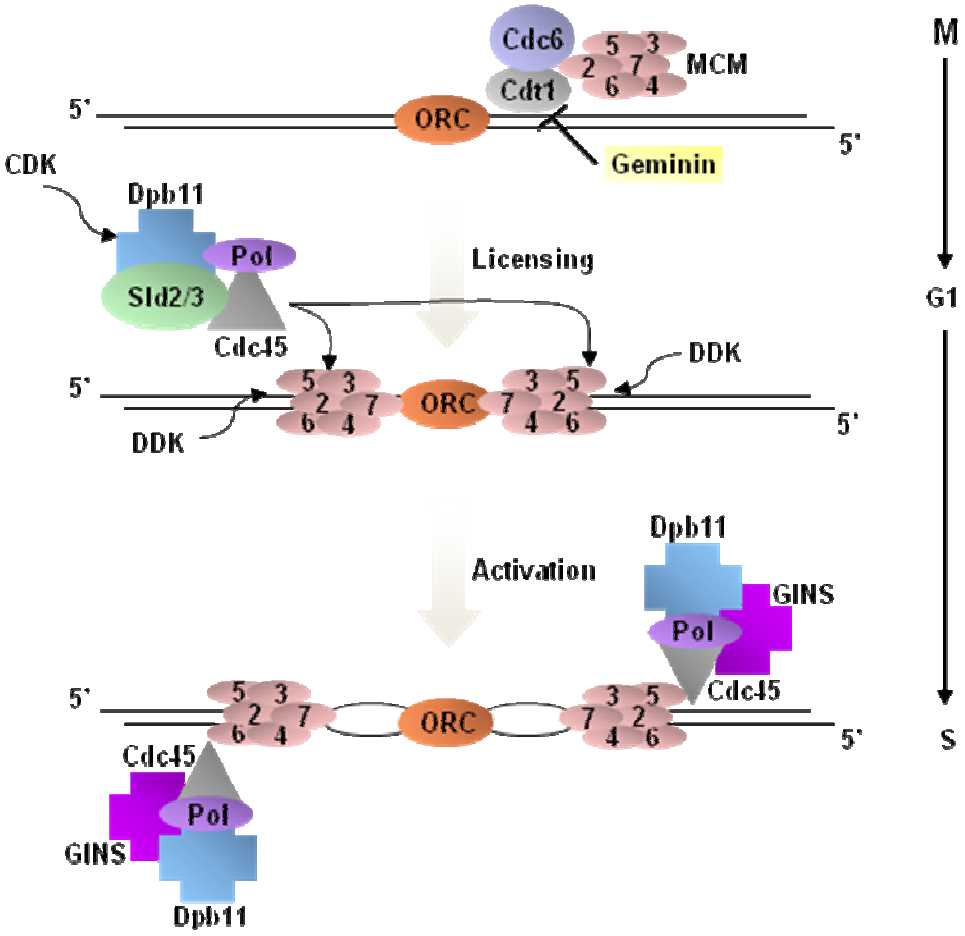


Figure 1.2 Initiation of DNA replication in eukaryotes. Replication origins are recognized by ORC. Cdt1, Cdc6 and the MCM helicase complex subsequently associate with origins in G1 to form pre-RCs. At this point, origins are "licensed" but not yet activated. In late G1, CDK and DDK become activated and in turn trigger origin activation by phosphorylating Sld2, Sld3 and the MCM proteins. This leads to the activation of the hexameric MCM helicase (MCM2-7), DNA unwinding and recruitment of other enzymes into the "replisome". In order to prevent inappropriate pre-RC formation and re-replication in S, G2 and M phases, CDK also phosphorylate and inactivate certain components of pre-RC and inhibit further re-licensing after origin firing. In vertebrates, another level of regulation to block re-replication is mediated through geminin, a protein that inhibits Cdt1 and prevents the reloading of MCM proteins without interfering with ORC and Cdc6 association in G1 (Sclafani, 2007).

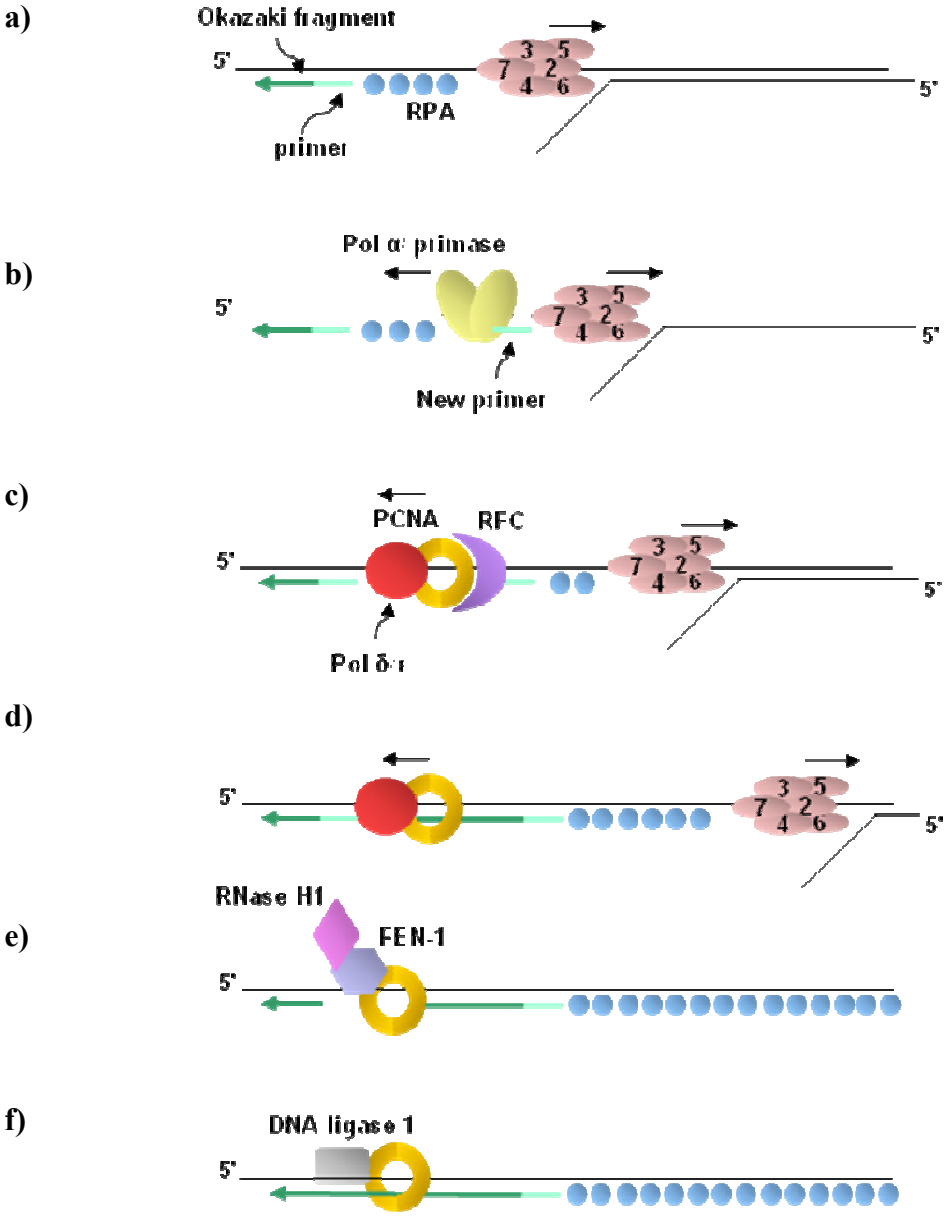


Figure 1.3 The process of DNA replication in S-phase. The events taking place during DNA synthesis of the lagging strand are shown. Synthesis of the leading strand (not shown) requires only a subset of these events because it is synthesized in a continuous manner from a single primer, rather than in short fragments known as Okazaki fragments. These fragments need to be joined together by Okazaki fragment maturation enzymes whose action is only needed during lagging strand synthesis. **a)** The DNA helix is unwound by the MCM helicase and the resulting single stranded DNA (ssDNA) is coated with RPA proteins. **b)** Pol α/primase is recruited to the RPA coated ssDNA and synthesises a short

RNA/DNA hybrid primer, displacing the RPA molecules as it travels along the template. **c)** The clamp loader RFC recognizes the completed primer-template junction, displaces Pol α/primase and catalyses the loading of the PCNA sliding clamp and DNA polymerases δ or ε. **d)** Through their binding to PCNA, pol δ/ε processively synthesise DNA from the RNA/DNA primer until they reach the 5' end of the next Okazaki fragment. Similar events take place at the leading strand, however since DNA synthesis occurs in 3'-5' direction, leading strand synthesis is continuous. **e)** The RNA/DNA primer of the previous Okazaki fragment is removed by FEN-1 and RNase H1. **f)** The single-stranded gap generated by primer removal is filled in and the two DNA segments are joined by DNA ligase 1 (Morgan, 2007).

Some evidence indicates that phosphorylation of the MCM protein complex by Cdc7-Dbf4 is required for Cdc45 loading, which is an essential step for origin activation (Sclafani, 2002; Zou, 2000). It has been hypothesized that phosphorylation of MCM proteins induces a conformational change in the MCM2-7 complex that activates the helicase and is a signal for Cdc45 binding to origins (Fletcher, 2003; Sclafani, 2002). The essential role of CDKs in promoting origin activation has also been determined. CDKs phosphorylate Sld2 and Sld3 which enable them to interact with Dpb11, a subunit of the DNA polymerase ε holoenzyme (Tanaka, 2007; Yabuuchi, 2006; Zegerman, 2007). CDK phosphorylation of Sld2 is absolutely required for replication initiation and for Cdc45 loading onto origins (Masumoto, 2002). The GINS complex (consisting of Sld5, Psf1, Psf2 and Psf3) is also required for loading of Cdc45 onto origins and replication initiation (Gambus, 2006; Takayama, 2003; Yabuuchi, 2006). The combined action of CDKs and DDK is required to load other components of the replisome and activation of the DNA helicase, which unwinds DNA and allows DNA replication to proceed.

1.2.1.2. Replication elongation

Much of our knowledge of the steps taking place during DNA replication comes from the study of the simian virus 40 (SV40). In this system, DNA is replicated by the viral T antigen (T Ag) that acts in concert with other cellular proteins (Waga, 1998). After site-specific recognition of the origin of DNA replication by T antigen, its intrinsic helicase activity unwinds the DNA double helix near the origin. This step is necessary for binding the heterotrimeric single-stranded DNA (ssDNA) binding protein, known as replication protein A (RPA) (Figure 1.3a). RPA binds to stretches of ssDNA to protect and stabilize DNA strands exposed during DNA replication or repair (Wold, 1997). In a subsequent step called primosome assembly, the DNA polymerase α/primase complex (pol α/primase) is loaded on DNA through a direct interaction with RPA (Walter, 2000) (Figure 1.3b). The pol α/primase complex contains four evolutionary conserved subunits (p180, p70, p58 and p48 in human cells). The p180 and p48 subunits harbour the polymerase and primase activity, respectively. Once the replisome is assembled, pol α/primase synthesizes *de novo* a short hybrid RNA-DNA primer that contains about 30 nucleotides (nts) (Figure 1.3b). DNA pol α is the only eukaryotic enzyme that can initiate synthesis *de novo*, *i.e.* without the help of a primer (Hübscher, 2002). The two activities of DNA pol α/primase, DNA primase and DNA polymerase, act sequentially to generate the RNA-DNA primer. Key to this reaction is the primase activity, which can lay down a short RNA primer that acts as a seed for DNA polymerase α. Next, replication factor C (RFC), a heteropentameric ATPase protein complex, binds tightly to the primer template junction and locally promotes the assembly of the primer recognition complex (Mossi, 1998). This consists in loading the ring-shaped DNA sliding clamp Proliferating Cell Nuclear Antigen (PCNA) and its subsequent association with the pol δ holoenzyme (Figure 1.3c). The PCNA ring is able to slide freely along the DNA and, in conjunction with the replicative DNA polymerase, promotes processive DNA synthesis on the leading strand for at least 5-10 kilobases (kbs) (Matsumoto, 1990). Since DNA synthesis always occurs in the 5`to 3`direction, the lagging

strand needs to be synthesized in the direction opposite from that of the replication fork. As a result, lagging strand synthesis is discontinuous and divided into short, 180 nts Okazaki fragments (Hübscher, 2001) (Figure 1.3d). *In vitro* in the SV40 system, Pol δ seems to be the main replicative polymerase and is able to support DNA synthesis of both the leading and lagging strands (Fukui, 2004; Hübscher, 2002; Waga, 1998). Although pol ε is not essential for SV40 DNA replication *in vitro*, it is essential for cellular DNA replication (Karthikeyan, 2000; Waga, 2001; Zlotkin, 1996). Even though more studies are required to delineate the precise function of Pol ε in DNA replication, it has become clear that, *in vivo*, both Pol δ and Pol ε play essential roles as replicative DNA polymerases (Fuss, 2002). Recently, the first evidence for Pol ε to support DNA synthesis of the lagging strand was demonstrated (Pursell, 2007). In addition, Pol ε plays a unique role in a checkpoint that signals DNA damage during replication (Fuss, 2002).

Finally, the lagging strand is converted into a long continuous DNA product by a process called Okazaki fragment maturation. After the replicative DNA polymerase encounters an RNA primer at the 5` end of the previous Okazaki fragment, the primer is completely removed (Figure 1.3e) by two different nucleases, RNase HI and a 5`-3` endonuclease called Flap endonuclease 1 (FEN-1) (Hübscher, 2001; Waga, 1998). RNase HI cleaves the RNA primer at the 5`end of each Okazaki fragment leaving a single ribonucleotide which gets removed by FEN-1. PCNA also binds to FEN-1 and stimulates FEN-1 activity by at least 50-fold (Li., 1995). Recent biochemical and genetic studies strongly support the existence of a redundant pathway involving the participation of the helicase/endonuclease Dna2 with FEN-1 (Bambara, 1997; Hübscher, 2001). One hypothesis is that Dna2 displaces the RNA segment from the DNA template which then gets cleaved off by FEN-1. The gap created is then filled in by either pol δ or pol ε. The nicks left between adjacent Okazaki fragments are effectively ligated by another PCNA binding protein, DNA ligase 1 (Figure 1.3f). The superhelical strain that is caused by the unwinding of the DNA double helix is released by DNA topoisomerase I and then

replicated DNA molecules are decatenated by DNA topoisomerase II (Hübscher, 2001; Montecucco, 1998; Wang, 2002).

As mentioned above, "*lagging*" strand synthesis requires several reactions that must be tightly coordinated with synthesis of the "*leading*" strand. This coordination depends, at least in part, on PCNA and is required in order to avoid synthesis of the lagging strand from falling behind by hundred of nucleotides from leading strand synthesis events (Hübscher, 2001; Lee, 1998; Prelich, 1988). Furthermore, the DNA polymerase on the lagging strand is transferred from one Okazaki fragment to the other, rather than having new molecules being reloaded every 200bp, which also favours processivity of lagging strand DNA synthesis. Finally cells have evolved a looping back mechanism on the lagging strand that allows both strands to extend in the same direction and at similar rates. Studies using the bacteriophage T7 replication system have recently highlighted the importance of the primase-helicase T7-gp4 in limiting the speed of the leading strand polymerase. In fact, the looping of the lagging strand has been shown to allow the T7-gp4 primase-helicase to maitain contact with the nascent primer on the lagging strand as well as allowing the primer to be withing physical reach of the lagging strand replication complex. This mechanism of coordination provides efficient primer utilization and ensures adequate primer hand-off to the lagging strand T7 polymerase (Pandey, 2003).

1.2.2. DNA repair

The integrity of DNA is continuously challenged by DNA replication errors, metabolic products that damage DNA (e.g. oxidizing species) and genotoxic agents from the environment (e.g. chemical from tobacco smoke). Efficient detection of DNA damage and nucleotides misincorporated during DNA replication is particularly vital for dividing cells as replication and segregation of unrepaired DNA can seriously compromise genomic integrity. In order to ensure accurate transmission of genetic information, eukaryotic cells

have evolved elaborate surveillance and repair mechanisms that rapidly recognize and repair damaged DNA at the replication fork (Dinant, 2008).

1.2.2.1. Cell cycle checkpoint activation

In response to DNA damage, cells trigger an evolutionary conserved signalling cascade involving proteins of the DNA damage checkpoint response. These are designed to detect damaged DNA and coordinate cell cycle progression with DNA repair. More precisely, activation of the DNA damage checkpoint results in a number of downstream biological consequences whereby cell cycle progression through the G1, S or M-phase is transiently delayed and ongoing DNA replication is slowed down, providing cells more time to repair DNA lesions (Paulovich, 1995; Weinert, 1988). In addition, various genes and proteins involved in the repair of DNA lesions are activated by transcriptional and post-transcriptional mechanisms and become localized to sites of repair (Brush, 1996; D'Amours and Jackson, 2001; Gasch et al., 2001). Chromatin also undergoes structural changes to allow access of the repair machinery to damaged sites. These changes are mediated by histone post-translational modifications or recruitment of non-histone proteins, such as ATP-dependent chromatin remodelling factors and histone chaperones, to sites of DNA damage (Escargueil, 2008; Osley, 2006; Polo, 2006). At the top of the signalling cascade is a specific group of checkpoint proteins called the phospho-inositol kinase-related proteins. These include the protein kinases ATM/ATR in humans and Mec1/Tel in yeast (Lowndes, 2000; Rouse, 2002). These kinases play a central role in all checkpoint responses as they can respond to DNA damage generated by ultraviolet (UV) irradiation, gamma irradiation, DNA alkylating agents such as methyl methanosulfonate (MMS) and the replication inhibitor hydroxyurea (HU) that acts by depleting deoxyribonucleoside triphosphates (dNTP) (Abraham, 2001; Foiani, 2000; Rhind, 1998). Two classes of checkpoint effector kinases, CHK1 and CHK2 in mammals (Chk1 and Rad53 in *Saccharomyces cerevisiae*) function downstream of the phospho-inositol related kinases. These kinases become rapidly phosphorylated and activated in a Mec1/Tel1 (ATR/ATM)-

dependent manner. Once activated, they modulate the activities of key effector proteins of the DNA damage response (Lowndes, 2000; Rouse, 2002). In budding yeast, an additional kinase named Dun1 is activated by Rad53 in response to DNA damage. Dun1 is involved in the control of transcriptional regulation and is needed to increase dNTP pools (Zhou, 1993).

1.2.2.2. DNA repair mechanisms

A number of highly sophisticated DNA repair pathways have evolved to cope with various types of DNA lesions. Single strand damages are repaired by three major damage repair pathways that include nucleotide excision repair (NER), base excision repair (BER), and mismatch repair (MMR). These processes function by excising a short stretch of the damaged DNA strand and then using the undamaged complementary strand to re-synthesize the correct DNA sequence (Ataian, 2006).

The NER pathway is specifically responsible for removing intra-strand helix distorting lesions such as UV-induced 6-4 photoproducts (6-4PPs) and cyclobutane pyrimidine dimers (CPDs). During NER, a large number of repair as well as replication proteins get sequentially deployed at the site of DNA damage to perform the incision as well as the gap-filling DNA synthesis and ligation steps. Incision is mediated by the Xeroderma pigmentosum G (XPG) endonuclease, a FEN-1 related nuclease that cleaves at the 3' side of the DNA lesion and is followed by an ERCC1-XPF mediated incision at the 5` side of the lesion (de Laat, 1999). Similarly to FEN-1, XPG also interacts with PCNA (Gary, 1997). The DNA repair synthesis step is carried out by DNA polymerase δ or ε, which fill in the gap with the correct nucleotides in a PCNA-dependent manner.

The BER pathway is mainly implicated in the repair of small non-distorting chemical alterations such as alkylation and oxidative damage to DNA bases (Zharkov, 2008). BER is initiated by the action of a DNA glycosylase that binds to the damaged DNA

base and hydrolyzes the N-glycosydic bond to release the damaged base from the deoxyribose. This cleavage generates an apyrymidinic/apurinic (AP) site that is subsequently processed by an AP-endonuclease. The synthesis step in BER is carried out by a short patch (single base) or a long patch (2-7 nts) pathway. In short patch BER, the missing nucleotide is filled in by DNA polymerase β, which also interacts directly with PCNA (Kedar, 2002). On the other hand, long patch BER utilizes the replication elongation apparatus, where FEN-1, PCNA, DNA polymerase β or δ and DNA ligase 1 are required for filling the gap and sealing the nicks (Klungland, 1997).

MMR targets mismatched bases and insertion-deletions mispairs (IDLs) that arise from replication errors (Iyer, 2005). Two distinct MutS-related heterodimeric complexes that have different mismatch recognition specificity can act to promote MMR. MSH2-MSH3 recognizes mispaired bases, while MSH2-MSH6 targets IDLs. The process also requires MLH1 which interact with MSH2 and MSH3 as well as other components of the replication machinery, including PCNA. Several studies have highlighted the important function of PCNA in MMR (Genschel, 2003; Gu, 1998; Umar, 1996). Importantly, PCNA is not only required for the DNA repair synthesis process, where it forms part of the DNA elongation apparatus, but appears to play a critical role in the initiation step as well. Recent observations revealed that PCNA enhances the mispair binding specificity of MSH2-MSH6 when a replication error is encountered (Flores-Rozas, 2000; Lau, 2003) This has led to the proposal that PCNA helps the MMR machinery mediate an efficient search for mismatches on newly replicated regions of DNA.

Finally, when the damaged DNA cannot be repaired, a fourth mechanism known as DNA damage bypass or DNA damage tolerance allows completion of DNA replication without removal of the damaged base. Although the DNA polymerase δ and ε replicate DNA in a highly processive and accurate manner, they are unable of copying DNA containing lesions such as CPDs, damaged bases and AP sites. When the replication machinery encounters such lesions, specialized polymerases termed translesion synthesis

(TLS) polymerases are recruited to stalled replication forks. Despite the conformational constraints imposed by the damaged bases, TLS polymerases can accommodate distorted template structures within their active site and replicate past these lesions. However, in contrast to replicative polymerases, TLS polymerases function with reduced fidelity on damaged DNA templates and are therefore mutagenic (Waters, 2009). Interestingly, several studies have suggested that stalling of the replication machinery at sites of DNA damage triggers post-translational modification of PCNA (Hoege, 2002). More specifically, PCNA ubiquitination promotes polymerase switching, where the replicative enzymes pol δ and ε are replaced by TLS polymerases, such as pol η which specialises in the bypass of UV-damaged CPDs (Kannouche, 2004; Stelter, 2003; Xiao, 2000).

1.3. The role of PCNA in genome duplication and maintenance of chromatin structure

Proliferating cell nuclear antigen (PCNA) is a member of the DNA sliding clamp family, whose structure and function are evolutionary well conserved from yeast to humans (Moldovan, 2007). Alignment of amino acid sequences of PCNA from different organisms shows a relatively low level of sequence conservation. Human and yeast PCNA share only 35% amino acid sequence identity (Bauer, 1990). Nevertheless, crystallographic studies have shown that their three-dimensional structures are highly superimposable even with the homodimeric structure of the functional homologue of PCNA in bacteria, the β-subunit of DNA polymerase III (Krishna, 1994). The latter forms a homodimer and shows essentially no sequence similarity with PCNA in eukaryotes. In contrast to its functional equivalent in prokaryotes, eukaryotic PCNA forms a homotrimeric ring shaped complex with pseudo-six-fold symmetry. Because PCNA contains a central cavity, it is capable of encircling double-stranded DNA and sliding freely along it. Replication factor C (RFC) catalyzes the loading of PCNA by transiently disrupting the PCNA ring and rapidly re-closing it when it encircles DNA (Bowman, 2004; Majka, 2004). Its topological link to DNA allows PCNA

to firmly tether DNA polymerases to the DNA template. *In vitro*, the presence of PCNA increases the processivity of DNA polymerases during elongation of DNA synthesis allowing it to synthesize long strands of DNA without dissociating from the template (Prelich, 1987; Prelich, 1988; Tan, 1986). Accordingly, PCNA is an important co-factor for DNA synthesis and essential for cell viability.

Each PCNA monomer has two very similar globular domains that are linked by a long loop called the interdomain connecting loop (IDCL). Each monomer is arranged head-to-tail to form a ring with two non-equivalent surfaces (Figure 1.4). The inner cavity of the ring is formed exclusively by α-helices and contains positively charged residues that interact with the negatively charged sugar-phosphate backbone of the DNA and facilitates its sliding along the DNA duplex. The outer surface is composed of β-sheets and is mainly responsible for mediating protein interactions (Moldovan, 2007). In addition to interacting with DNA polymerases, PCNA directly binds to a multitude of proteins (Table 1.1) that function concomitantly with replication and serves as a molecular platform for other factors that act in processes that are not temporally linked to DNA replication. These are involved in DNA repair, cell cycle regulation as well as in the propagation of epigenetic information (Maga, 2003). PCNA is also crucially involved in the establishment of sister chromatid cohesion during S-phase (Moldovan, 2006).

One of the most fascinating, yet poorly studied property of PCNA is its ability to imprint newly replicated DNA for subsequent DNA processing events such as new histone deposition and chromatin assembly. This is supported by the observation that PCNA shows little dissociation from established replication sites, suggesting that it remains topologically associated with DNA throughout replication of the DNA leading and lagging strands (Sporbert, 2002). Moreover, while investigating the mechanisms and control of chromatin duplication, Shibahara and Stillman discovered that PCNA indeed marks newly replicated DNA for chromatin assembly and that unloading of PCNA from replicated DNA by RFC reverses that mark (Shibahara, 1999). More specifically, chromatin assembly mediated by

the molecular histone chaperone CAF-1 is inhibited when PCNA is artificially removed from replicated DNA using RFC. CAF-1 and PCNA directly bind each other and, based on immunofluorescence, colocalise with newly replicated DNA during S-phase (Bravo, 1987; Krude, 1995; Shibahara, 1999). Genetic studies performed in yeast have also suggested that PCNA participates in epigenetic inheritance of heterochromatin structures by binding directly to CAF-1 (Zhang, 2000). Several other proteins whose function is associated with the maintenance and remodelling of chromatin structure also interact with PCNA. For instance, the DNA (5-cytosine) methyltransferase 1 (DNMT1) enzyme, which preserves epigenetic methylation patterns following DNA replication and repair, is targeted to the replication fork through PCNA binding via its PIP-box (PCNA interaction peptide) (Chuang, 1997). Another example is the histone deacetylase 1 (HDAC1) enzyme which deacetylates newly deposited histone proteins. Interestingly, HDAC1 is also recruited to sites of DNA replication through a direct interaction with PCNA (Milutinovic, 2002). Taken together, the observation that PCNA interacts with CAF-1, DNMT1 and HDAC1 indicate that PCNA plays an essential role in coupling DNA replication to chromatin assembly and epigenetic inheritance.

a) **b)**

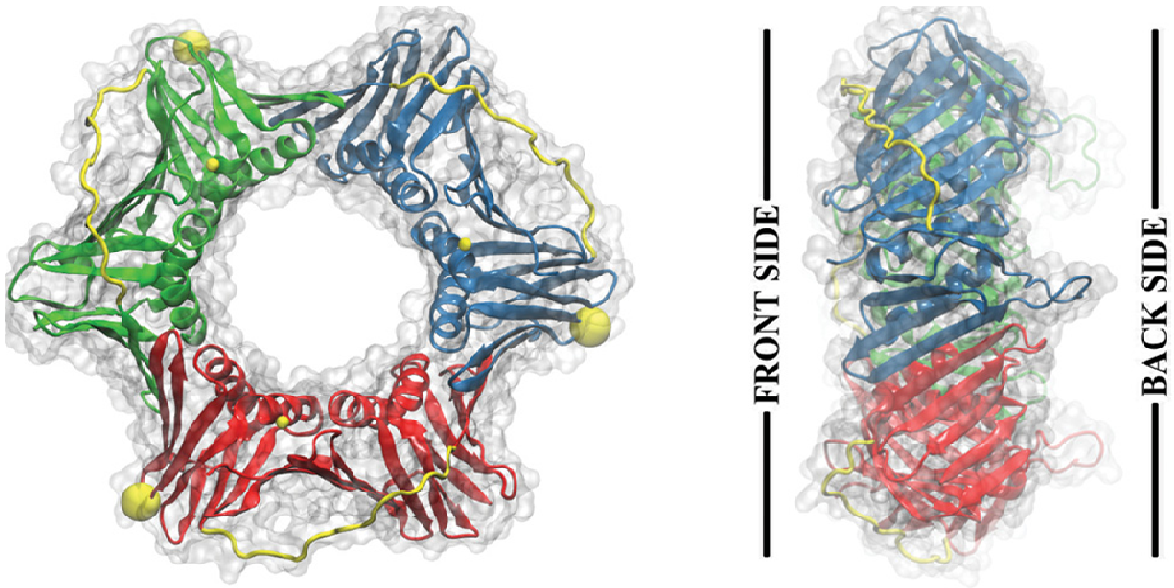


Figure 1.4 Crystal structure of human PCNA. a) Front view of the homotrimeric ring consisting of three subunits coloured in red, blue and green. Each subunit is composed of two domains (N- and C-terminal) that are connected by a long linker called the interdomain connecting loop (IDCL) coloured in yellow. Post-translationally modified residues are indicated with small (phosphorylated Tyr-211) and big (ubiquitinated and SUMOylated Lys-164) yellow spheres. **b)** Side view of human PCNA showing the front and back surfaces. The front side always points in the direction of DNA synthesis and all PIP-box containing proteins bind to the same surface on the front side of PCNA (Stoimenov, 2009).

Most proteins that bind PCNA contain a conserved motif termed the PCNA-interacting peptide (PIP) box with the consensus sequence Q-X-X-**h**-X-X-**a**-**a**, where « h » is a hydrophobic residue (L, I or M) and « a » is an aromatic residue (F or Y but also, more rarely H) and is occasionally N-terminally flanked by the KA-box motif (Warbrick, 1998a; Warbrick, 1998b; Xu, 2001). Exhaustive mutational and structural analyses have revealed key sites involved in mediating interactions between PCNA and PIP-containing proteins (Bruning, 2004; Eissenberg, 1997; Gulbis, 1996; Matsumiya, 2002; Shamoo, 1999). The PIP-box adopts a 3_{10} helix structure, which allows insertion of the PIP box hydrophobic and aromatic residues into a hydrophobic cavity of PCNA situated beneath the IDCL. Residues N-terminal of the PIP-box form a short antiparallel β-zipper interaction where the conserved glutamine residue engages in numerous interactions with residues near the C-terminus of PCNA. The C-terminal end of the PIP-box forms an extended β-sheet with the IDCL of PCNA, thereby stabilizing the overall base of the interaction. Remarkably, all PIP-box containing proteins bind to the front side of PCNA (Figure 1.4b), which is the side that associates with DNA polymerases and faces the direction of DNA synthesis, suggesting the presence of regulatory mechanisms that spatially and temporally control the association of PCNA with different partners that play different roles in DNA metabolism. Several studies have indicated that affinity-driven competition (Bruning, 2004), proteolysis (Wang, 2006) and phosphorylation of PIP-box containing proteins (Henneke, 2003), as well as post-translational modification of PCNA (Hoege, 2002), regulate the interaction of PCNA with its protein partners.

Table 1.1 PCNA binding proteins

Function	Activities	PCNA binding proteins	PIP-box Sequences[§ ¶]
DNA replication	DNA polymerases	<u>DNA Pol δ</u>	
		h p12	KR**L**ITDS**Y**
		h p66	**Q**VS**I**TG**FF**
		h p125	**KA**T**Q**CQ**L**[*]
		sc Pol32	**Q**GT**L**ES**FF**
		sp Cdc27	**Q**KS**I**MS**FF**
		<u>DNA Pol ε</u>	
		sc Pol2	**Q**TS**L**TK**FF**
		sc Dpb2	**Q**MF**L**TK**YY**
	Flap-endonuclease	h FEN-1	**Q**GR**L**DD**FF**
	DNA ligase	h DNA ligase 1	**Q**RS**I**MS**FF**
	Clamp loader	<u>RFC</u>	
		h Rfc1	MD**I**RK**FF**
		sc Rfc1	MVN**I**SD**FF**[*]
			NMS**VV**G**YF**
		sc Rfc3	FASTRQI**F**
		sc Rfc4	Other[#]
	Replication factors	sc Mcm10	**Q**HT**L**DV**Y**I
		sp Cdc24	**Q**YLLIP**FW**[*]
	Replication licensing	xp Cdt1	**Q**MR**V**TD**FF**
	Topoisomerase	g Topo IIα	**Q**TT**L**P**F**K
DNA damage bypass	E3 ubiquitin ligases	sc Rad18	?
		sc Rad5	?
	TLS polymerases	sc DNA Pol η	SKN**IL**S**FF**
		h DNA Pol η	**Q**T**L**ES**FF**
		h DNA Pol ι	KKG**LI**D**YY**
		(h) DNA Pol κ	KHT**L**DI**FF**
		(h) DNA Pol λ	SVP**VL**EL**F**
		(m) REV1	≠
		(sc) DNA Pol ζ	≠

Prevention of sister chromatin recombination	E2 SUMO-conjugating enzyme	(sc) Ubc9	?
	Helicase	(sc) Srs2	?
Mismatch repair (MMR)	MMR sensors	(sc) Msh3 (sc) Msh6	**Q**PT**I**SR**FF** **Q**SS**LL**S**FF**
	MMR transducer	(sc) Mlh1	**Q**IG**L**TD**F**A
	Exonuclease	(h) EXO1	?
Base excision repair (BER)	Glycosylases	(h) UNG2 (h) MPG (h) NTH1 (h) MYH	**Q**KT**L**YS**FF* **F**CMN**I**SS**Q** (inverted) ≠ **Q**QV**L**DN**FF**
	Endonucleases	(sc) Apn2 (h) APE1 (h) APE2	NKS**L**DS**FF** not investigated **Q**KN**L**KS**YF**
	DNA polymerases	DNA pol β	**Q**LQK**V**H**F**
	BER organizer	(h) XRCC1	≠
Nucleotide excision repair (NER)	Endonuclease	(h) XPG	≠
DNA repair/DNA damage tolerance	Terminal deoxy-nucleotidyl-transferase	(h) Tdt	≠
	Poly (ADP-ribose) polymerase	(h) PARP-1	**Q**D**LI**KMI**F***
	Helicases	(h) WRN (h) BLM (h) RECCQ5 (sc) Rrm3	**Q**WK**LL**RD **Q**QR**V**KD**FF* **Q**N**LI**RH**FF* **Q**QT**L**SS**FF**
	ATPase	(sc) Mgs1	?
Chromatin assembly	Histone chaperone	(sc) Cac1	**Q**SR**I**GN**FF**
Epigenetic inheritance and chromatin remodelling	Chromatin remodelling factor	(h) WSTF	**Q**DE**I**AED**Y*** **Q**D**II**HSI**H*** **Q**AS**VI**KK**F*** **Q**C**LV**ALL**H***
	DNA methylase	(h) DNMT1	**Q**TT**I**TS**HF**
	Histone deacetylase	(h) HDAC1	?
	Histone acetyltransferase	(h) p300	?

Sister chromatid cohesion	Acetyltransferases	(sc) Eco1 (h) ESCO1/2	**Q**SK**L**QVN **Q**L**II**DAGQ*
	Helicase	(sc) Chl1	**Q**DS**I**NSE**H***
	Clamp loader	(h) Ctf18	?
Cell cycle control and survival	Kinases	(h) CDK2 (h) EGFR	≠ ?
	CDK inhibitors	(h) p21 (h) p57 (xp) Xic1	**Q**TSMTD**FY** SGP**LI**SD**FF** TTP**I**TD**YF**
	CDK activator	(h) Cyclin D1	?
	Cell growth inhibiting factor	(h) MCL1	**Q**RNHETA**F**
	Cell growth promoting factor	(h) p15	**Q**KG**I**GE**FF**
	Phosphatase	(h) Cdc25	?
	Pro-apoptotic factor	(h) ING1b (h) Gadd45 (h) MyD118 (h) CR6 (h) p53	**Q**LH**LV**N**Y**V ? ? ? **Q**ETFSDL**W**
	E3 ubiquitin ligase	(h) MDM2	**Q**M**IVL**T**YF**
Transcription	Transcription factor	(h) RARα	≠
Transposition	Transposase	(d) Pogo	**Q**KK**I**TD**YF**

¶ Example of PIP-box sequences from different organisms.
§ Due to the wealth of literature on PCNA binding proteins, a list of quotes can be retrieved at the following reference (Moldovan et al., 2007).
* Binding through this PIP-box has not been confirmed experimentally.
‡ PIP-box consensus sequences are marked in bold red.
# Rfc4 binds PCNA through a mechanism that is different from PIP-box proteins.
≠: Upon inspection, a PIP-box sequence could not be identified.
?: PIP-box and interaction between PIP-box and PCNA has not been characterised.
(h): Homo sapiens, (sc): *Saccharomyces cerevisiae*; (sp): *Schizosaccharomyces pombe*; (d): *Drosophila melongaster*; (xp): *Xenopus laevis*; (m): *Mus musculus*; (g): *Gallus gallus*.

1.4. A general overview of chromatin assembly

The duplication of chromatin structures involves three major reactions. First, pre-existing (also known as parental) nucleosomes located ahead of replication forks are transiently disrupted by the advancing replication machinery. Second, parental histones are rapidly transferred onto either sister chromatids behind replication forks to regenerate nucleosomes. Third, the gaps in chromatin structure created by DNA duplication are filled in by the formation of new nucleosomes onto nascent DNA. Nucleosomes that are disrupted throughout the genome must be efficiently and accurately reassembled onto the duplicated DNA to propagate epigenetic information contained in modifications of parental histones. In addition, epigenetic marks must be copied from the parental histones to the new histones. Defects in these mechanisms can have detrimental consequences for the faithful transmission of epigenetic information from parental cells to daughter cells (Groth, 2007).

1.4.1. Replication-coupled chromatin assembly

During DNA replication, the progressing replication fork triggers a localised disassembly of parental nucleosomes that allows the replication machinery to access DNA despite the nucleosomal barrier (Gasser, 1996). Although the molecular basis behind parental nucleosome disruption is poorly understood, mounting evidence argues that this process is facilitated by histone chaperones. Histone chaperones could act as transient histone acceptors that promote the transfer of parental histones, along with their epigenetic post-translational modifications, onto the nascent sister chromatids. Several lines of biochemical and genetic evidences have established a role for the evolutionary conserved H2A/H2B histone chaperone complex FACT in promoting DNA replication and transcription through nucleosomes (Formosa, 2008). Importantly, FACT was recently shown to co-purify with MCM proteins from both yeast and human cells (Gambus, 2006;

Tan, 2006). The function of MCM proteins in unwinding DNA in front of the replication fork places FACT in a key position in mediating nucleosome disruption and reassembly (Gambus, 2006; Takahashi, 2005).

Behind the replication fork, nucleosomes are very rapidly and efficiently assembled as soon as approximately 150-400 bps of DNA emerges out of the replisome (Gasser, 1996; Sogo, 1986). The transfer of parental core histones onto either of the two nascent chromatids accounts for half of the nucleosomes that are assembled. This process, known as parental histone segregation, likely serves to propagate histone modification-based epigenetic information from parental histones to newly synthesized histones. The other half of the nucleosomes is assembled from newly synthesized and acetylated histones in a process known as *de novo* nucleosome assembly (often referred to as replication-coupled nucleosome assembly). This process occurs in a stepwise mechanism in which a tetramer of histones H3 and H4 is initially deposited onto DNA and two H2A-H2B heterodimers are added subsequently to complete the nucleosome. Parental and newly synthesized histones are distributed randomly between the two sister chromatids (Jackson, 1990; Sogo, 1986). Following this, formation of mature chromatin is achieved through a number of steps including nucleosome spacing and recruitment of histone H1 and non-histone DNA binding proteins (Adkins, 2004; Brown, 2003; Corona, 2004).

In addition to being coupled to cell cycle progression, histone synthesis is tightly regulated and balanced with the rate of DNA synthesis. Core histones are synthesized during S-phase and exit from S-phase results in a rapid decrease in histone mRNA levels (Marzluff, 2002). In *Saccharomyces cerevisiae*, overexpression of pairs of H2A/H2B or H3/H4 genes leads to frequent mitotic chromosome loss, indicating that the stoichiometry between the histone pairs needs to be carefully controlled (Meeks-Wagner, 1986). Assembly of histones into nucleosomes is essential for viability and failure to synthesize core histones and assemble them into nucleosomes during DNA replication causes cells to arrest in G2 and ultimately leads to cell death (Han, 1987; Kim, 1988). Even slight defects

in nucleosome assembly lead to genomic instability and sensitivity to DNA damaging agents (Emili, 2001; Game, 1999; Myung, 2003; Tyler, 1999). As stated earlier, several pathways of DNA repair necessitate synthesis of a short stretch of DNA and *de novo* nucleosome assembly coupled to the DNA synthesis step also appears to play an important role in restoring chromatin structure following DNA damage repair (Green, 2002; Lewis, 2005; Lichten, 1995). Finally, *de novo* nucleosome assembly also occurs during transcription (Lewis, 2005; Lichten, 1995; Maldonado, 1999; Tyler, 1999).

1.4.2. Histone chaperone

Biochemical studies of histone-DNA interactions have provided key insights into the mechanism of *de novo* nucleosome assembly (Loyola, 2004; Tyler, 2002). Histones are highly enriched in positively charged basic amino acids, such as lysine and arginine. As a result, core histones have an intrinsically high affinity for the negatively charged phosphate groups of DNA. When mixed together *in vitro* at physiological ionic strength, histone and DNA associate together non-specifically and form insoluble aggregates which could potentially be detrimental to many aspects of cell physiology if they were allowed to occur *in vivo*. The assembly of newly synthesized histones into nucleosomes is mediated by histone chaperones which deposit histones at sites of DNA synthesis (DNA replication or repair). Histone chaperones generally contain acidic domains that interact with histones and shield part of their positive charge from DNA, allowing chromatin assembly to occur in a regulated and ordered manner by avoiding the formation of histone-DNA aggregates *in vivo*.

1.4.2.1. CAF-1

To date, chromatin assembly factor 1 (CAF-1) is the only known histone chaperone with the unique ability to couple nucleosome assembly with DNA synthesis. Accordingly,

CAF-1 function is tightly coupled with DNA replication and DNA repair. CAF-1 was originally identified as a factor that supported chromatin assembly during *in vitro* DNA replication of the simian virus 40 (SV40) genome (Smith, 1989). Consistent with its role in *de novo* nucleosome assembly, CAF-1 is associated with newly synthesized histones H3 and H4 that are acetylated on lysines 5, 8 and 12 of the histone H4 tail (Tyler, 1999; Verreault, 1996). In *Saccharomyces cerevisiae*, these modifications contribute to nucleosome assembly and cell viability (Ma, 1998). However, during *in vitro* replication of SV40, CAF-1 binding to histones H3 and H4 does not require acetylation of either histone and its ability to promote replication-dependent nucleosome assembly is not hampered by complete truncations of H3 and H4 N-terminal tails (Shibahara, 2000). Hence, the precise function of the acetylation of new histones in the context of CAF-1 dependent nucleosome assembly remains largely unknown.

In most species, CAF-1 is composed of three polypeptide subunits corresponding to p150, p60 and p48 in humans and Cac1, Cac2 and Cac3 in budding yeast (Figure 1.5a). In addition, functional homologues have also been identified in *Schizosaccharomyces pombe*, *Arabidopsis thaliana*, *Drosophila melanogaster, Mus musculus* and *Xenopus laevis*, highlighting the evolutionary conservation of CAF-1 (Ridgway, 2000). All three CAF-1 subunits associate with histones *in vitro* (Kaufman, 1995; Shibahara, 2000; Verreault, 1996). In contrast to yeast, CAF-1 is essential for viability in human cells (Hoek, 2003; Nabatiyan, 2004). This difference is most likely due to the existence of another redundant histone H3/H4 chaperone in yeast, the Hir protein complex (Kaufman, 1998; Sharp, 2001). Although the four subunits of the Hir complex also exist in human cells, it appears to be dedicated to replication-independent nucleosome assembly whereas, based on genetic evidence, the yeast Hir complex can contribute to replication-coupled nucleosome assembly (Kaufman, 1998; Sharp, 2001). In human and mouse cells, CAF-1 localizes to essentially all replication foci during S-phase and deposits histones H3 and H4 behind replication forks in nascent euchromatic and heterochromatic DNA (Krude, 1995; Marheineke, 1998; Smith, 1989; Taddei, 1999; Verreault, 1996).

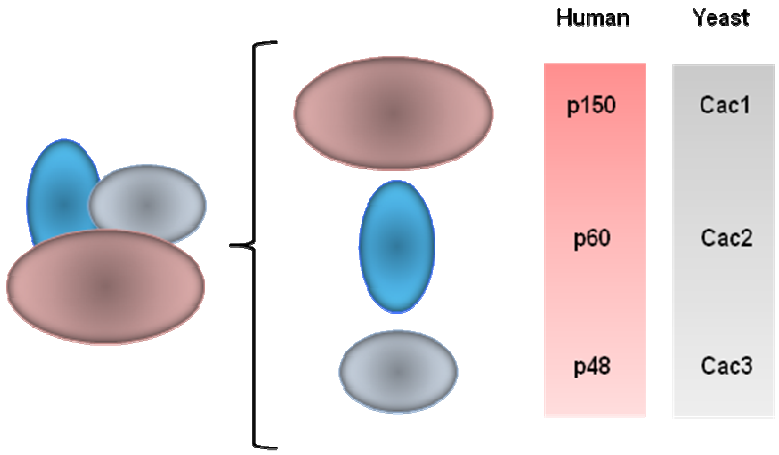
During both DNA replication and repair, yeast and human CAF-1 are targeted to sites of DNA synthesis through a direct interaction between their largest subunits (Cac1 and p150, respectively) and PCNA (Green, 2003; Krawitz, 2002; Moggs, 2000; Sharp, 2001; Sharp, 2002; Shibahara, 1999; Zhang, 2000). As illustrated in figure 1.5b, the human p150 subunit contains several domains that are important for CAF-1 function. The p150 subunit of CAF-1 contains two fundamentally distinct PIP-boxes termed PIP1 and PIP2, which are partially conserved through evolution and mediate direct interactions between CAF-1 PCNA (Ben-Shahar, 2009). PIP2 is different from canonical PIPs since the highly conserved glutamine residue which, based on several crystal structures is involved in several interactions with PCNA, is replaced by a lysine (Figure 1.5c). Although PIP2 has low affinity for PCNA, it is nonetheless essential for replication-coupled nucleosome assembly in the SV40 system as well as for targeting CAF-1 to DNA replication foci during early and mid S-phase. Surprisingly, although PIP1 binds strongly to PCNA, it is dispensable for nucleosome assembly during DNA synthesis *in vitro*. An appealing hypothesis to explain these findings is that the intrinsic features of the CAF-1 PIPs might enable CAF-1 to deposit histones on replicating DNA without interfering with other PCNA-binding proteins during *in vivo* DNA replication.

The acidic KER and ED domains of p150 were shown to mediate p150 interaction with newly synthesized and acetylated histones H3 and H4 (Kaufman, 1995). The p150 protein also contains a PEST region, a domain rich in proline (P), glutamic acid (E), serine (S) and threonine (T) residues. The presence of PEST domains in other proteins has been correlated with cell cycle dependent degradation (Rogers, 1986) and was therefore proposed to be involved in CAF-1 turnover (Kaufman, 1995), although the function of this turnover is currently unknown. The PXVXL motif in the N-terminal domain of p150 mediates CAF-1 interaction with the heterochromatin protein 1 (HP1) family of proteins, which are structural components of pericentric and other forms of heterochromatin (Murzina, 1999; Thiru, 2004). An interaction between CAF-1 and HP1 is essential for delivery of HP1

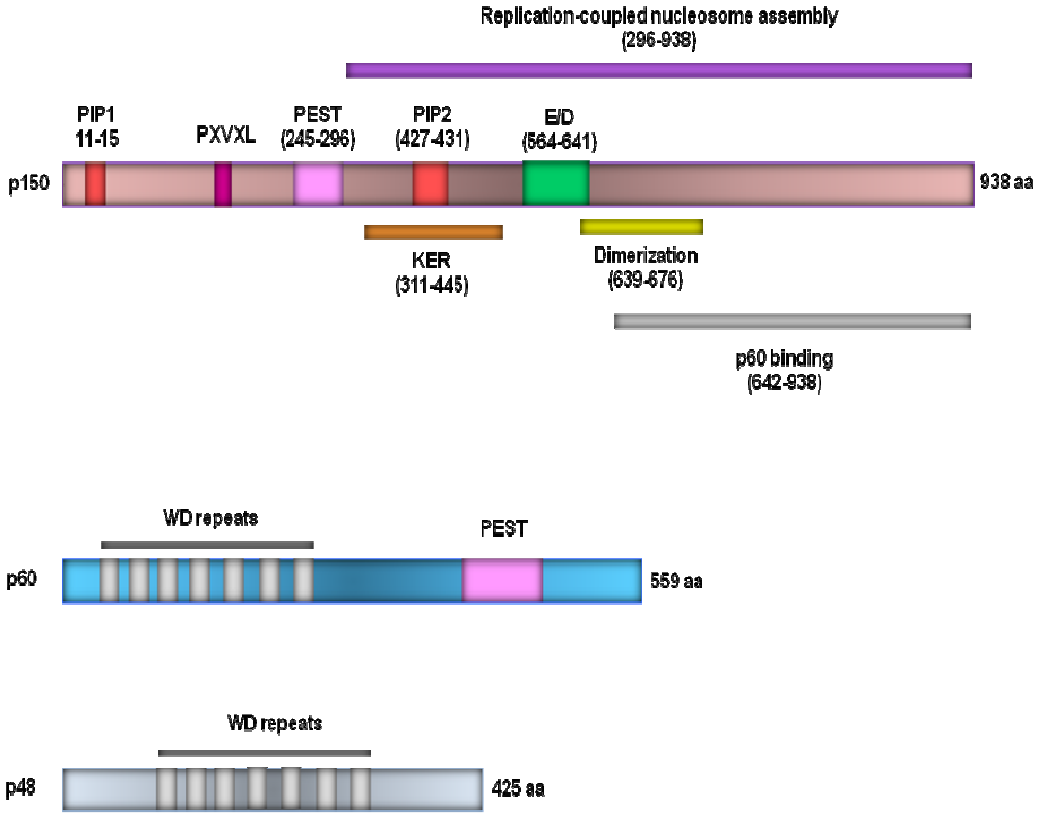
molecules to sites of heterochromatin replication (Quivy, 2004). In budding yeast, HP1 does not exist. However the yeast homologue of human p150, known as Cac1, interacts with the heterochromatin related protein, silent information regulator 1 (Sir1) and the MYST family HAT, something about silencing 2 (Sas2), which deactylates H4K16 to create a flexible boundary between euchromatin and heterochromatin at telomeres (Meijsing, 2001; Sharp, 2003). The first 296 amino acids of human p150 are not required for PCNA-dependent nucleosome assembly during SV40 DNA replication *in vitro* (Kaufman, 1995). The C-terminal region of p150 is however crucial for binding to the p60 subunit and is therefore necessary for chromatin assembly *in vitro* (Kaufman, 1995).

In higher eukaryotes, a unique region of 36 residues, in the C-terminal region of p150 is critical for p150 homodimerization (Quivy, 2001). Deletion of this region abolishes the formation of p150 dimers and abrogates its ability to promote chromatin assembly *in vitro*. Furthermore, expression of human p150 acts in a dominant negative fashion in early *Xenopus laevis* embryos. The presence of both human and *Xenopus* p150 in the same cells results in severe cell cycle progression defects during the rapid cleavage stages of *Xenopus* laevis early development and eventually leads to early embryonic lethality. However, expression of a form of human p150 containing a deletion of the dimerization domain does not affect *Xenopus laevis* embryo development. In a *Xenopus* laevis extract that carries DNA repair-coupled nucleosome assembly, human p150 interferes with *Xenopus* laevis p150 in a dimerization domain-dependent manner (Quivy, 2001). In human cells, expression of a dominant negative inhibitor of CAF-1 activity triggers DNA damage and activates the intra S-phase checkpoint, which leads to DNA replication and S-phase arrest (Ye, 2001). These results highlight the importance of CAF-1 in replication-coupled nucleosome assembly *in vivo* and provide insights into its dimerization properties and the mechanistic regulation of its activity in higher eukaryotes.

a)


Human
Yeast
p150
Cac1
p60
Cac2
p48
Cac3

b)


Replication-coupled nucleosome assembly
(296-938)
PIP1
11-15
PXVXL
PEST
(245-296)
PIP2
(427-431)
E/D
(564-641)
p150
938 aa
KER
(311-445)
Dimerization
(639-676)
p60 binding
(642-938)
WD repeats
PEST
p60
559 aa
WD repeats
p48
425 aa

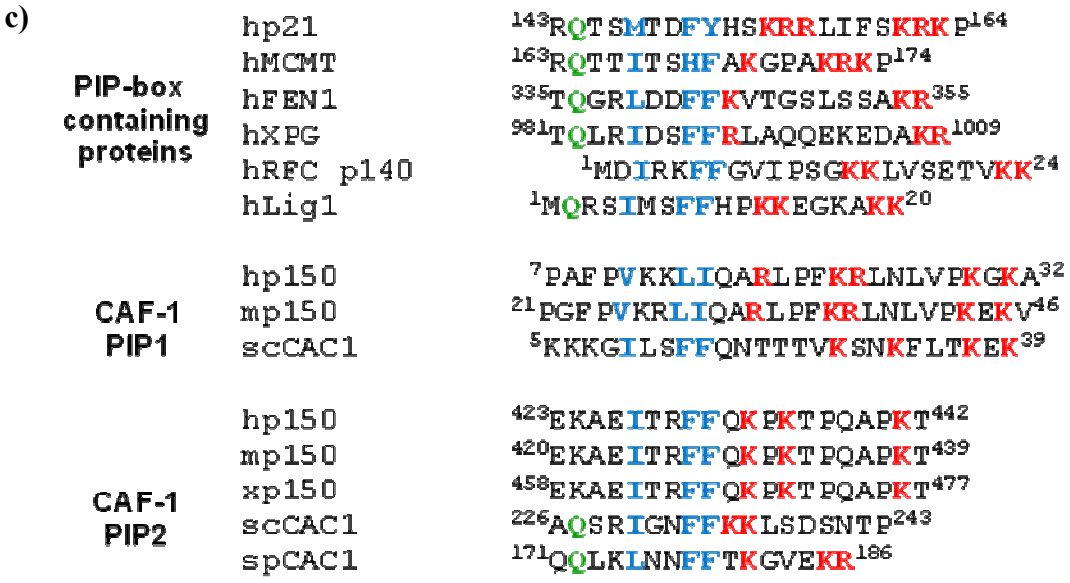


Figure 1.5 CAF-1 is composed of three evolutionary conserved subunits. a) The different subunits of CAF-1 are schematically illustrated and their respective names are indicated. **b)** The domain organization of the three CAF-1 subunits in human is shown. In the p150 subunit, the PIP boxes (PIP1 and PIP2) are in red. The PXVXL consensus motif for binding to HP1 is in purple. The KER and ED domains, rich in charged residues, are in orange and green, respectively. p150 can dimerise through residues 639-676 in yellow and binding to p60 is mediated through residues 642-938 in grey. The region required for chromatin assembly *in vitro* spans residues 296-938 represented in purple. The PEST domains in p150 and p60, which are thought to be important for protein turnover, are illustrated in pink. The 7 WD repeats of p60 and p48 are represented in light grey. **c)** Amino acid sequence alignment of PIP-box-containing PCNA binding proteins with the PIP-box sequences from CAF-1. The CAF-1 PIP-box sequence is evolutionarily conserved. h: *Homo sapiens*; m: *Mus musculus*; sc. *Saccharomyces cerevisiae*; x: *Xenopus laevis*; sp: *Schizosaccharomyces pombe*. The conserved hydrophobic and aromatic residues are highlighted in blue, the conserved glutamine is in green and basic residues are highlighted in red.

The middle subunit of CAF-1, p60, is a member of the WD repeat family of proteins. The WD repeats are regions of approximately 40 amino acid residues containing a conserved tryptophan-aspartate (WD) di-peptide. WD repeats are generally involved in mediating protein-protein interactions (Neer, 1994; Smith, 1999b). The p60 subunit has a total of seven conserved WD repeats and, like p150, also contains a PEST domain

(Kaufman, 1998, 1995). Similarly to p60, the evolutionary conserved p48 subunit of CAF-1 is also a member of the WD repeat protein family and is able to accommodate many interactions with various proteins. Both p48, and another highly related protein known as p46, are often referred to as RbAp48 and RbAp46 because they were initially discovered as Retinoblastoma (Rb) binding proteins *in vitro* (Qian, 1995; Qian, 1993). The RbAp46/48 proteins bind directly to helix 1 in the histone fold of histone H4. Helix 1 binds directly to nucleosomal DNA and adopts a dramatically different conformation when bound to RbAp46/48 than when H4 is part of nucleosome (Luger, 1997; Murzina, 2008). Interestingly, in a broad range of species, RbAp46/48 are part of many multi-protein complexes involved in histone modification, nucleosome assembly and ATP-dependent nucleosome remodelling (Barak, 2003; Cao, 2002; Kuzmichev, 2002; Taunton, 1996; Zhang, 1997; Zhang, 1999). For instance, in human cells RbAp46, but not the highly related RbAp48, is the histone H4 binding subunit of the Hat1 complex, a type-B histone acetyltransferase that acetylates newly synthesized H4 to promote chromatin assembly during DNA replication (Verreault, 1998). Both RbAp46 and p48 are intrinsic subunits of enzymes that deacetylate histones and contain HDAC1 and HDAC2 as catalytic subunits. In addition, RbAp46/48 are found in the ATP-dependent nucleosome remodelling complex NuRF, the nucleosome remodelling and histone deacetylase complex NuRD, the Sin3 deacetylase complex and the Polycomb repressor complex PRC2 that promotes epigenetic silencing by methylating histone H3. The Rb tumour suppressor also functions in transcriptional repression by recruiting HDACs that contain RbAp46/48 (Nicolas, 2001). This has led to the proposal that RbAp46/48 generally act as histone binding subunits and simply target ATP-dependent chromatin remodelers, and enzymes that acetylate, deacetylate or methylate histones to their histone substrates.

The p150 and p60 subunits of CAF-1 are both phosphorylated *in vivo* (Keller, 2000; Marheineke, 1998; Martini, 1998; Smith, 1991). Human CAF-1 isolated from mitotic cells does not support replication-dependent nucleosome assembly *in vitro* and this correlates with hyperphosphorylation of the p60 subunit. *In vitro* dephosphorylation of mitotic CAF-1

restores its ability to promote DNA synthesis-coupled nucleosome assembly (Keller, 2000). Active CAF-1 can be isolated from G1, S and G2 phase cell nuclei where the p60 subunit is found in moderately phosphorylated and hypophosphorylated forms. Both p150 and p60 are phosphorylated by CDK2 and dephosphorylated by protein phosphatase 1 (PP1) *in vitro* (Gerard, 2006; Keller, 2000). Experiments using synthetic and natural inhibitors of either CDKs or protein phosphatases suggest that the G1/S-phase specific cyclin/CDK2 kinases, which are required for the initiation of DNA replication in isolated cell nuclei (Krude, 2000; Krude, 1997), are also required for CAF-1 dependent nucleosome assembly coupled to DNA synthesis *in vitro*. CAF-1 activity during DNA replication in human cell extracts is also dependent on PP1 activity (Keller, 2000). These results provide functional evidence of the requirement for continuous cycles of reversible protein phosphorylation events for efficient CAF-1 activity *in vitro*. In addition, *in vitro* data have also shown that the p150 subunit of CAF-1 interacts directly with and is phosphorylated by Cdc7-Dbf4 (Gerard, 2006), a kinase that is essential for initiation of DNA replication (Bell, 2002). Phosphorylation by Cdc7-Dbf4 has been reported to disrupt the p150 dimer, stabilizing the monomeric form of p150, and promoting its interaction with PCNA. Thus, according to this hypothesis, the phosphorylation of p150 would target CAF-1 to sites of DNA synthesis through PCNA binding. This phosphorylation might contribute to coordinate CAF-1 function with other PCNA-dependent processes when DNA replication is proceeding.

Yeast genetic assays using reporter genes to monitor gene silencing in sub-telomeric chromosomal regions, the silent mating type loci and the rDNA locus revealed that cells lacking any of the CAF-1 subunits have subtle defects in heterochromatin-mediated transcriptional silencing (Enomoto, 1998, 1997; Kaufman, 1997; Monson, 1997; Smith, 1999a). The residual silencing observed in *cac*Δ mutant strains suggests that CAF-1 is partially redundant with other factors that participate in heterochromatin formation, such as the Hir proteins and Asf1. Significantly, yeast cells lacking CAF-1 and either Hir proteins or Asf1 exhibit more pronounced growth and silencing defects than cells lacking CAF-1 only (Kaufman, 1998; Qian, 1998; Sharp, 2002; Singer, 1998; Tyler, 1999). Insights into

the *in vivo* importance of CAF-1 in the maintenance of gene silencing and the inheritance of epigenetic chromosomal states were also obtained in higher eukaryotes. The expression of a dominant negative truncated form of mammalian CAF-1 increases by more than 500-fold the frequency at which transcriptional gene silencing can revert to an active state. Endonuclease protection assays further revealed that reversion to the active state correlates with increased accessibility to chromatin structure (Tchenio, 2001). Moreover, studies in *Saccharomyces cerevisiae* demonstrated that CAF-1 dependent transcriptional silencing is strongly dependent on PCNA. Two classes of hypomorphic mutations in PCNA showed reduced silencing of reporter genes integrated near telomeres (Zhang, 2000). One class of PCNA mutations crippled silencing to the same extent in the absence or presence of CAF-1 (Sharp, 2001), indicating that those PCNA mutations exert their effects by impairing the ability of CAF-1 to promote silencing. These silencing defective PCNA mutants showed reduced binding to CAF-1 *in vitro* and significantly impaired the association of Cac1 to chromatin *in vivo*. In contrast, a second group of PCNA mutations synergistically reduced silencing in cells lacking CAF-1, implying that those mutations impaired a pathway that functions in parallel with CAF-1 to promote silencing. Consistent with this, the residual (CAF-1 independent) silencing observed in these PCNA mutants, was dependent on Asf1 and Hir1, which functionally overlap with CAF-1 to promote gene silencing in a PCNA-dependent manner (Sharp, 2001).

CAF-1 is also involved in the response to DNA damage. Eukaryotic cells need to remodel chromatin structure in order to allow repair machineries to access potentially lethal DNA lesions. Following repair of UV-induced DNA lesions, chromatin structure needs to be restored and, therefore, new histone H3 and H4 molecules are incorporated at sites of DNA damage during repair in a process that is dependent on both NER and CAF-1 (Polo, 2006). Accordingly, both *in vitro* and *in vivo* experiments have established a role for CAF-1 in promoting chromatin assembly during the processes of NER and single strand break (SSB) repair (Gaillard, 1996; Green, 2003; Martini, 1998; Okano, 2003). Remarkably, phosphorylated CAF-1 is also recruited to chromatin following UV irradiation of human

cells outside of S-phase (Martini, 1998). This is consistent with the observation that CAF-1 isolated from G1 and G2 is able to promote nucleosome assembly *in vitro* (Keller, 2000; Marheineke, 1998). Importantly, CAF-1 is recruited to sites of UV-induced DNA damage in a PCNA-dependent manner and depletion of PCNA from a cell-free system during repair disrupts CAF-1-dependent chromatin assembly *in vitro* (Moggs, 2000). Genetic studies performed in budding yeast revealed that *cac*Δ mutant strains display a mild sensitivity to UV-irradiation, highlighting the necessity for CAF-1 in survival of UV-induced DNA damage (Game, 1999; Kaufman, 1997). Interestingly, deletion of *CAC1* increases the UV sensitivity of nucleotide excision repair mutants (*rad1*Δ and *rad14*Δ) and recombinational repair mutants (*rad51*Δ), but does not increases the UV sensitivity of *rad6*Δ or *rad18*Δ mutant strains (Game, 1999). Both Rad6 and Rad18 are involved in tolerance of UV-induced DNA lesions because they are necessary for error-free and error-prone DNA lesion bypass (Lawrence, 1994; Prakash, 1993). Thus, one proposed hypothesis is that CAF-1 promotes rapid nucleosome assembly during the Rad6/Rad18-dependent DNA damage bypass of UV-induced lesions.

In a manner that is dependent upon PCNA, CAF-1 also plays a role in nucleosome assembly following repair of DNA double-strand breaks by either homologous recombination or non-homologous DNA end joining (Linger, 2005; Nabatiyan, 2006). Yeast CAF-1 mutants display mild sensitivity to a variety of double-strand DNA damaging agents. In yeast, a systematic screen for viable gene deletions revealed that *cac2*Δ mutants are sensitive to the alkylating agent MMS. Likewise, mutations in the genes encoding *Arabidopsis thaliana* subunits of CAF-1 (*fas1* and *fas2*) also confer MMS sensitivity (Takeda, 2004). Finally, deletion of genes encoding *Saccharomyces cerevisiae* CAF-1 lead to an increase in the frequency of both spontaneous and MMS-induced gross chromosomal rearrangements (GCRs), which are initiated by double-strand DNA lesions that likely occur more frequently in cells lacking CAF-1 (Myung, 2003).

1.4.2.2. The Hir proteins

In *Saccharomyces cerevisiae*, the histone regulator (*HIR*) genes *HIR1*, *HIR2*, *HIR3* and *HPC2* encode proteins that were originally identified as transcriptional repressors of histone genes (Osley, 1987; Sherwood, 1991; Xu, 1992). Indeed, the *Saccharomyces cerevisiae* Hir proteins tightly regulate histone gene transcription during cell cycle progression by repressing transcription of six of the eight histone genes (*HTA1-HTB1*, *HHT1-HHF1* and *HHT2-HHF2*) outside of the G1/S phase, when the demand for histone synthesis is low (Osley, 1987). In addition, Hir proteins are required to repress transcription from the *HTA2-HTB2* locus when multiple copies of *HTA-HTB* genes are present or in response to inhibition of DNA replication (Moran, 1990; Osley, 1987; Recht, 1996; Sherwood, 1991; Sherwood, 1993). Although the molecular mechanism by which Hir proteins repress histone gene transcription is not clear, it likely involves the modulation of specialized chromatin structures that regulate accessibility to histone gene promoters (Dimova, 1999; Prochasson, 2005).

In yeast, mutations in *HIR* genes have essentially no effect on silencing of reporter genes integrated near telomeres or the silent mating type loci. In addition, cell proliferation is not affected when any of the *HIR* genes is mutated and the chromatin structure of *HIR* single mutants is essentially the same as in wild-type cells (Kaufman, 1998). However, cells lacking both Hir proteins and CAF-1 exhibit a more accessible chromatin structure at a number of chromosomal loci as well as severe silencing defects, elevated rates of Ty1 retrotransposition and sensitivity to the DNA alkylating agent MMS (Kaufman, 1998; Qian, 1998). *cac*Δ *hir*Δ double mutant cells also exhibit a prolonged delay through G2/M phase of the cell cycle. This delay arises from alterations in peri-centromeric chromatin structures that impair kinetochore function and activate the mitotic spindle assembly checkpoint (Sharp, 2002). Even though the function of the spindle checkpoint is to minimize misegregation, *cac*Δ *hir*Δ double mutants experience a dramatic increase in the frequency of mitotic chromosome loss (Sharp, 2002). The synergistic effect of mutating both Hir

proteins and CAF-1 demonstrates that, in the absence of CAF-1, the Hir proteins become important for efficient chromatin assembly which, in turn, is important for heterochromatin-mediated transcriptional silencing, pericentric chromatin function and chromosome segregation. Thus, in *Saccharomyces cerevisiae*, CAF-1 and Hir proteins are genetically redundant. Intriguingly, although none of the Hir proteins contains any obvious PIP-box sequences, specific mutations in PCNA block the contribution of Hir proteins to silencing of sub-telomeric reporter genes (Sharp, 2001), suggesting that in yeast, Hir proteins may promote nucleosome assembly in a PCNA-dependent manner.

Several lines of evidence suggest that, like CAF-1, the Hir proteins bind histones H3 and H4 and promote their assembly into nucleosomes (Kaufman, 1998; Sharp, 2001). Furthermore, in higher eukaryotes, HIRA, a homologue of yeast Hir1 and Hir2, also binds to core histones (Lorain, 1998). Sequence comparison of mammalian HIRA, yeast Hir1 and Hir2, and the subunits of CAF-1 Cac2 (yeast) and p60 in higher eukaryotes showed that they form a distinct sub-family of WD-repeat proteins that are highly related to each other (Kaufman, 1998; Lamour, 1995; Wilming, 1997). Furthermore, HIRA forms a complex with UBN1 and Cabin 1, the homologues of yeast Hpc2 and Hir3, respectively (Banumathy, 2009; Tagami et al., 2004). Interestingly, HIRA and UBN1 are needed for the formation of senescence-associated heterochromatin foci (SAHF), where chromatin undergoes dramatic remodelling through the formation of domains of facultative heterochromatin (Banumathy, 2009). Taken together, these observations suggest that HIRA might have some role in chromatin assembly and histone gene regulation.

The role of Hir proteins in replication-dependent nucleosome assembly is however not conserved in human cells. Unlike CAF-1, HIRA does not localise to DNA replication foci and promotes nucleosome assembly independently of DNA synthesis in *Xenopus laevis* extracts (Ray-Gallet, 2003). In mouse, HIRA is essential for proper embryonic development (Roberts, 2002). Overexpression of HIRA leads to transcriptional repression of histone genes and inhibits DNA synthesis causing human cells to arrest during S-phase

of the cell cycle (Hall, 2001; Nelson, 2002). In human cells, HIRA is an *in vivo* substrate of the S-phase specific cyclin-Cdk2 kinase (cyclin A or E) (Hall, 2001). Consistent with the notion that HIRA may not function in nucleosome assembly during S-phase, HIRA associates exclusively *in vivo* with histone H3 and H4 dimers that specifically contain the histone variant H3.3 (Loppin, 2005; Nakayama, 2007; Tagami, 2004). Unlike H3.1 and H3.2, H3.3 is not only expressed during S-phase, but also in the G1 and G2 phases of the cell cycle in Chinese hamster ovary cells (Wu, 1982).The replication-independent chromatin assembly pathway promotes deposition of histone H3.3/H4 to transcriptionally active loci throughout the cell cycle as a mechanism to fill in gaps in chromatin left after the passage of the transcription machinery and nucleosome displacement (Ahmad, 2002a, b; McKittrick, 2004; Ray-Gallet, 2003; Tagami, 2004). In addition, replication-independent deposition of H3.3 may serve to activate gene expression by replacing H3 histone molecules that are methylated at lysine 9 (H3K9me), a metabolically stable modification that is associated with heterochromatin formation and gene repression (Ahmad, 2002a, b; Nakayama, 2007).

Taken together, these current data indicate that in yeast, where only one form of histone H3 exists, both CAF-1 and Hir proteins function in replication-dependent nucleosome assembly. The genetic evidence does not however rule out the possibility that the yeast Hir complex may also promote nucleosome assembly in a replication-independent manner. Although the yeast and mammalian Hir complexes are both involved in aspects of chromatin assembly, there has been significant divergence in the specialized role of HIRA in higher eukaryotes. While CAF-1 is still involved in replication-dependent nucleosome assembly, HIRA has become specialized in replication-independent nucleosome assembly and specifically associates with the H3.3 variant. This loss of redundancy between HIRA and CAF-1 in higher eukaryotes is consistent with the severe phenotypes that are observed following genetic perturbation of CAF-1 induction (Hoek, 2003; Nabatiyan, 2004; Roberts, 2002; Ye, 2001).

1.5. Mass spectrometry as a tool to study protein phosphorylation

The major aim of the research conducted during the course of this Ph.D. thesis was to understand the regulatory mechanisms that control the process of chromatin assembly during DNA replication. More specifically, the focus of this thesis is centered on the identification as well as the analysis of post-translational modifications, such as phosphorylation, of factors implicated in the process of nucleosome formation and chromatin assembly. To achieve these ends, we employ mass spectrometry as a tool to characterize proteins of interest. In recent years, mass spectrometry has become a powerful tool used in molecular and cell biology. More specifically, mass spectrometry is currently the only available method that allows the systematic characterization of molecular alterations at the protein level (Walther, 2010). The following sections describe the overall workflow of a typical mass spectrometry (MS) experiment (Figure 1.6).

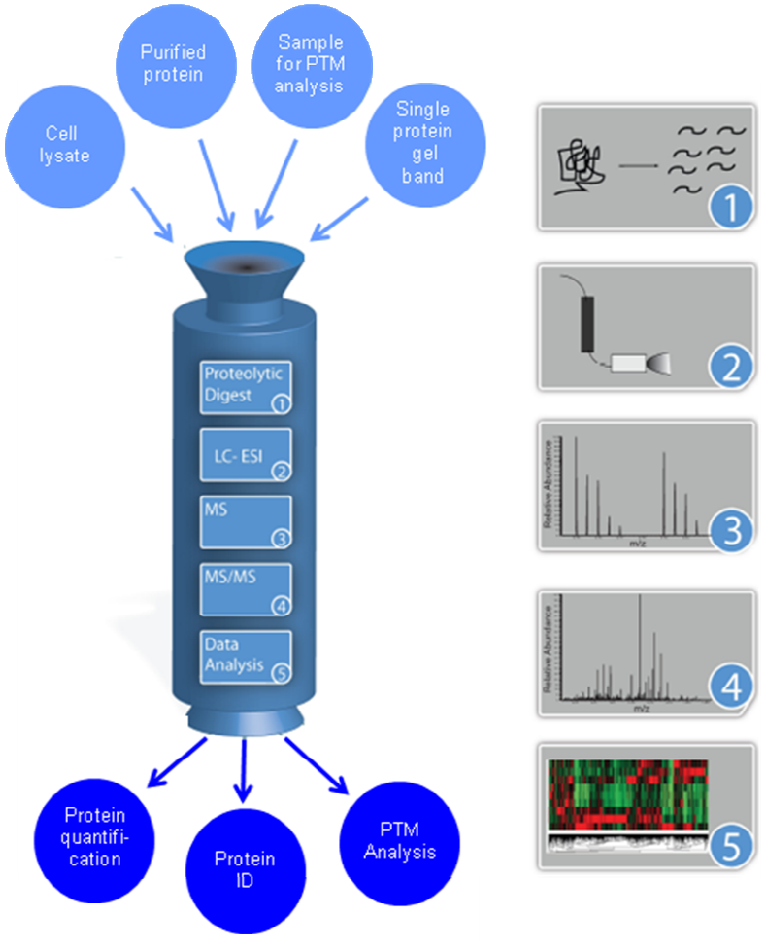


Figure 1.6 General workflow of a typical mass spectrometry experiment. Samples can be prepared differently according to the specific application. The typical workflow consists of several steps. First, proteolytic digestion (1); second, separation of peptides by liquid chromatography and electrospray ionization (ESI) (2); third, MS analysis (3); fourth, peptide fragmentation and MS/MS analysis (4); fifth, data analysis including protein identification, quantification and characterization of post-translational modifications (5) (Walther, 2010).

1.5.1. Mass spectrometry principles

1.5.1.1. Sample preparation

One of the most prominent challenges in mass spectrometry is to resolve and detect peptides in complex mixtures. In fact, the quality of the MS analysis is crucially dependent on the amount of sample introduced into the mass spectrometer as well as on its level of purity (Granvogl, 2007; Washburn, 2001). This is particularly true for low abundant peptides, such as phosphorylated peptides. As a result, mass spectrometry, although being a highly robust and sensitive method for the analysis of proteins and peptides, often requires simplification of the sample mixture (Granvogl, 2007).

The most widely used method to separate proteins is sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which can be performed in one (1D) or two (2D) dimensions. When phosphorylation sites are to be mapped on a specific protein, immunoprecipitation using a set of expression vectors yielding FLAG- or TAP-tagged proteins can be effectively used in conjunction with SDS-PAGE to obtain thorough sequence coverage of the protein in question and identify a number of *in vivo* phosphorylation sites (Goshe, 2006).

In 1D-electrophoresis, SDS molecules bind and impart a negative charge (through the SDS sulfate group) to the proteins. This is done in a constant proportion to the molecular mass of the protein. Upon application of an electric field, proteins will separate based on their ability to fit within the pores of the polyacrylamide matrix. Short proteins will migrate faster than larger ones, thereby resolving proteins based on their individual molecular weights (Liebler, 2002). Occasionally, a phosphoprotein will exhibit an altered electrophoretic migration compared to its non-phosphorylated counterpart. One explanation is that the negatively charged phosphate group on the protein might disrupt uniform binding of SDS to the phosphoprotein, altering the charge density and the migration rate of the

phosphoprotein (Gafken, 2006). As a result, 1D-electrophoresis can sometimes resolve modified proteins from non-modified ones, even when the mass difference between the two isoforms is as low as 80Da. Alternatively, in 2D-electrophoresis, proteins are initially separated in the first dimension based on their isoelectric point (pI) by preparative isoelectric focusing (IEF) and then in the second dimension on the basis of their molecular weight (Hanash, 2000; Rabilloud, 2002). Multi-dimensional gel separation of proteins is ideally suited for more complex protein mixtures and provides higher resolution for the separation of modified proteins. One example is phosphorylation, which not only alters the molecular weight of the protein but also affects its pI (Mann, 2003). Nonetheless, some problems encountered with 2D-electrophoresis have limited the use of this technique. These include large amount of sample handling, low reproducibility and the incompatibility of some proteins such as low abundance, hydrophobic, acidic and basic proteins, with the first dimension (Liebler, 2002).

Following electrophoresis, separated proteins can be visualized by use of several MS-compatible staining methods. Conventional staining techniques include fluorescent stains as well as Coomassie brilliant blue (CBB) and silver stains. The latter two are most popular due to their ease of use, simple protocols and high sensitivity. CBB can typically detect 30 ng per gel band (Candiano, 2004; Neuhoff, 1988; Wang, 2007) whereas silver staining can reach detection limits in the range of 1-10 ng (Graham, 2005; Shevchenko, 1996). In order to improve the possibility of detecting and characterizing post-translational modifications, proteins should be at least visible by CBB staining. Hence, several picomoles (or ~1μg) of the target protein should be isolated (Mann, 2003). Prior to MS analysis, stained bands of interests are excised and proteins are proteolytically digested inside the gel to generate peptides. These are subsequently extracted from the gel and analyzed by ESI-MS.

Trypsin is by far, the most commonly used protease in proteomics applications. Trypsin cleaves proteins specifically after lysine and arginine residues except when these

are followed by a proline residue (Liebler, 2002). The use of trypsin offers several advantages. First, the biological distribution of lysine and arginine residues in most proteins is such that the resulting tryptic peptides are of appropriate mass for MS analysis (Olsen, 2004). Furthermore, trypsin digestion ensures that at least one basic amino acid (either lysine or arginine), with high proton affinity, is retained within the peptide sequence. This is particularly important for peptide ionization in positive-ion mode and is ideal for sequencing by collision induced dissociation (CID) (Swaney, 2010). Finally, trypsin can be used directly to digest proteins within the gel matrix.

1.5.1.2. Chromatographic separation of peptides

Depending on the complexity of the sample, peptides can then be analyzed differently. One of the most effectively used approaches to analyse peptide mixtures is high performance liquid chromatography (HPLC) (Domon, 2006; Yates, 2004). The principle behind this system is that a very heterogeneous mixture of peptides is converted to a more homogeneous mixture, which is more easily analysed by the mass spectrometer (Liebler, 2002). In a typical MS experiment, the HPLC system is directly coupled to the mass spectrometer via the electrospray ionization (ESI) source. This setup introduces a fast and reliable way of probing complex mixtures of peptides because it allows the sample to be directly and continuously analysed by the MS instrument as it elutes from the chromatographic column. The separation of peptides is generally accomplished using reverse-phase (RP) HPLC containing C_{18} material (Shen, 2002a). In RP-HPLC, the amount of organic solvent present in the mobile phase, usually acetonitrile (ACN), is increased over time to elute peptides of progressively higher hydrophobicity. One advantage of RP-HPLC is that the mobile phase is compatible with ESI and can therefore be easily coupled to the mass spectrometer (Shen, 2002a). Typically, the HPLC separation is used in nano scale (nanoLC-MS/MS), where capillary columns are used instead of the traditional large-bore HPLC columns (Liebler, 2002). The use of long and narrow capillary columns (typically between 50-150 μm internal diameter) allows the RP-HPLC system to operate at very low

flow rates (in the order of 100-600 nL/min). This greatly improves the loading capacity, the sensitivity, the reproducibility as well as the dynamic range of LC-MS analysis (Shen, 2002b).

Finally, the use of combined separation methods is alternatively used to separate complete proteomes. In contrast to purified proteins, these generate highly complex peptide mixtures that are best analyzed by two-dimensional (2D) nanoLC-MS/MS analysis (Fournier, 2007). The 2D chromatography method generally features a strong cation exchange (SCX) column that is added prior to the RP-HPLC column. The complex sample is initially loaded onto the SCX column and eluted using a series of increasing salt concentrations to elute peptides based on their net positive charge. Each fraction is subsequently loaded onto the RP column and directly eluted and applied to the mass spectrometer using ESI (Link, 1999; Wolters, 2001). This multidimensional separation approach addresses the limited peak capacity imposed by the one-dimensional approach and increases the resolving power as well as the sensitivity of the analysis (Fournier, 2007; Liu, 2002).

1.5.1.3. Mass spectrometry instrumentation

Since its establishment more than 100 years ago, mass spectrometry has become one of the most comprehensive and versatile tool used for the characterisation of biomolecules (Cañas, 2006). In addition to their enhanced sensitivity and selectivity, MS instruments now provide rich structural information of compounds such as proteins and peptides with high accuracy and in a very short time. The principle of MS consists of separating, detecting and accurately measuring ions in the gas phase based on their mass-to-charge (*m/z*) ratio (Aebersold, 2003). Basically, mass spectrometers are composed of three essential parts. The first part is the ionization source which converts the sample into ions. The second part, the mass analyzer, resolves every *m/z* peptide ions, induces their fragmentation and generates fragment ions that are subsequently separated and also

analyzed based on their *m/z* ratio. The third part, the detector, detects the resolved ions and in combination with sophisticated softwares provides qualitative as well as quantitative data for analysis (Watson, 2007). Finally, the typical MS instrument delivers a full spectrum of precursor ions, also known as an MS spectrum, which provides data for measuring the mass (or *m/z*) and the abundance of each peptide ion, as well as a «product» ion spectrum, also known as the MS/MS spectrum. This allows the sequence of the peptide to be elucidated and post-translational modifications to be identified (Steen, 2004; Watson, 2007).

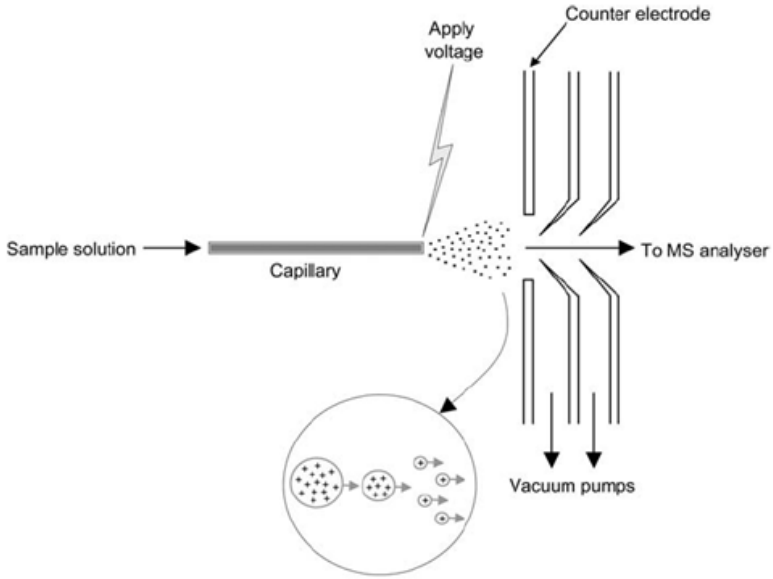
1.5.1.3.1. Electrospray ionization (ESI) source

Many types of ionization methods for proteins have been developed in mass spectrometry. These include fast atom bombardment (FAB), plasma desorption (PD) and thermospray (Cañas, 2006). However, the development of a new ionization technique called electrospray ionization (ESI) by Fenn in the late eighties proved to be one of the most important breakthrough in the field (Fenn, 1989). The versatility of various mass analyzers with ESI sources and its ability to be continuously coupled to the HPLC and the MS instrument makes ESI extensively used in biological research today (Aebersold, 2003).

ESI is characterized as a «soft» ionization technique because it is capable of ionizing large bio-molecules without fragmenting them therefore allowing intact molecular ions to be produced (Fenn, 1989). In ESI, a solution containing the analyte is introduced through a capillary column upon which a high voltage (2-5 kV) is applied. The electric field gradient between the end of the capillary and the atmospheric inlet on the instrument creates a spray of charged droplets (Taylor cone) (Figure 1.7). The charged droplets become smaller as the solvent evaporates. This desolvation process is facilitated by a heated capillary or with the assistance of an inert heated carrier gas such as nitrogen. As the droplets evaporate and become smaller, the charge density increases. Once the droplets have reached Rayleigh`s limit, which is the point where the charges within the confined small solvent area start to repel each other, a Coulombic explosion occur. This explosion

creates smaller droplets that undergo the same process until only the charged ions remains and all the solvent is removed. Once the analyte is free of solvent, it moves to the mass analyzer. This process is very fast and occurs in milliseconds.

a)



b)

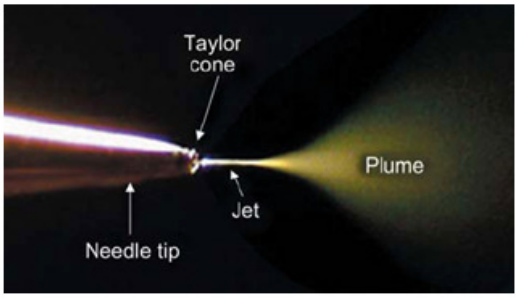


Figure 1.7 The electrospray ionization (ESI) process. a) Schematic diagram of an ESI interface**. b)** The ESI process viewed through a high-powered microscope (Lane, 2005).

ESI is unique in its ability to produce a distribution of multiply charged ions for the same species, such that the *m/z* of the ions is within the mass range of many analyzers (Cech, 2001). Finally, one important development in ESI is nano-ESI, which uses smaller flow rates (in the orders of nL/min). Nano-ESI offers higher ionization efficiency by at least two orders of magnitude than conventional ESI (Shen, 2002b; Wilm, 1996). Overall, these features provide higher sensitivity and minimize sample loss.

1.5.1.3.2. Q-q-LIT and LTQ-Orbitrap mass analyzers

In the past two decades, major technical developments have taken place in the field of mass spectrometry. These have mainly been focused on designing high performing mass spectrometers that offer three main characteristics. The first is sensitivity with the growing need for instruments that are capable of obtaining data on femtomole levels or less. The second is high resolution with instruments that can reliably distinguish peptide ions with very similar *m/z* values. The third characteristic is mass accuracy, which is particularly important when identifying peptides by comparing them with theoretical database values. There are four main mass analyzers widely used in proteomics, including the linear ion trap (LIT), the triple quadrupole (Q-q-Q), the time of flight (TOF) and the Fourier transform ion cyclotron resonance (FT-ICR). The mass analyzer is the part of the mass spectrometer in which an electric and/or magnetic field is used to separate ions in the gas phase. In order to answer specific needs during analysis, mass spectrometers are often today constructed with more than one mass analyzer. These are termed hybrid mass spectrometers and combine different performance characteristics (Table 1.2), offered by different types of analyzers, into one mass spectrometer to produce a highly versatile scientific tool (Aebersold, 2003; Cañas, 2006; Domon, 2006). During the course of this Ph.D. thesis, two hybrid instruments, the Q-q-LIT and the linear trap quadrupole (LTQ)-Orbitrap, containing each distinct performance features and applications (Table 1.2), were used.

Table 1.2 Performance comparison of different instruments commonly used in mass spectrometry-based proteomics (Yates, 2009).

Instrument	Application	Resolution	Mass accuracy	Sensitivity	Dynamic range	Scan rate
LIT	CID, ETD, MS^n	2,000	100 ppm	Femtomole	1×10^4	Fast
Q-q-Q	CID, MS^2, quantification: MRM, P, p and NL scans	2,000	100 ppm	Attomole	1×10^6	Moderate
Q-q-LIT	CID, MS^n, quantification: MRM, P, p and NL scans	2,000	100 ppm	Attomole	1×10^6	Moderate
Q-TOF	CID, MS^2, quantification	10,000	2-5 ppm	Attomole	1×10^6	Moderate
LTQ-Orbitrap	CID, ETD, MS^n, quantification	100,000	2 ppm	Femtomole	1×10^4	Moderate
LTQ-FTICR	CID, ETD, MS^n, quantification	500,000	< 2 ppm	Femtomole	1×10^4	Slow

LIT: linear ion trap; Q-q-Q: triple quadrupole; Q-q-LIT: triple quadrupole linear ion trap; Q-TOF : quadrupole time-of-flight; LTQ-Orbitrap: linear trap quadrupole-orbitrap; LTQ-FTICR: linear trap quadrupole-Fourier transform ion cyclotron resonance; CID: collision induced dissociation; ETD: electron transfer dissociation; MRM: multiple reaction monitoring; P: precursor ion scan; p: product ion scan; NL: neutral loss scan.

First, the Q-q-LIT is a popular hybrid MS instrument that combines the capabilities of two quadrupole mass filters with a linear ion trap (Hopfgartner, 2004). A typical triple quadrupole instrument has the first (Q1) and third quadrupole (Q3) made up of four parallel rods which form a hyperbolic or cylindrical inner surface (Figure 1.8). Opposite rods are electrically connected where each set has a direct current (DC) and an alternating current (AC) or a radiofrequency voltage (RF) applied. By adjusting the voltages, ions of a given *m/z* value will have the correct trajectory to fly through the quadrupole. Ions that have unstable trajectories will collide with the rods. In that manner, ions that may be entering at the same time are separated by continuously varying the applied voltage. The middle quadrupole (Q2) is employed as a gas phase collision cell to fragment peptide ions. An RF-

only voltage is applied in the second quadrupole and therefore does not function as a mass filtering device because all ions are allowed to pass through the quadrupole. The fragment ions emerging from Q2 are analyzed in Q3 before finally reaching the detector. Triple quadrupoles can perform very selective scan modes such as precursor ion scanning, constant neutral loss scans and multiple reaction monitoring (MRM), which are generally not available on other types of mass spectrometers (Cañas, 2006; Lane, 2005) .

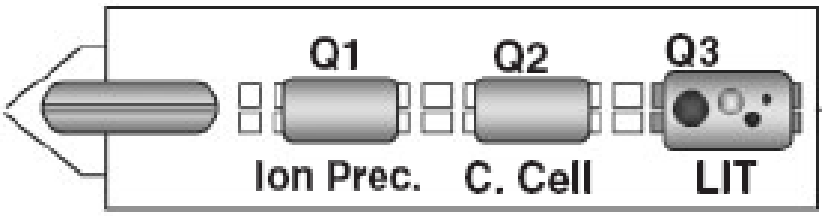


Figure 1.8 Main components of a triple quadrupole linear ion trap (Q-q-LIT) hybrid mass spectrometer (Cañas, 2006).

In the Q-q-LIT, the third component can be either employed as a RF/DC mass filter (Q3) or as a linear ion trap (Figure 1.8). In the trap, ions are trapped radially and axially by the RF and DC voltages respectively and may remain held inside for various periods of time. After a brief cooling period, the *m/z* of ions are scanned before they are mass-selectively ejected into the detector. One advantage of the LIT is that ions that are held in the linear ion trap can be subjected to enhanced MS (EMS) scan and undergo ion-trap fragmentation to generate an enhanced product ion (EPI) MS/MS scan (Hager, 2004; Hopfgartner, 2004). Combining the specificity of the quadrupole and the sensitivity of the ion trap into one instrument, the Q-q-LIT mass spectrometer provides improved performance and enhanced sensitivity. Furthermore, the ability to interconnect triple-quadrupole and ion trap MS scans on a chromatographic timescale is especially beneficial for selective peptide identification and structure characterization (Le Blanc, 2003).

Second, the LTQ-Orbitrap, recently invented by Alexander Makarov, is one of the most efficient hybrid mass analyzer used for proteomics applications (Makarov, 2006a; Scigelova, 2006). The LTQ-Orbitrap bears similarity to the FT-ICR instrument in terms of performance characteristics, but at a much lower cost, making the LTQ-Orbitrap a widespread instrument in proteomics. This hybrid instrument combines three main elements: a quadrupole linear ion trap, a c-trap and an orbitrap (Figure 1.9). Positioned at the front end of the instrument, the linear ion trap first accumulates ions, isolates and fragments them. The LTQ is able to perform MS^3 scans with very high sensitivity, but relatively low resolution and mass accuracy. Accumulated ions are transferred into the C-trap, an RF-only quadrupole that stores and collisionally cools ions prior to their injection into the orbitrap (Makarov, 2006a, b). The orbitrap mass analyzer consists of an outer barrel-like electrode and an inner spindle-like electrode that form an electrostatic field (Scigelova, 2006). Ions are pulsed tangentially into the electric field between the electrodes, where they orbit around the central electrode in circles (Makarov, 2006b). Similarly to the FT-ICR, the orbitrap is used as a mass analyzer by sensing the axial frequency of ion oscillation which is proportional to the square root of *m/z* (Makarov, 2006a).

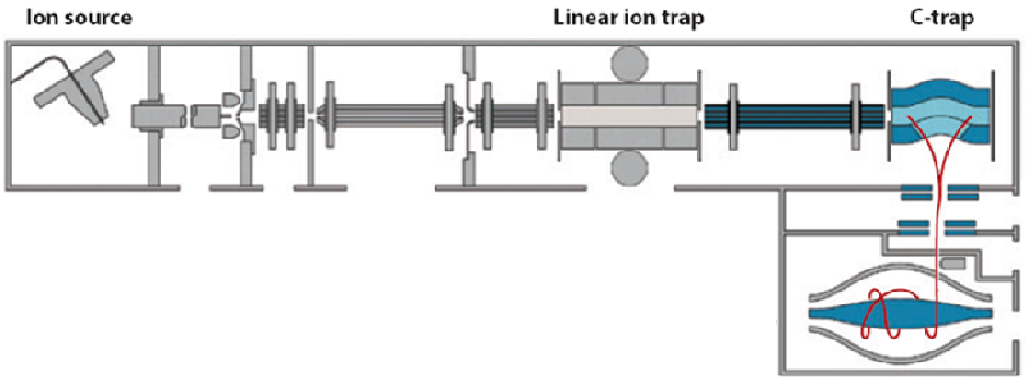


Figure 1.9 Main components of a LTQ-Orbitrap hybrid mass spectrometer (Yates, 2009).

As a result, the LTQ-Orbitrap provides excellent mass accuracy (2-5 ppm), high resolving power (>100,000) and a wide dynamic range (~5000). Finally, combined with the speed

and sensitivity of the LIT, the LTQ-Orbitrap has the advantage of reducing the rate of false positive identifications as well as facilitating the unambiguous identification of low abundant post-translational modifications in complex samples (Makarov, 2006a; Scigelova, 2006).

1.5.1.3.3. Peptide fragmentation under low energy collision

As mentioned earlier, to determine the identity of a peptide, the peptide itself must get fragmented along its backbone. This fragmentation process generates an MS/MS spectrum that basically contains a list of all *m/z* ratios corresponding to different peptide fragments. The difference between each *m/z* value corresponds to the mass of one amino acid (Steen, 2004). The type of fragment ions that get produced depends on the decomposition process. The most common gas-phase dissociation technique used to generate fragment ions from peptides is collision induced dissociation (CID), a low energy regimen. The process typically involves isolating a particular peptide ion, which then gets fragmented as it collides with an inert gas such as nitrogen, helium or argon at low pressure (Steen, 2004). In CID, the fragmentation mainly occurs at the peptide amide bond where two types of fragments are generated (Figure 1.10). The first fragment, where the charge is retained on the N-terminus of the peptide, is termed a «b-ion». The second fragment, where the charge is retained on the C-terminus part of the peptide is termed a «y-ion». Connecting the fragments with increasing size in the MS/MS spectrum, either from the N- or the C-terminus (b and y ions series, respectively), allows for the complete amino acid sequence of the peptide to finally be deduced. In addition, a-ions, corresponding to the loss of a C=O group from b-ions, can sometimes be produced. However, this is most often observed only for the b_2-ion. (Budzikiewicz, 2005; Liebler, 2002; Steen, 2004).

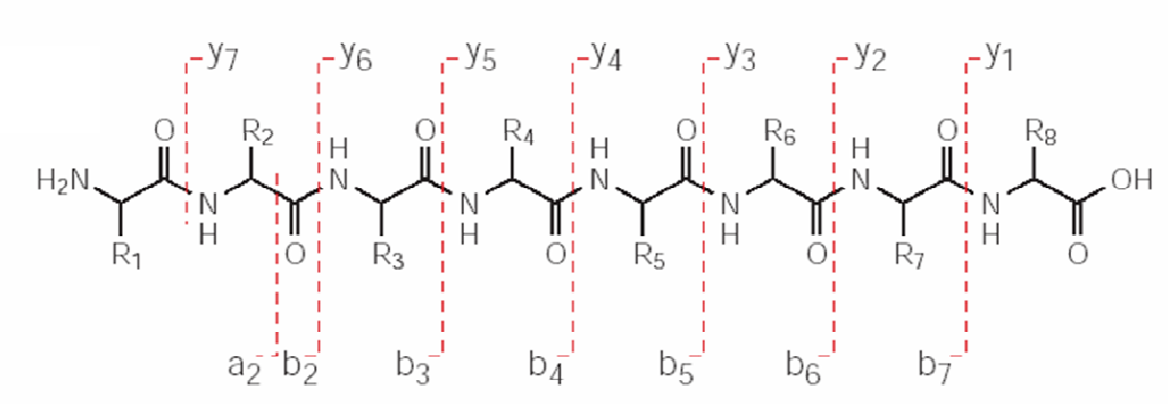


Figure 1.10 Nomenclature of peptide fragmentation patterns with collision induced dissociation (CID). Cleavage along the peptidic bond generates two ions containing the C-terminal (y-ions) end and the N-terminal (b-ions) ends (Steen, 2004).

Since the CID technique will fragment a peptide at its weakest bonds, some labile side chain modifications, such as phosphorylation, can be cleaved by this dissociation technique because the phosphodiester bond is more labile than the amide linkage. As a result neutral loss fragments are also frequently generated under CID and observed in MS/MS spectra (Cañas, 2006). How this feature can be used to detect phosphopeptides and identify the exact site of modification will be discussed in more details in future sections.

1.5.1.3.4. Database searching using MS/MS spectral data

There are two ways in which peptides can be sequenced. First, MS/MS spectra can be manually interpreted. This *de novo* sequencing approach is usually performed without prior knowledge of the peptide or protein primary structure. Specialized softwares (i.e., PEAKS) have been developed to facilitate assignment of peptide sequences without database correlation (Ma, 2003). However, accurate manual sequencing can be difficult and is largely dependent on the complexity of the MS/MS spectrum in terms of resolution and

mass accuracy of the fragment ions. In addition, searching and assigning every MS/MS spectra by hand is very challenging and time consuming as mass spectrometers can collect thousands of spectra per analysis (Liebler, 2002; Steen, 2004). As a result, database searching has become a valuable tool in mass spectrometry. There currently exist many different programs for database searching including MASCOT (Perkins, 1999), SEQUEST (Eng, 1994) and OMSSA (Barnsnes, 2009). These database searching algorithms perform an *in silico* digestion of the protein or the complete proteome to be investigated and finds peptides with theoretical masses corresponding to masses measured by the MS instrument. The algorithm then correlates experimental MS/MS spectra with theoretical spectra generated *in silico*. Based on the probability that a match is not random, each of these programs uses their own scoring system to then evaluate how well the experimental and theoretical spectral data match (Liebler, 2002; Steen, 2004). These scores help discriminate between true and false peptide assignments. The search space is typically limited by parameters such as mass tolerance, enzyme specificity, number of missed cleavages and post-translational modifications. These parameters help to minimize the rate of false positive identifications and increase the confidence level of each peptide assignment.

Overall, database searching tools have the advantages of being fast and offer high-throughput analysis of peptides and proteins. Furthermore, only partial interpretation of the spectrum is enough for database searching. It is therefore possible to sequence and identify peptides of very low abundance. Nevertheless, further manual validation of search results, especially for low abundant and modified peptides (such as phosphopeptides) is highly recommended since false positives do occasionally arise.

1.5.1.3.5. Relative quantification of peptide abundance

There is an increasing requirement in mass spectrometry for rapid, highly reproducible and accurate quantification strategies that allow protein (or peptide) abundance changes to be determined in complex biological samples. These changes, which can result from genetic, environmental or chemical perturbations, are currently measured using two distinct approaches (Figure 1.11): differential isotopic labelling and label-free quantification (Bantscheff, 2007).

The differential isotope labelling technique (Figure 1.11a) is based on the theory that the chemical and physical properties of an isotope-labelled peptide are identical to its unlabelled native counterpart (Bantscheff, 2007). The two peptides behave identically during their chromatographic separation. However, due to the differential isotope labelling, the two peptides will differ in mass. The mass spectrometer can recognize the difference in mass between the labelled and unlabelled peptides and compare their signal intensities to achieve accurate quantification. Isotope labels can be incorporated either *in vitro*, by using light or heavy versions of isotope-tagging reagents or *in vivo*, by the incorporation of isotope-labelled amino acids through metabolic labelling (Gstaiger, 2009). Although isotope-labelling techniques are thought to measure abundance with high accuracy, they still face several limitations. These include increase in sample complexity, requirement for higher sample concentration, incomplete labelling and comparison is often limited to only 2 or 3 different sample conditions. Finally, the cost of reagents in labelling experiments is quite expensive (Bantscheff, 2007; Gstaiger, 2009; Patel, 2009).

The label-free quantification method (Figure 1.11b) is based on the comparison of peptide precursor ions intensities across different samples. In this approach, the ion chromatogram of each peptide is extracted and their mass spectrometric peak areas are integrated over the chromatographic time scale. It has been shown that the signal intensity

measured from ESI is directly proportional with peptide ion concentration, therefore allowing relative abundance changes in peptides to be calculated (Zhu, 2010). Compared to the labelling approach, the label-free method has several advantages (Bantscheff, 2007; Gstaiger, 2009; Patel, 2009; Zhu, 2010). First, this method is less time consuming as well as less expensive since the labelling step is omitted. In addition, there is no limit in the number of experimental samples that can be compared and unlike introducing a label, the sample complexity is not affected. The label-free approach also provides higher dynamic range in terms of quantification. Nevertheless, label-free quantification can still be challenging, particularly during analysis of very complex mixtures.

Two main experimental parameters affect the analytical accuracy of the quantification and the measurement of ion intensities. First, high mass accuracy MS instruments are required in order to minimize assignment of intensities to co-eluting peptides with similar *m/z* values. Second, reproducibility of the LC-system must be maintained to allow corresponding peptides from different samples to be easily paired. Also, run to run analysis can result in differences in peak intensities even between identical samples. This can be caused by experimental variations such as sample preparation, sample injection as well variation in MS performance. To account for these kinds of variations, normalization is often required. Several bioinformatics softwares, such as our in-house Mass Sense and Peptide cluster label-fee softwares, which allow automated comparative analysis of different samples to be performed simultaneously have been developed (Jaitly, 2007). These automated bioinformatics platforms computationally map peptides based on their *m/z*, charge state, retention time and intensity and align them across different samples to generate expression profiles.

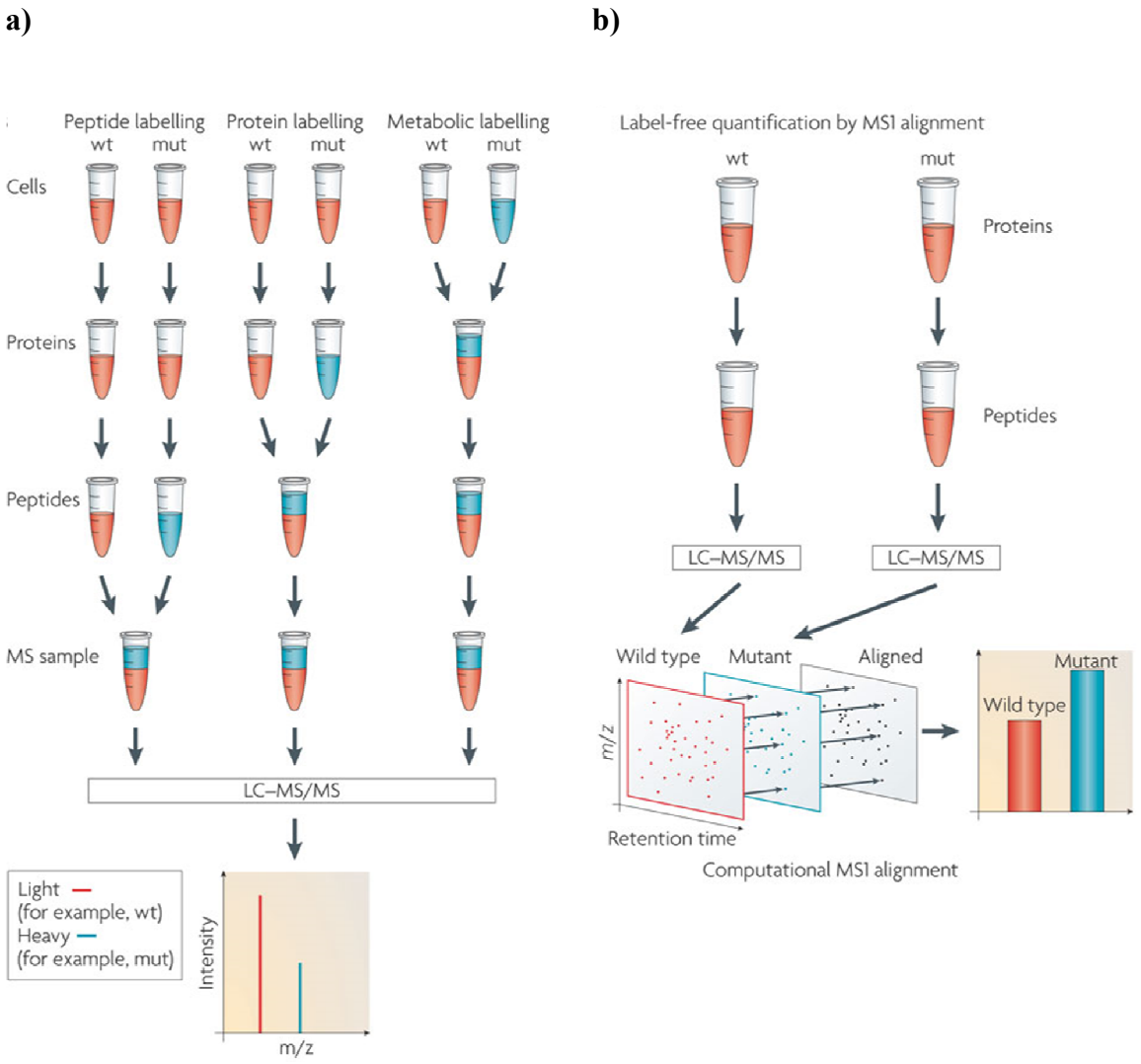


Figure 1.11 General approach of quantitative mass spectrometry. a) Isotope-labelling approach. Proteins or peptides are differentially labelled with heavy or light isotopes. This can be done by *in vitro* chemical or *in vivo* metabolic labelling. After enzymatic digestion, labelled peptides are combined, separated and analysed by mass spectrometry. Differential labelling produces a characteristic mass shift that is used to calculate and quantify the intensity ratio of isotope-labelled peptide pairs. **b) Label-free approach.** Different samples are subjected to individual LC-MS/MS analysis. MS spectra are then computationally aligned to quantify and compare peak intensities for peptide pairs across different conditions (Gstaiger, 2009).

One disadvantage of both the isotope-labelling and label-free approaches is that these methods are somehow biased toward the detection of high abundance peptides. These quantitative methods are also more suited for easily observable intensity changes (Gstaiger, 2009). Recently, targeted MS analysis, which allows the selective detection and quantification of pre-determined low abundant peptides, has been introduced to overcome the challenges of analyzing peptides that fall below the detection limit of conventional shotgun MS analysis (Lange, 2008; Yocum, 2009). This process, also known as multiple reaction monitoring (MRM) has allowed the detection of yeast proteins that are expressed at less than 50 copies per cell in whole cell lysates. This number can even be lowered to 10 copies per cell for fractionated proteins (Picotti, 2009, 2008). This sensitive technique is also ideally suited for the quantification of low abundant post-translational modifications (Lange, 2008; Unwin, 2005; Yocum, 2009). Its application in profiling protein phosphorylation is further explained in details in section 1.5.2.4.

1.5.2. Methods to detect protein phosphorylation

One of the most studied regulatory post-translational modification of proteins is the phosphorylation of serine, threonine and tyrosine residues. Regulation by phosphorylation can alter various functional properties of the substrate protein such as enzymatic activity, sub-cellular localization, stability, binding specificity, protein interaction and DNA interaction (Salazar, 2009; Swarup, 1998). Radiolabel studies suggest that approximately one third of all eukaryotic proteins are phosphorylated at one time or another (Zolnierowicz, 2000). The variety is enormous and as a result, a broad range of cellular processes such as cell cycle control, DNA replication, transcription and signal transduction are regulated by phosphorylation events (Salazar, 2009).

Mass spectrometry has a particular strength in its ability to characterize protein phosphorylation. The effectiveness of mass spectrometry lies in its ability to sequence phosphorylated residues and simultaneously identify all phosphoproteins and

phosphopeptides in one single analysis (Mann, 2002). However, even with its excellent sensitivity and high resolving power, MS analysis of phosphoproteins is still an inherently difficult task. One reason is that phosphoproteins of interest are generally of low abundance. As a result, phosphopeptides are grossly underrepresented and selectively suppressed in the presence of high abundant unmodified peptides. Second, phosphopeptides are negatively charged and therefore have lower ionization and detection efficiencies than their unphosphorylated counterparts in positive mode ESI-MS (Aebersold, 2001; Mann, 2002; Steen, 2006). To overcome these drawbacks, several strategies for detecting and characterizing phosphoproteins by MS have been developed (Figure 1.12).

1.5.2.1. Enrichment techniques

Enrichment strategies are often used to isolate and concentrate the fraction of phosphorylated peptides in the sample to be studied (Mann, 2002; Yates, 2009). Two of the major methods used to capture phosphopeptides include immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) (Figure 1.12) (Dunn, 2010). The IMAC approach exploits the high affinity that negatively charged phosphopeptides exhibit towards positively charged metal ions (Fe^{3+} and Ga^{3+}) (Posewitz, 1999). These are complexed with metal binding ligands such as the iminodiacetate (IDA) and nitrilotriacetate (NTA), with Fe^{3+}- NTA being the most frequently employed complex used to enrich phosphopeptides (Dunn, 2010). One challenge encountered in IMAC is the high level of non specific binding of peptides containing acidic residues such as glutamic acid and aspartic acid. In order to overcome this problem conversion of the carboxylic acid group of amino acids to methyl esters by reaction with methanolic acetyl chloride is often performed, thereby reducing nonspecific binding by at least two orders of magnitude (Ficarro, 2002). However, incomplete esterification also occurs and some carboxylic residues remain unmethylated. This has been shown to lead to an increase in sample complexity originating from peptides with different degrees of methylation and occasionally sample loss (Haydon, 2003; Trinidad, 2006; Tsai, 2008). Another way of

increasing the selectivity of IMAC is by extraction at low pH. Using a strong acid like TFA before enrichment, protonates the carboxylic residues of peptides and therefore reduces their non specific binding (Kokubu, 2005).

Alternatively, the use of the MOAC method for enriching phosphopeptides has been growing in the past years mainly because it offers high selectivity and recovery without requiring a methyl esterification step (Dunn, 2010). Extraction of phosphopeptides with titanium dioxide (TiO_2) as the metal oxide resin is currently the most popular method used (Thingholm, 2006). Although not a lot is known about the surface chemistry and binding properties of TiO_2. Several investigations have shown that the bidentate binding of TiO_2 to phosphate is pH dependent (Connor, 1999). Also, TiO_2 possesses strong Lewis acid properties and high coordination numbers, which in combination, may be a possible explanation for the affinity of phosphate molecules to metal oxides (Grün, 1996). Under acidic conditions, the positively charged surface of TiO_2 exhibit outstanding selectivity and enrichment behaviour towards phosphopeptides over non-modified acidic peptides. Mainly, the use of additives such as lactic acid and 2,5-dyhydroxybenzoic (DHB) acid have been shown to improve the selectivity of enrichment (Jensen, 2007; Larsen, 2005; Sugiyama, 2007; Thingholm, 2006). Lactic acid and DHB are known as displacer molecules that compete with the non-modified acidic peptides for TiO_2 binding while virtually unaffecting the enrichment of phosphopeptides (Larsen, 2005). Finally, compared to IMAC, which has higher affinity for multi-phosphorylated peptides, TiO_2 exhibits higher selectivity toward mono-phosphorylated peptides (Larsen, 2005).

Summing up, IMAC and TiO_2 are generally performed off-line using tips or micro-columns (Hsieh, 2007; Posewitz, 1999). However, automated on-line enrichment coupled to ESI-MS can also be performed. In addition to obtaining fast and more reproducible data acquisition, this online method also simplifies the enrichment procedure and minimizes sample loss (Pinkse, 2004; Wang, 2006). The IMAC and TiO_2 enrichment steps also have the advantage of pre-concentrating the sample and removing reagents not compatible with

MS (Corthals, 2005). However, TiO_2 has shown to be more tolerant against salts, detergents and denaturing components than IMAC. Compared to IMAC, TiO_2 also provides much greater benefits for the enrichment of phosphopeptides in terms of phosphate affinity as well as chemical and physical stability (Jensen, 2007).

1.5.2.2. Chemical derivatization

Enrichment techniques decrease sample complexity and consequently, significantly improves the detection of phosphopeptides by MS. Still, due to the chemical features of the phosphate group, MS analysis of phosphorylated peptides remains limited. As mentioned previously, the presence of the negatively charged phosphate group decreases the ionization efficiency of the modified peptides. Additionally, due to the labile nature of phosphoserine and phosphothreonine residues, MS-based sequence identification of phosphopeptides is often impeded (Mann, 2002). The problem principally occurs during the process of peptide fragmentation. Upon collision with the chemically inert gas, peptide ions become particularly weakened at protonated amide linkages (Dongré, 1996). The amide bonds then randomly dissociate along the peptide backbone, producing fragment ions that allow MS/MS interpretation of the amino acid sequence (McLafferty, 2001; Steen, 2004). However, in the presence of phosphoserine or phosphothreonine residues, the phosphoryl group competes with the amide linkage for the preferred site of protonation. As a result, cleavage is often directed towards the phosphoryl group rather than the amide bond. The fragmentation results in a dominant loss of H_3PO_4, such that fragments generated from amide bond cleavage along the peptide backbone are negligible. Consequently, MS/MS spectra of phosphorylated peptides often do not contain sufficient information to allow confident assignment of amino acid sequence and determination of exact site of phosphorylation (Carr, 1996).

One way by which interpretation of MS/MS spectra becomes more straightforward is by removing and/or replacing the phosphate group by a more stable and less acidic

group. The phosphate moiety can be chemically modified through a β-elimination reaction which occurs when phosphorylated residues are exposed to strong alkaline conditions (Byford, 1991; McLachlin, 2003). Strong bases such as NaOH or $Ba(OH)_2$ are used to cleave the phosphodiester bond of phosphoserine and phosphothreonine residues converting them to their respective dehydroalanine and dehydrobutyric acid analogs (Figure 1.12). This derivatization results in a mass shift of -18Da (compared to the unmodified residue) than can be detected by MS/MS and directly used to locate the exact site of phosphorylation. Alternatively, the α,β-unsaturated residue generated can also act as a Michael acceptor and react with a variety of thiol-based nucleophiles such as dithiothreitol (DTT) or ethanedithiol (EDT) (Jaffe, 1998). The free sulhydryls can then be coupled to a biotinylated affinity tag allowing purification of the phosphoprotein (Figure 1.12) (Oda, 2001). The appropriate mass shit can then be detected by MS/MS.

One weakness of this strategy is the low reactivity of phosphothreonine compared to phosphoserine and the non-reactivity of phosphotyrosine residues which, because of its stable aromatic ring, cannot undergo β-elimination (Ahn, 2001). In addition, O-glycosylated and sulfated serine and threonine residues can also undergo β-elimination. This can yield to false positive identifications because the end product is identical to phosphorylated serine and threonine residues (Medzihradszky, 2004; Reinders, 2005). Finally, another complication arises as incomplete derivatization and side reactions of cysteine and methionine residues which increase the complexity of the sample and affect phosphopeptide intensity and detection (Oda, 2001; Reinders, 2005).

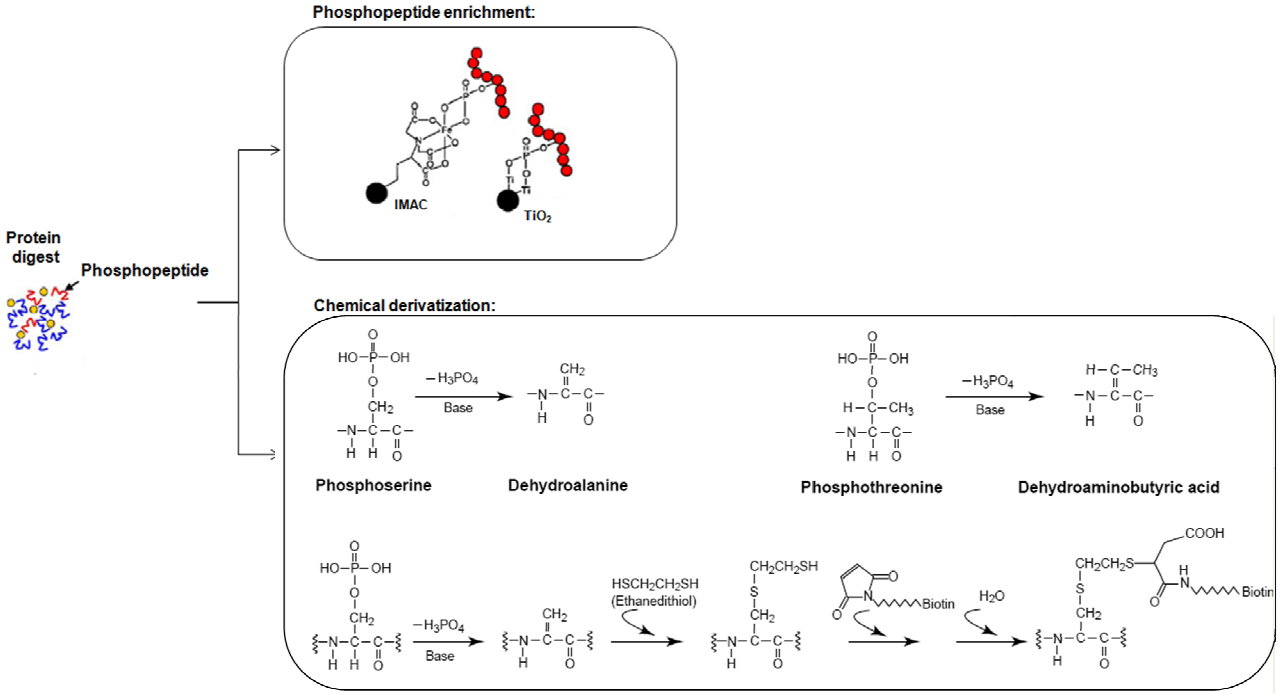


Figure 1.12 Overview of common approaches used for phosphopeptide analysis by mass spectrometry. The top panel describes IMAC and TiO_2 strategies for the specific enrichment of phosphopeptides (Dunn, 2010). The bottom panel describes the chemical derivatization approach. Under strong basic conditions, phosphoserine and phosphothreonine residues undergo a β-elimination reaction. Due to the loss of phosphoric acid, an α,β-unsaturated bond is formed where the specific products are dehydroalanine and dehydroaminobutyric acid, respectively. The α,β-unsaturated bond can then undergo Michael addition with a thiol group and a biotin tag for enrichment on avidin column (Mann, 2002).

1.5.2.3. Fragmentation methods for phosphopeptide identification

Taking advantage of the chemical labile nature of the phosphoester bond of phosphorylated residues, various mass spectrometric methods for detecting phosphopeptides and phosphorylated amino acids have been developed. In CID, peptides

containing phosphoserine and phosphothreonine residues commonly undergo a gas-phase β-elimination reaction which results in a neutral loss of phosphoric acid (H_3PO_4; 98 Da) from the precursor ion (DeGnore, 1998; Mann, 2002). Doubly and triply charged phosphopeptides would show a noticeable loss of 49 and 32.66 Da in the MS/MS spectrum. When fragmentation along the peptide backbone is not negligible, amino acid sequence interpretation often shows a spacing of 69 Da due to dehydroalanine or a spacing of 83 Da due to dehydroaminobutyric acid, pin-pointing the exact site of the phosphorylated serine and threonine residues, respectively (Cañas, 2006; Mann, 2002). Phosphotyrosine residues are generally more resistant to this loss (Aebersold, 2001). In fact, when tyrosine residues are phosphorylated and undergo CID, an increased mass shift of 80 Da, which is the mass of a phosphate group, can be observed on the corresponding tyrosine fragment. Consequently, phosphotyrosine with a molecular weight of 243Da can be identified (Cañas, 2006). Occasionally, a neutral loss of a phosphate molecule (HPO_3; 80 Da) can also occur, but is less common (DeGnore, 1998; Thingholm, 2009). Compared to the loss of H_3PO_4, loss of HPO_3 does not generate a mass shift that allows the identification of the exact site of phosphorylation. Alternatively, electron transfer dissociation (ETD) is a recently introduced method that fragments peptides by transferring electrons from fluoranthene radical ions generated in a chemical ionization source to the protonated peptide (Syka, 2004). This reaction induces fragmentation along the peptide backbone by specifically cleaving the C_α-N bond, producing complementary c- and z-type fragment ions instead of the typical b- and y-type ions generated by CID. One important advantage of ETD over CID is that labile post-translational modifications, such as phosphorylation, are preserved, allowing the complete amino acid sequence and site of phosphorylation to be deduced with relative ease (Mikesh, 2006).

Phosphopeptides can also be selectively targeted using the neutral loss as an indicator. Using ion trap hybrid mass spectrometry, the fragment ion originating from the neutral loss can also be selected for further fragmentation by a process called data-dependent neutral loss MS^3 (Figure 1.13) (Beausoleil, 2004).

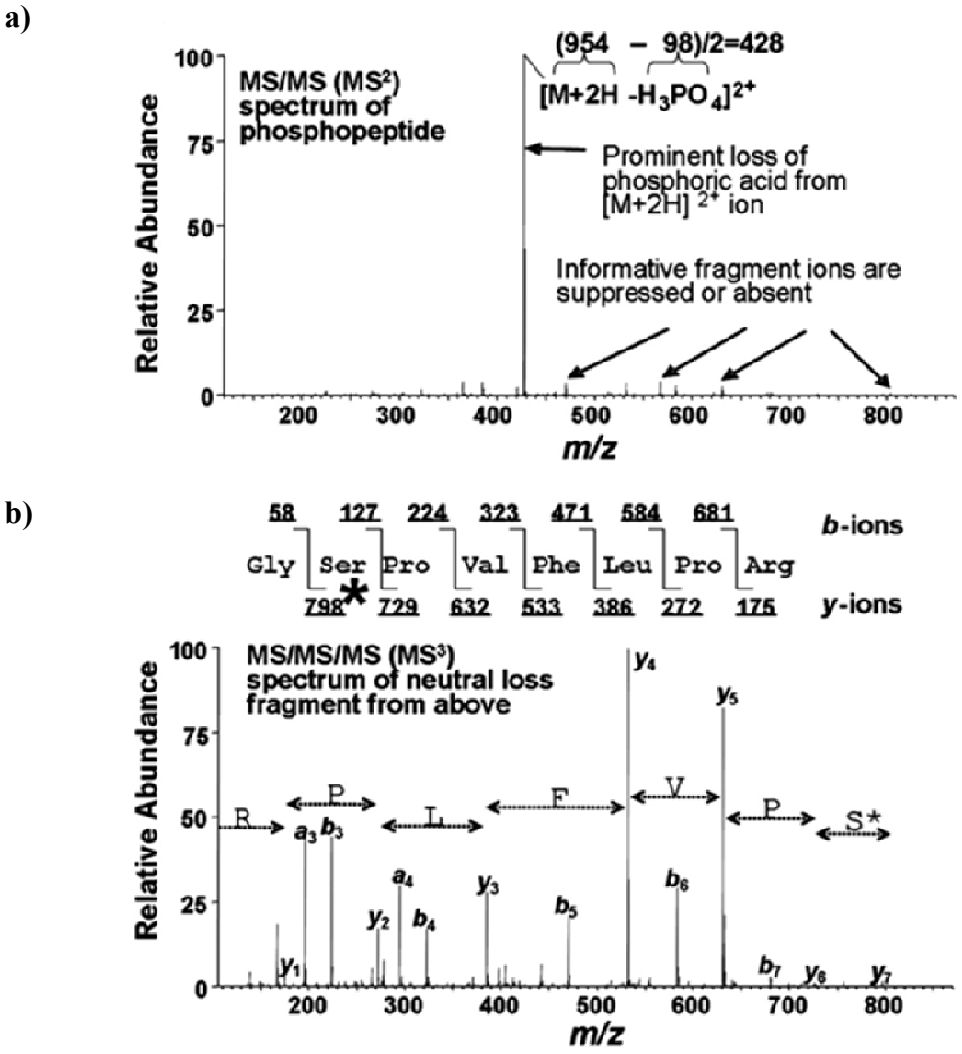


Figure 1.13 Example of a phosphopeptide analyzed by data-dependent neutral loss MS^3. a) MS/MS spectrum of a phosphopeptide presenting a strong neutral loss of phosphoric acid (H_3PO_4) and negligible fragmentation along the peptide backbone. **b)** MS^3 spectrum generated from the isolation and activation of the prominent neutral loss fragment generated in a. Unambiguous identification of the peptide sequence and phosphorylation site is possible due to the abundant peptide bond fragmentation (Beausoleil, 2004).

One advantage of this method is that the fragmentation of the product ion resulting from the neutral loss is sequence specific and allows phosphorylation site identification by

detection of dehydroalanine and dehydroaminobutyric acid. One disadvantage is that sometimes low abundant phosphopeptides do not provide interpretable MS^3 spectrum, instead this technology is well suited for relatively higher abundant phosphopeptides. This is attributed to insufficient trapping of ions over several MS stages (Beausoleil, 2004). In order to circumvent this problem, an alternative method called pseudo-MS^3 or multistage activation (MSA) was developed (Schroeder, 2004). In MSA, the neutral loss ion product is activated by CID while fragments from the precursor ion are still in the trap, in a single MS stage. The resulting spectrum contains fragments from the precursor and the neutral loss product. One advantage of MSA is that the analysis time per peptide is reduced since the ion trap does not require a refilling step for the MS^3 scan. Finally, MSA generates information rich spectra with increased signal intensity and a greater number of structurally diagnostic ions (Schroeder, 2004; Ulintz, 2008).

1.5.2.4. Multiple reaction monitoring (MRM)

MS-based analysis of peptides in general, including phosphopeptides, is usually conducted in data-dependent acquisition (DDA) mode, where the mass spectrometer continuously repeats a cycle consisting of a full scan MS spectrum followed by the selection and fragmentation of the most abundant *m/z* ratios. Depending on the type of instrument used, the number of selected precursor ions that get fragmented can vary between 1 and 10 (Wolf-Yadlin, 2007). One main disadvantage of this technique lies in the inconsistent reproducibility of the method (Elias, 2005). Affected by the complexity of the sample and coupled by variability in automated peak selection, phosphopeptide identification and quantification often suffer.

Multiple reaction monitoring (MRM), using a triple quadrupole linear ion trap (Q-q-LIT) instrument, is ideally suited for obtaining maximum sensitivity and reproducibility for the targeted identification and quantification of proteins of interest. Knowing the mass, primary sequence, modification type and the preferred fragmentation pathway of peptides

of interest, specific precursor ions (proteotypic peptides) that generate specific fragment ions, known as MRM transitions, can be monitored (Cox, 2005). The transitions are designed such that the first mass resolving quadrupole (Q1) is set to transmit only an ion of specific *m/z* for its subsequent fragmentation by CID in the second quadrupole (Q2). The third quadrupole (Q3) is then set to transmit a specific fragment ion (Figure 1.14). These are termed Q1 and Q3 transition masses (Yocum, 2009). In the case of phosphopeptide analysis, an MRM transition of 98 Da is used to monitor loss of phosphoric acid from a specific peptide (Cox, 2005; Unwin, 2005).

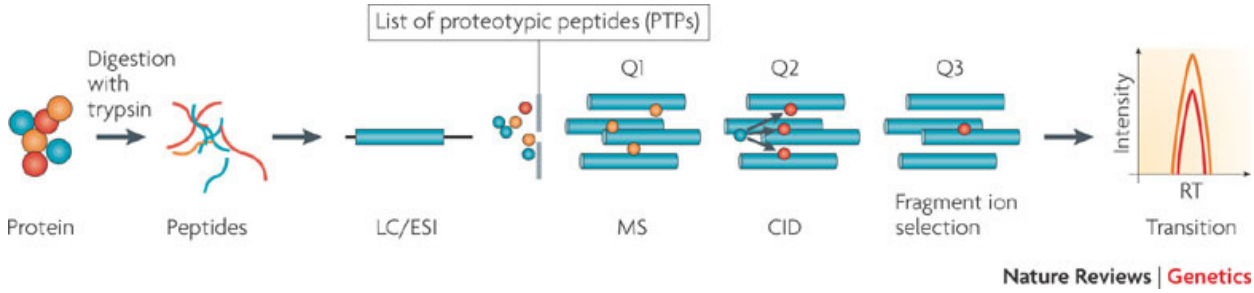


Figure 1.14 Schematic representation of the MRM technique using a Q-q-LIT instrument. Proteins are digested and introduced into the mass spectrometer by LC-ESI. The mass spectrometer is programmed such that precursor ions are sequentially and selectively transmitted through Q1, fragmented following collision induced dissociation (CID) in Q2. Finally, the third quadrupole (Q3) selectively transmits the appropriate fragment ion (Gstaiger, 2009).

The detection of two ions is required to generate an MRM signal, making it very selective, as it is highly unlikely that isobaric molecules that may co-elute with the target peptide will also generate an identical fragment mass (Cox, 2005). It is also possible to further confirm the identity of a phosphopeptide detected by MRM by designing the MS instrument in such a way to obtain MS and MS/MS spectra for each potential candidate. As the signal corresponding to a specific MRM transition reaches a pre-defined threshold, several linear ion trap scans can be triggered, including an enhanced resolution scan as well as an enhanced product ions scan that confirm the charge state and the mono-isotopic mass

of the targeted phosphopeptide as well as its sequence and the location of the phosphoylation site (Cox, 2005). Several MRM transitions can be monitored simultaneously in a single experiment. It becomes therefore possible to probe for several phosphopeptides from a known protein sequence. The time taken to analyse each transition is termed dwell-time. The time taken to analyze all transitions is termed cycle-time. Both affect the sensitivity of the analysis. At the cost of decreasing the productivity of the analysis (i.e., a decrease in the number of transitions that can be monitored simultaneously), a more sensitive measurement is usually obtained by increasing the dwell-time. Other MS parameters, such as the collision energy, can be adjusted to generate optimal peptide fragmentation and maximise the transmission of the desired product ions (Yocum, 2009).

In summary, the main advantage of the targeted MRM method lies in its throughput as well as in its increased specificity, reproducibility and sensitivity (Unwin, 2005). The fact that only specific ions with defined *m/z* values are fragmented and only specific fragments are allowed to reach the detector means that there is very little chemical background. The enhanced signal-to-noise (S/N) ratio allows for a more sensitive detection of low levels phosphopeptides as well as the ability to accurately quantify them over a much greater dynamic range (5 orders of magnitude) (Wolf-Yadlin, 2007).

1.6. Research objectives

Using mass spectrometry as a tool to identify and characterize protein phosphorylation, the major aim of the research conducted during the course of my Ph.D. was to understand the regulatory post-translational mechanisms that control the process of chromatin assembly during DNA replication and repair. My studies were mainly centered on the phosphorylation of CAF-1 and Hir proteins.

Initially, within the context of protein phosphorylation analysis by mass spectrometry (MS), the silver staining procedure, which is frequently used for visualizing gel-separated proteins prior to MS analysis, was evaluated. The goal of this project was to follow up on a previous study that was focused on indentifying and characterizing the phosphorylation of another protein involved in chromatin assembly namely, proliferating cell nuclear antigen (PCNA). Our data indicated to us that PCNA peptides which were initially believed to be phosphorylated *in vivo* were instead artifactually sulfated during sample preparation. As phosphorylation and sulfation of proteins both impart a mass shift of 80 Da on serine, threonine and tyrosine residues, distinction between these two types of modifications by MS is quite challenging. The data collected in chapter 2 highlight important pitfalls with silver staining in the assignment of protein phosphorylation. Based on the propensity of the sulfation artifact, the silver and Coomassie staining techniques were compared with each other. In order to avoid that phosphorylation sites be mis-identified, MS instruments, such as the LTQ-orbitrap, which provides higher resolution and better mass accuracy, should be used and, more importantly, silver staining should be completely avoided for accurate identification of phosphorylation sites. Hence, Coomassie brilliant blue (CBB) staining is the recommended staining method for all future protein phosphorylation analysis.

Chromatin assembly factor, CAF-1, plays an important role in *de novo* nucleosome assembly by specifically bringing histones H3 and H4 molecules to newly synthesized DNA, as opposed to non-replicating DNA. CAF-1's unique ability to promote nucleosome assembly onto newly synthesized DNA is achieved through a direct physical interaction between its largest subunit and PCNA. As highlighted in the introduction, CAF-1 has been identified as a phosphoprotein *in vivo*. Furthermore, *in vitro* studies have revealed that CAF-1 phosphorylation is controlled by cyclin/CDK2 (CDK) and Cdc7-Dbf4 (DDK). The goal of this project, presented in chapter 3, was to understand more clearly how phosphorylation regulates CAF-1 function during DNA replication. In order to do so, we used a combination of SDS-PAGE followed by LC-MS/MS analysis of immunoprecipitated CAF-1 from yeast to show that the largest subunit, Cac1, is phosphorylated on novel serine and/or threonine residues. In the second part, we set out to assess the functional significance of specific phosphorylation sites identified in CAF-1. This was achieved by performing phenotypic assays using mutated versions of Cac1 as well as kinase assays to further confirm the protein kinase that phosphorylates specific residues on Cac1 *in vitro*.

Finally, as mentioned, the Hir protein complex in yeast plays a dual role in the nucleosome assembly pathway. In addition to binding to histones H3 and H4 and promoting their assembly into nucleosomes, Hir proteins contribute to histone gene regulation during normal cell cycle progression and DNA repair by transcriptionally repressing histone genes when demand for histone synthesis is low. Recent studies have demonstrated that the DNA damage checkpoint kinase, Rad53, prevents the accumulation of excess histones in response to DNA damage or replication stress during S-phase (Gunjan, 2003). In chapter 4, we set out to determine if Hir proteins are a direct target of phosphorylation by the DNA damage checkpoint kinases. For this specific goal, we used a targeted quantitative MS approach to determine if the level of Hir protein phosphorylation is specifically altered during activation or inhibition of the DNA damage checkpoint kinases. Furthermore, we also performed an LC-MS/MS analysis of the Hir protein

complex in order to gain insights into how phosphorylation might be implicated in regulating the function of Hir proteins in histone gene regulation during normal cell cycle progression.

1.7. References

Abraham, R.T. (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes & Development *15*, 2177-2196.

Adkins, N.L., Watts, M., and Georgel, P.T. (2004). To the 30-nm chromatin fiber and beyond. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *1677*, 12-23.

Aebersold, R., Goodlett, D.R. (2001). Mass Spectrometry in Proteomics. Chemical Reviews *101*, 269-296.

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

Ahmad, K., Henikoff, S. (2002a). Histone H3 variants specify modes of chromatin assembly. Proceedings of the National Academy of Sciences of the United States of America *99*, 16477-16484.

Ahmad, K., Henikoff, S. (2002b). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Molecular Cell *9*, 1191-1200.

Ahn, N.G., Resing, K.A. (2001). Towards the phosphoproteome. Nature Biotechnology *19*, 317-318.

Ai, X., Parthun, M.R. (2004). The nuclear Hat1p/Hat2p complex: a molecular link between type B histone acetyltransferases and chromatin assembly. Molecular Cell *14*, 195-205.

Aparicio, O.M., Billington, B.L., Gottschling, D.E. (1991). Modifiers of position effect are shared between telomeric and silent mating-type loci in *Saccharomyces cerevisiae*. Cell *66*, 1279-1287.

Ataian, Y., Krebs, J.E. (2006). Five repair pathways in one context: chromatin modification during DNA repair. Biochemistry & Cell Biology *84*, 490-504.

Bambara, R.A., Murante, R.S., Henricksen, L.A. (1997). Enzymes and reactions at the eukaryotic DNA replication fork. Journal of Biological Chemistry *272*, 4647-4650.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. Analytical and Bioanalytical Chemistry *389*, 1017-1031.

Banumathy, G., Somaiah, N., Zhang, R., Tang, Y., Hoffmann, J., Andrake, M., Ceulemans, H., Schultz, D., Marmorstein, R., Adams, P.D. (2009). Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Molecular & Cellular Biology *29*, 758-770.

Barak, O., Lazzaro, M.A., Lane, W.S., Speicher, D.W., Picketts, D.J., and Shiekhattar, R. (2003). Isolation of human NURF: a regulator of Engrailed gene expression. The EMBO Journal *22*, 6089-6100.

Barsnes, H., Huber, S., Sickmann, A., Eidhammer, I., Martens, L. (2009). MOSA Pareser: an open-source library to parse and extract data from OMSSA MS/MS search results. Proteomics *14*, 3772-3774.

Bauer, G.A., and Burgers, M.J. (1990). Molecular cloning, structure and expression of the yeast proliferating cell nuclear antigen gene. Nucleic Acids Research *18*, 261-265.

Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villén, J., Li, J., Cohn, M.A., Cantley, L.C., Gygi, S.P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. Proceedings of the National Academy of Sciences of the United States of America *101*, 12130-12135.

Bell, S.P., Dutta, A. (2002). DNA replication in eukaryotic cells. Annual Review of Biochemistry *71*, 333-374.

Bell, S.P., Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature *357*, 128-134.

Ben-Shahar, T.R., Castillo, A.G., Osborne, M.J., Borden, K.L.B., Kornblatt, J. and Verreault, A. (2009). Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. Molecular & Cellular Biology *29*, 6353-6365.

Bernard, P., Maure, J-F., Partridge, J.F., Genier, S., Javerzat, J-P., Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. Science *294*, 2539-2542.

Bowman, G.D., O'Donnell, M., and Kuriyan, J. (2004). Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. Nature *429*, 724-730.

Bravo, R., and Macdonald-Bravo, H. (1987). Existence of two populations of cyclin/proliferating cell nuclear antigen during the cell cycle: association with DNA replication sites. The Journal of Cell Biology *105*, 1549-1554.

Brown, D.T. (2003). Histone H1 and the dynamic regulation of chromatin function. Biochemistry & Cell Biology *81*, 221-227.

Brownell, J.E., Allis, C.D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Current Opinion in Genetics & Development *6*, 176-184.

Bruning, J.B., and Shamoo, Y. (2004). Structural and Thermodynamic Analysis of Human PCNA with Peptides Derived from DNA Polymerase-´ p66 Subunit and Flap Endonuclease-1. Structure *12*, 2209-2219.

Brush, G.S., Morrow, D.M., Hieter, P., Kelly, T.J. (1996). The ATM homologue MEC1 is required for phosphorylation of replication protein A in yeast. Proceedings of the National Academy of Sciences of the United States of America *93*, 15075-15080.

Budzikiewicz, H. (2005). J.H. Gross: Mass Spectrometry. A Textbook, Vol 381 (Springer Berlin / Heidelberg).

Byford, M.F. (1991). Rapid and selective modification of phosphoserine residues catalysed by Ba2+ ions for their detection during peptide microsequencing. Biochemical Journal *280*, 261-265.

Cañas, B., López-Ferrer, D., Ramos-Fernández, A., Camafeita, E., Calvo, E. (2006). Mass spectrometry technologies for proteomics. Briefings in Functional Genomics and Proteomics *4*, 295-320.

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G. M., Carnemolla, B., Orecchia, P., Zardi, L., Righetti, P. G. (2004). Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis *25*, 1327-1333.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science *298*, 1039-1043.

Carmen, A.A., Milne, L., Grunstein, M. (2002). Acetylation of the yeast histone H4 N terminus regulates its binding to heterochromatin protein SIR3. Journal of Biological Chemistry *277*, 4778-4781.

Carr, S.A., Huddleston, M.J., Annan, R.S. (1996). Selective Detection and Sequencing of Phosphopeptides at the Femtomole Level by Mass Spectrometry. Analytical Biochemistry *239*, 180-192.

Cech, N.B., Enke, C. G. (2001). Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrometry Reviews *20*, 362-387.

Chakravarthy, S., Park, Y.-J., Chodaparambil, J., Edayathumangalam, R.S., and Luger, K. (2005). Structure and dynamic properties of nucleosome core particles. FEBS Letters *579*, 895-898.

Chicoine, L.G., Schulman, I.G., Richman, R., Cook, R.G., and Allis, C.D. (1986). Nonrandom utilization of acetylation sites in histones isolated from Tetrahymena. Evidence for functionally distinct H4 acetylation sites. Journal of Biological Chemistry *261*, 1071-1076.

Chuang, L.S.-H., Ian, H.-I., Koh, T.-W., Ng, H.-H., Xu, G., and Li, B.F.L. (1997). Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for p21WAF1. Science *277*, 1996-2000.

Connor, P.A., McQuillan, A. J. (1999). Phosphate adsorption onto TiO2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study. Langmuir *15*, 2916-2921.

Corona, D.F.V., and Tamkun, J.W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *1677*, 113-119.

Corthals, G.L., Aebersold, R., Goodlett, D.R. (2005). Identification of Phosphorylation Sites Using Microimmobilized Metal Affinity Chromatography. In Methods in Enzymology, A.L. Burlingame, ed. (Academic Press), pp. 66-81.

Cox, D.M., Zhong, F., Du, M., Duchoslav, E., Sakuma, T., McDermott, J.C. (2005). Multiple reaction monitoring as a method for identifying protein post-translational modifications. Journal of Biomolecular Techniques *16*, 83-90.

D'Amours, D., and Jackson, S.P. (2001). The yeast Xrs2 complex functions in S phase checkpoint regulation. Genes & Development *15*, 2238-2249.

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

de Laat, W.L., Jaspers, N.G.J., and Hoeijmakers, J.H.J. (1999). Molecular mechanism of nucleotide excision repair. Genes & Development *13*, 768-785.

DeGnore, J.P., Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. Journal of the American Society for Mass Spectrometry *9*, 1175-1188.

Dillon, N., and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. Trends in Genetics *18*, 252-258.

Dimova, D., Nackerdien, Z., Furgeson, S., Eguchi, S., Osley, M.A. (1999). A role for transcriptional repressors in targeting the yeast Swi/Snf complex. Molecular Cell *4*, 75-83.

Dinant, C., Houtsmuller, A., and Vermeulen, W. (2008). Chromatin structure and DNA damage repair. Epigenetics & Chromatin *1*, 9.

Domon, B., Aebersold, R. (2006). Mass Spectrometry and Protein Analysis. Science *312*, 212-217.

Dongré, A.R., Jones, J.L., Somogyi, Á., Wysocki, V.H. (1996). Influence of peptide composition, gas-phase basicity and chemical modification on fragmentation efficiency: evidence for the mobile proton model. Journal of the American Chemical Society *118*, 8365-8374.

Driscoll, R., Hudson, A., and Jackson, S.P. (2007). Yeast Rtt109 Promotes Genome Stability by Acetylating Histone H3 on Lysine 56. Science *315*, 649-652.

Dunn, J.D., Reid, G. E., Bruening, M. L. (2010). Techniques for phosphopeptide enrichment prior to analysis by mass spectrometry. Mass Spectrometry Reviews *29*, 29-54.

Eissenberg, J., Ayyagari, R., Gomes, X., and Burgers, P. (1997). Mutations in yeast proliferating cell nuclear antigen define distinct sites for interaction with DNA polymerase delta and DNA polymerase epsilon. Molecular & Cellular Biology *17*, 6367-6378.

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

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

Eng, J.K., McCormack, A.L., Yates III, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. Journal of the American Society for Mass Spectrometry *11*, 976-989.

Enomoto, S., Berman, J. (1998). Chromatin assembly factor I contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. Genes & Development *12*, 219-232.

Enomoto, S., McCune-Zierath, P.D., Gerami-Nejad, M., Sanders, M.A., Berman, J. (1997). RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function in vivo. Genes & Development *11*, 358-370.

Escargueil, A.E., Soares, D.G., Salvador, M., Larsen, A.K., Henriques, J.A. (2008). What histone code for DNA repair. Mutatation Research *658*, 259-270.

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

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

Fillingham, J., Recht, J., Silva, A.C., Suter, B., Emili, A., Stagljar, I., Krogan, N.J., Allis, C.D., Keogh, M.-C., and Greenblatt, J.F. (2008). Chaperone Control of the Activity and Specificity of the Histone H3 Acetyltransferase Rtt109. Molecular & Cellular Biology *28*, 4342-4353.

Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M., and Chen, X.S. (2003). The structure and function of MCM from archaeal M. Thermoautotrophicum. Nature Structural & Molecular Biology *10*, 160-167.

Flickinger, R. (2001). Replication timing and cell differentiation. Differentiation *69*, 18-26.

Flores-Rozas, H., Clark, D., and Kolodner, R.D. (2000). Proliferating cell nuclear antigen and Msh2p-Msh6p interact to form an active mispair recognition complex. Nature Genetics *26*, 375-378.

Foiani, M., Pellicioli, A., Lopes, M., Lucca, C., Ferrari, M., Liberi, G., Muzi Falconi, M., and Plevani, P. (2000). DNA damage checkpoints and DNA replication controls in *Saccharomyces cerevisiae*. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *451*, 187-196.

Formosa, T. (2008). FACT and the reorganized nucleosome. Molecular Biosystems *11*, 1085-1093.

Fournier, M.L., Gilmore, J.M., Martin-Brown, S.A., Washburn, M.P. (2007). Multidimensional Separations-Based Shotgun Proteomics. Chemical Reviews *107*, 3654-3686.

Fukui, T., Yamauchi, K., Muroya, T., Akiyama, M., Maki, H., Sugino, A., and Waga, S. (2004). Distinct roles of DNA polymerases delta and epsilon at the replication fork in Xenopus egg extracts. Genes to Cells *9*, 179-191.

Fuss, J., and Linn, S. (2002). Human DNA Polymerase ε Colocalizes with Proliferating Cell Nuclear Antigen and DNA Replication Late, but Not Early, in S Phase. Journal of Biological Chemistry *277*, 8658-8666.

Gafken, P.R., Lampe, P.D. (2006). Methodologies for characterizing phosphoproteins by mass spectrometry. Cell Communication and Adhesion *13*, 249-262.

Gaillard, P.H., Martini, E.M., Kaufman, P.D., Stillman, B., Moustacchi, E., Almouzni, G. (1996). Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. Cell *86*, 887-896.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome

progression complexes at eukaryotic DNA replication forks. Nature Cell Biology *8*, 358-366.

Game, J.C., Kaufman, P.D. (1999). Role of *saccharomyces cerevisiae* chromatin assembly factor-I in repair of ultraviolet radiation damage in vivo. Genetics *151*, 485-497.

Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D.*, et al.* (2007). Organismal Differences in Post-translational Modifications in Histones H3 and H4. Journal of Biological Chemistry *282*, 7641-7655.

Gary, R., Ludwig, D.L., Cornelius, H.L., MacInnes, M.A., and Park, M.S. (1997). The DNA Repair Endonuclease XPG Binds to Proliferating Cell Nuclear Antigen (PCNA) and Shares Sequence Elements with the PCNA-binding Regions of FEN-1 and Cyclin-dependent Kinase Inhibitor p21. Journal of Biological Chemistry *272*, 24522-24529.

Gasch, A.P., Huang, M., Metzner, S., Botstein, D., Elledge, S.J., and Brown, P.O. (2001). Genomic Expression Responses to DNA-damaging Agents and the Regulatory Role of the Yeast ATR Homolog Mec1p. Molecular Biology of the Cell *12*, 2987-3003.

Gasser, R., Koller, T., and Sogo, J.M. (1996). The Stability of Nucleosomes at the Replication Fork. Journal of Molecular Biology *258*, 224-239.

Genschel, J., Modrich, P. (2003). Mechanism of 5-Directed Excision in Human Mismatch Repair. Molecular Cell *12*, 1077-1086.

Gerard, A., Koundrioukoff, S., Ramillon, V., Sergere, J-C., Mailand, N., Quivy, J-P., Almouzni, G. (2006). The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. EMBO Reports *7*, 817-823.

Goren, A., and Cedar, H. (2003). Replicating by the clock. Nature Reviews Molecular Cell Biology *4*, 25-32.

Goshe, M.B. (2006). Characterizing phosphoproteins and phosphoproteomes using mass spectrometry. Briefings in Functional Genomics & Proteomics *4*, 363-376.

Graham, D.R.M., Elliott, S.T., van Eyk, J.E. (2005). Broad-based proteomic strategies: a practical guide to proteomics and functional screening. The Journal of Physiology *563*, 1-9.

Granvogl, B., Plöscher, M., Eichacker, L. (2007). Sample preparation by in-gel digestion for mass spectrometry-based proteomics. Analytical and Bioanalytical Chemistry *389*, 991-1002.

Green, C., and Almouzni, G. (2003). Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair in vivo. The EMBO Journal *22*, 5163 - 5174.

Green, C.M., Almouzni, G. (2002). When repair meets chromatin. First in series on chromatin dynamics. EMBO Reports *3*, 28-33.

Grewal, S.I.S., and Elgin, S.C.R. (2002). Heterochromatin: new possibilities for the inheritance of structure. Current Opinion in Genetics & Development *12*, 178-187.

Grewal, S.I.S., and Moazed, D. (2003). Heterochromatin and Epigenetic Control of Gene Expression. Science *301*, 798-802.

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

Grün, M., Kurganov, A. A., Schacht, S., Schüth, F., Unger, K. K. (1996). Comparison of an ordered mesoporous aluminosilicate, silica, alumina, titania and zirconia in normal-phase high-performance liquid chromatography. Journal of Chromatography A *740*, 1-9.

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

Gu, L., Hong, Y., McCulloch, S., Watanabe, H., and Li, G.-M. (1998). ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. Nucleic Acids Research *26*, 1173-1178.

Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., Kuriyan, J. (1996). Structure of the C-terminal region of p21WAF1/CIP1 complexed with human PCNA. Cell *87*, 297-306.

Gunjan, A., Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *Saccharomyces cerevisiae*. Cell *115*, 537-549.

Hager, J.W. (2004). Q TRAP(TM) mass spectrometer technology for proteomics applications. Drug Discovery Today: TARGETS *3*, 31-36.

Hall, C., Nelson, D.M., Ye, X., Baker, K., DeCaprio, J.A., Seeholzer, S., Lipinski, M., Adams, P.D. (2001). HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel cyclin-cdk2 substrate whose expression blocks S-phase progression. Molecular & Cellular Biology *21*, 1854-1865.

Hall, I.M., Shankaranarayana, G.D., Noma, K.-i., Ayoub, N., Cohen, A., and Grewal, S.I.S. (2002). Establishment and Maintenance of a Heterochromatin Domain. Science *297*, 2232-2237.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.-M., and Zhang, Z. (2007a). Rtt109 Acetylates Histone H3 Lysine 56 and Functions in DNA Replication. Science *315*, 653-655.

Han, J., Zhou, H., Li, Z., Xu, R.-M., and Zhang, Z. (2007b). Acetylation of Lysine 56 of Histone H3 Catalyzed by RTT109 and Regulated by ASF1 Is Required for Replisome Integrity. Journal of Biological Chemistry *282*, 28587-28596.

Han, M., Chang, M., Kim, U.J., Grunstein, M. (1987). Histone H2B repression causes cell-cycle-specific arrest in yeast: Effects on chromosomal segregation, replication, and transcription. Cell *48*, 589-597.

Hanash, S.M. (2000). Biomedical applications of two-dimensional electrophoresis using immobilized pH gradients: Current status. Electrophoresis *21*, 1202-1209.

Haydon, C.E., Eyers, P.A., Aveline-Wolf, L.D., Resing, K.A., Maller, J.L., Ahn, N.G. (2003). Identification of Novel Phosphorylation Sites on *Xenopus laevis* Aurora A and Analysis of Phosphopeptide Enrichment by Immobilized Metal-affinity Chromatography. Molecular & Cellular Proteomics *2*, 1055-1067.

Henneke, G., Koundrioukoff, S., and Hubscher, U. (2003). Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation. Oncogene *22*, 4301-4313.

Hoege, C., Pfander, B., Moldovan, G.-L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141.

Hoek, M., Stillman, B. (2003). Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. Proceedings of the National Academy of Sciences of the United States of America *100*, 12183-12188.

Hopfgartner, G., Varesio, E., Tschäppät, V., Grivet, C., Bourgogne, E., Leuthold, L. A. (2004). Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. Journal of Mass Spectrometry *39*, 845-855.

Hsieh, H.-C., Sheu, C., Shi, F.-K., Li, D.-T. (2007). Development of a titanium dioxide nanoparticle pipette-tip for the selective enrichment of phosphorylated peptides. Journal of Chromatography A *1165*, 128-135.

Hübscher, U., Maga, G., Spadari, S. (2002). Eukaryotic DNA polymerases. Annual Review of Biochemistry *71*, 133-163.

Hübscher, U., Seo, Y-S. (2001). Replication of the lagging strand: a concert of at least 23 polypeptides. Molecules and Cells *12*, 149-157.

Hyrien, O., Marheineke, K., and Goldar, A. (2003). Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. Bioessays *25*, 116-125.

Imai, S.-i., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. Nature *403*, 795-800.

Iyer, R.R., Pluciennik, A., Burdett, V., and Modrich, P.L. (2005). DNA Mismatch Repair: Functions and Mechanisms. Chemical Reviews *106*, 302-323.

Jackson, V. (1990). In vivo studies on the dynamics of histone-DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. Biochemistry *29*, 719-731.

Jaffe, H.V., Pant, H.C. (1998). Characterization of serine and threonine phosphorylation sites in β-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. Biochemistry *37*, 16211-16224.

Jaitly, G., Bonneil, E., Jaitly, N., Eng, K., Pomiès, C., Thibault, P. (2007). Comprehensive profiling of unlabeled peptide ions from large-scale proteomics experiments using one and two dimensional nanoLC-MS/MS. Paper presented at: Proceedings of the 55th ASMS Conference on Mass Spectrometry (Indianapolis, IN).

Jensen, S.S., Larsen, M. R. (2007). Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. Rapid Communications in Mass Spectrometry *21*, 3635-3645.

Jenuwein, T., Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

Kannouche, P.L., Wing, J., and Lehmann, A.R. (2004). Interaction of Human DNA Polymerase with Monoubiquitinated PCNA: A Possible Mechanism for the Polymerase Switch in Response to DNA Damage. Molecular Cell *14*, 491-500.

Karthikeyan, R., Vonarx, E.J., Straffon, A.F.L., Simon, M., Faye, G., and Kunz, B.A. (2000). Evidence from mutational specificity studies that yeast DNA polymerases [delta] and [epsilon] replicate different DNA strands at an intracellular replication fork. Journal of Molecular Biology *299*, 405-419.

Kaufman, P., Kobayashi, R., and Stillman, B. (1997). Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. Genes & Development *11*, 345 - 357.

Kaufman, P.D., Cohen, J.L., Osley, M.A. (1998). Hir Proteins Are Required for Position-Dependent Gene Silencing in *Saccharomyces cerevisiae* in the Absence of Chromatin Assembly Factor I. Molecular & Cellular Biology *18*, 4793-4806.

Kaufman, P.D., Kobayashi, R., Naama, K., Stillman, B. (1995). The p150 and p60 subunits of chromatin assembly factor 1: A molecular link between newly synthesized histones and DNA replication. Cell *81*, 1105-1114.

Kaufman, P.D., Kobayashi, R., Stillman, B (1997). Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. Genes & Development *11*, 345-357.

Kedar, P.S., Kim, S.-J., Robertson, A., Hou, E., Prasad, R., Horton, J.K., and Wilson, S.H. (2002). Direct Interaction between Mammalian DNA Polymerase β and Proliferating Cell Nuclear Antigen. Journal of Biological Chemistry *277*, 31115-31123.

Keller, C., Krude, T. (2000). Requirement of Cyclin/Cdk2 and Protein Phosphatase 1 Activity for Chromatin Assembly Factor 1-dependent Chromatin Assembly during DNA Synthesis. Journal of Biological Chemistry *275*, 35512-35521.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259-272.

Kim, U.J., Han, M., Kayne, P., Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of *Saccharomyces cerevisiae*. The EMBO Journal *7*, 2211-2219.

Kimura, A., Umehara, T., and Horikoshi, M. (2002). Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. Nature Genetics *32*, 370-377.

Klungland, A., Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). The EMBO Journal *16*, 3341-3348.

Kokubu, M., Ishihama, Y., Sato, T., Nagasu, T., Oda, Y. (2005). Specificity of Immobilized Metal Affinity-Based IMAC/C18 Tip Enrichment of Phosphopeptides for Protein Phosphorylation Analysis. Analytical Chemistry *77*, 5144-5154.

Krawitz, D.C., Kama, T., Kaufman, P.D. (2002). Chromatin Assembly Factor I Mutants Defective for PCNA Binding Require Asf1/Hir Proteins for Silencing. Molecular & Cellular Biology *22*, 614-625.

Krishna, T.S.R., Kong, X.-P., Gary, S., Burgers, P.M., and Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell *79*, 1233-1243.

Krude, T. (1995). Chromatin Assembly Factor 1 (CAF-1) Colocalizes with Replication Foci in HeLa Cell Nuclei. Experimental Cell Research *220*, 304-311.

Krude, T. (2000). Initiation of Human DNA Replication in Vitro Using Nuclei from Cells Arrested at an Initiation-competent State. Journal of Biological Chemistry *275*, 13699-13707.

Krude, T., Jackman, M., Pines, J., Laskey, R.A. (1997). Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. Cell *88*, 109-119.

Kuo, M.-H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269-272.

Kurdistani, S.K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nature Reviews Molecular Cell Biology *4*, 276-284.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes & Development *16*, 2893-2905.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lamour, V., Lécluse, Y., Desmaze, C., Spector, M., Bodescot, M., Aurias, A., Osley, M.A., Lipinski, M. (1995). A human homolog of the *Saccharomyces cerevisiae* HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. Human Molecular Genetics *4*, 791-799.

Lane, C.S. (2005). Mass spectrometry-based proteomics in the life sciences. Cellular and Molecular Life Sciences *62*, 848-869.

Lange, V., Picotti, P., Domon, B., Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. Molecular Systems Biology *4*, 1-14.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., Jørgensen, T.J.D. (2005). Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. Molecular & Cellular Proteomics *4*, 873-886.

Lau, P.J., and Kolodner, R.D. (2003). Transfer of the MSH2·MSH6 Complex from Proliferating Cell Nuclear Antigen to Mispaired Bases in DNA. Journal of Biological Chemistry *278*, 14-17.

Lawrence, C. (1994). The RAD6 DNA repair pathway in *Saccharomyces cerevisiae*: What does it do, and how does it do it? Bioessays *16*, 253-258.

Le Blanc, J.C.Y., Hager, J. W., Ilisiu, A. M. P., Hunter, C., Zhong, F., Chu, I. (2003). Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications. Proteomics *3*, 859-869.

Lee, J., Chastain, P.D., Kusakabe, T., Griffith, J.D., Richardson, CC. (1998). Coordinated leading and lagging strand DNA synthesis on a minicircular template. Molecular Cell *1*, 1001-1010.

Lewis, L.K., Karthikeyan, G., Cassiano, J., Resnick, M.A. (2005). Reduction of nucleosome assembly during new DNA synthesis impairs both major pathways of double-strand break repair. Nucleic Acids Research *33*, 4928-4939.

Li, F., Chen, J., Izumi, M., Butler, M.C., Keezer, S.M., and Gilbert, D.M. (2001). The replication timing program of the Chinese hamster β-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. The Journal of Cell Biology *154*, 283-292.

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

Li, X., Li, J., Harrington, J., Lieber, M.R., and Burgers, P.M.J. (1995). Lagging Strand DNA Synthesis at the Eukaryotic Replication Fork Involves Binding and Stimulation of FEN-1 by Proliferating Cell Nuclear Antigen. Journal of Biological Chemistry *270*, 22109-22112.

Lichten, M., and Goldman, A.S.H. (1995). Meiotic Recombination Hotspots. Annual Review of Genetics *29*, 423-444.

Liebler, D.C. (2002). Introduction to Proteomics; Tools for the New Biology (Totowa, New Jersey, Humana Press).

Linger, J., Tyler, J.K. (2005). The Yeast Histone Chaperone Chromatin Assembly Factor 1 Protects Against Double-Strand DNA-Damaging Agents. Genetics *171*, 1513-1522.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., Yates, J.R. (1999). Direct analysis of protein complexes using mass spectrometry. Nature Biotechnology *17*, 676-682.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001). Correlation Between Histone Lysine Methylation and Developmental Changes at the Chicken beta - Globin Locus. Science *293*, 2453-2455.

Liu, H., Lin, D., Yates, J.R. 3rd. (2002). Multidimensional separations for protein/peptide analysis in the post-genomic era. Biotechniques *32*, 898, 900, 902 passim.

Loppin, B., Bonnefoy, E., Anselme, C., Laurencon, A., Karr, T.L., Couble, P. (2005). The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature *437*, 1386-1390.

Lorain, S., Quivy, J.-P., Monier-Gavelle, F., Scamps, C., Lecluse, Y., Almouzni, G., and Lipinski, M. (1998). Core Histones and HIRIP3, a Novel Histone-Binding Protein, Directly Interact with WD Repeat Protein HIRA. Molecular & Cellular Biology *18*, 5546-5556.

Lowndes, N.F., and Murguia, J.R. (2000). Sensing and responding to DNA damage. Current Opinion in Genetics & Development *10*, 17-25.

Loyola, A., and Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *1677*, 3-11.

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

Ma, X.-J., Wu, J., Altheim, B.A., Schultz, M.C., and Grunstein, M. (1998). Deposition-related sites K5/K12 in histone H4 are not required for nucleosome deposition in yeast. Proceedings of the National Academy of Sciences of the United States of America *95*, 6693-6698.

Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom. *17*, 2337-2342.

Maga, G., and Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. Journal of Cell Science *116*, 3051-3060.

Majka, J., Burgers, P.M.J., and Kivie, M. (2004). The PCNA-RFC Families of DNA Clamps and Clamp Loaders. In Progress in Nucleic Acid Research and Molecular Biology (Academic Press), pp. 227-260.

Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., Horning, S. (2006a). Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. Analytical Chemistry *78*, 2113-2120.

Makarov, A., Denisov, E., Lange, O., Horning, S. (2006b). Dynamic Range of Mass Accuracy in LTQ Orbitrap Hybrid Mass Spectrometer. Journal of the American Society for Mass Spectrometry *17*, 977-982.

Maldonado, E., Hampsey, M., and Reinberg, D. (1999). Repression: Targeting the Heart of the Matter. Cell *99*, 455-458.

Mann, M., Jensen, O.N. (2003). Proteomic analysis of post-translational modifications. Nature Biotechnology *21*, 255-261.

Mann, M., Ong, S-E., Grønborg, M., Steen, H., Jensen, O. N., Pandey, A. (2002). Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends in Biotechnology *20*, 261-268.

Marheineke, K., Krude, T. (1998). Nucleosome Assembly Activity and Intracellular Localization of Human CAF-1 Changes during the Cell Division Cycle. Journal of Biological Chemistry *273*, 15279-15286.

Martini, E., Roche, D.M.J., Marheineke, K., Verreault, A., Almouzni, G. (1998). Recruitment of Phosphorylated Chromatin Assembly Factor 1 to Chromatin after UV Irradiation of Human Cells. The Journal of Cell Biology *143*, 563-575.

Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. Current Opinion in Cell Biology *14*, 692-699.

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

Masumoto, H., Muramatsu, S., Kamimura, Y., and Araki, H. (2002). S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. Nature *415*, 651-655.

Matsumiya, S., Ishino, S., Ishino, Y., and Morikawa, K. (2002). Physical interaction between proliferating cell nuclear antigen and replication factor C from Pyrococcus furiosus. Genes to Cells *7*, 911-922.

Matsumoto, T., Eki, T., and Hurwitz, J. (1990). Studies on the initiation and elongation reactions in the simian virus 40 DNA replication system. Proceedings of the National Academy of Sciences of the United States of America *87*, 9712-9716.

McKittrick, E., Gafken, P.R., Ahmad, K., Henikoff, S. (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proceedings of the National Academy of Sciences of the United States of America *101*, 1525-1530.

McLachlin, D.T., Chait, B.T. (2003). Improved β-Elimination-Based Affinity Purification Strategy for Enrichment of Phosphopeptides. Analytical Chemistry *75*, 6826-6836.

McLafferty, F.W. (2001). Tandem mass spectrometric analysis of complex biological mixtures. International Journal of Mass Spectrometry *212*, 81-87.

McNairn, A.J., and Gilbert, D.M. (2003). Epigenomic replication: Linking epigenetics to DNA replication. Bioessays *25*, 647-656.

Medzihradszky, K.F., Darula, Z., Perlson, E., Fainzilber, M., Chalkley, R.J., Ball, H., Greenbaum, D., Bogyo, M., Tyson, D.R., Bradshaw, R.A., Burlingame, A.L. (2004). O-Sulfonation of Serine and Threonine. Molecular & Cellular Proteomics *3*, 429-440.

Meeks-Wagner, D., Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell *44*, 43-52.

Meijsing, S.H., Ehrenhofer-Murray, A.E. (2001). The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in *Saccharomyces cerevisiae*. Genes & Development *15*, 3169-3182.

Mikesh, L.M., Ueberheide, B., Chi, A., Coon, J.J., Syka, J.E.P., Shabanowitz, J., Hunt, D.F. (2006). The utility of ETD mass spectrometry in proteomic analysis. Biochimica and Biophysica Acta *12*, 1811-1822.

Milutinovic, S., Zhuang, Q., and Szyf, M. (2002). Proliferating Cell Nuclear Antigen Associates with Histone Deacetylase Activity, Integrating DNA Replication and Chromatin Modification. Journal of Biological Chemistry *277*, 20974-20978.

Moazed, M. (2001). Common themes in mechanisms of gene silencing. Molecular Cell *8*, 489-498.

Moggs, J.G., Grandi, P., Quivy, J-P., Jonsson, Z.O., Hubscher, U., Becker, P.B., Almouzni, G. (2000). A CAF-1-PCNA-Mediated Chromatin Assembly Pathway Triggered by Sensing DNA Damage. Molecular & Cellular Biology *20*, 1206-1218.

Moldovan, G.-L., Pfander, B., and Jentsch, S. (2007). PCNA, the Maestro of the Replication Fork. Cell *129*, 665-679.

Moldovan, G.-L., Pfander, B., Jentsch, S. (2006). PCNA controls establishment of sister chromatid cohesion during S phase. Molecular Cell *23*, 723-732.

Monson, E.K., de Bruin, D., Zakian, V.A. (1997). The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. Proceedings of the National Academy of Sciences of the United States of America *94*, 13081-13086.

Montecucco, A., Rossi, R., Levin, D.S., Gary, R., Park, M.S., Motycka, T.A., Ciarrocchi, G., Villa, A., Biamonti, G., and Tomkinson, A.E. (1998). DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. The EMBO Journal *17*, 3786-3795.

Moran, L., Norris, D., and Osley, M.A. (1990). A yeast H2A-H2B promoter can be regulated by changes in histone gene copy number. Genes & Development *4*, 752-763.

Morgan, D.O. (2007). The cell cycle: principles of control (London, England, New Science Press Ltd).

Mossi, R., and Hübscher, U. (1998). Clamping down on clamps and clamp loaders. European Journal of Biochemistry *254*, 209-216.

Murzina, N., Verreault, A., Laue, E., Stillman, B. (1999). Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. Cell *4*, 529-540.

Murzina, N.V., Pei, X.Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., Laue, E.D. (2008). Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. Structure *16*, 1077-1085.

Myung, K., Pennaneach, V., Kats, E.S., and Kolodner, R.D. (2003). *Saccharomyces cerevisiae* chromatin-assembly factors that act during DNA replication function in the maintenance of genome stability. Proceedings of the National Academy of Sciences of the United States of America *100*, 6640-6645.

Nabatiyan, A., Krude, T. (2004). Silencing of Chromatin Assembly Factor 1 in Human Cells Leads to Cell Death and Loss of Chromatin Assembly during DNA Synthesis. Molecular & Cellular Biology *24*, 2853-2862.

Nabatiyan, A., Szuts, D., and Krude, T. (2006). Induction of CAF-1 Expression in Response to DNA Strand Breaks in Quiescent Human Cells. Molecular & Cellular Biology *26*, 1839-1849.

Nakayama, J.-i., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I.S. (2001). Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. Science *292*, 110-113.

Nakayama, T., Nishioka, K., Dong, Y.-X., Shimojima, T., Hirose, S. (2007). *Drosophila* GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. Genes & Development *21*, 552-561.

Neer, E.J., Schmidt, C.J., Nambudripad, R., and Smith, T.F. (1994). The ancient regulatory-protein family of WD-repeat proteins. Nature *371*, 297-300.

Nelson, D.M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G.D., Yen, T.J., Harper, J.W., Adams, P.D. (2002). Coupling of DNA Synthesis and Histone Synthesis in S Phase Independent of Cyclin/cdk2 Activity. Molecular & Cellular Biology *22*, 7459-7472.

Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis *9*, 255-262.

Nicolas, E., Ait-Si-Ali, S., and Trouche, D. (2001). The histone deacetylase HDAC3 targets RbAp48 to the retinoblastoma protein. Nucleic Acids Research *29*, 3131-3136.

Oda, Y., Nagasu, T., Chait, B.T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nature Biotechnology *19*, 379-382.

Okano, S., Lan, L., Caldecott, K.W., Mori, T., and Yasui, A. (2003). Spatial and Temporal Cellular Responses to Single-Strand Breaks in Human Cells. Molecular & Cellular Biology *23*, 3974-3981.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nature Reviews Molecular Cellular Biology *4*, 809-814.

Olsen, J.V., Ong, S.-E., Mann, M. (2004). Trypsin cleaves exclusively c-terminal to arginine and lysine residues. Molecular & Cellular Proteomics *3*, 608-614.

Osley, M.A., Lycan, D. (1987). Trans-acting regulatory mutations that alter transcription of *Saccharomyces cerevisiae* histone genes. Molecular & Cellular Biology *7*, 4204-4210.

Osley, M.A., Shen, X. (2006). Altering nucleosomes during DNA double-strand break repair. Trends in Genetics *22*, 671-677.

Pandey, M., Syed, S., Donmez, I., Ha, T., Patel, S. (2003). Coordinating DNA replication by means of priming loop and differential synthesis rate. Nature *462*, 940-943.

Patel, V.J., Thalassinos, K., Slade, S.E., Connolly, J.B., Crombie, A., Murrell, J.C., Scrivens, J.H. (2009). A Comparison of Labeling and Label-Free Mass Spectrometry-Based Proteomics Approaches. Journal of Proteome Research *8*, 3752-3759.

Paulovich, A.G., Hartwell, L.H. (1995). A checkpoint regulates the rate of progression through S phase in *Saccharomyces cerevisiae* in response to DNA damage. Cell *8*, 841-847.

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

Picotti, P., Bodenmiller, B., Mueller, L.N., Domon, B., Aebersold, R. (2009). Full dynamic range proteome analysis of *Saccharomyces cerevisiae* by targeted proteomics. Cell *138*, 795-806.

Picotti, P., Lam, H., Campbell, D., Deutsch, E.W., Mirzaei, H., Ranish, J., Domon, B., Aebersold, R. (2008). A database of mass spectrometric assays for the yeast proteome. Nature Methods *5*, 913-914.

Pinkse, M.W.H., Uitto, P.M., Hilhorst, M.J., Ooms, B., Heck, A.J.R. (2004). Selective Isolation at the Femtomole Level of Phosphopeptides from Proteolytic Digests Using 2D-NanoLC-ESI-MS/MS and Titanium Oxide Precolumns. Analytical Chemistry *76*, 3935-3943.

Polo, S.E., Roche, D., Almouzni, G. (2006). New histone incorporation marks sites of UV repair in human cells. Cell *127*, 481-493.

Posewitz, M.C., Tempst, P. (1999). Immobilized Gallium(III) Affinity Chromatography of Phosphopeptides. Analytical Chemistry *71*, 2883-2892.

Prakash, S., Sung, P., and Prakash, L. (1993). DNA Repair Genes and Proteins of *Saccharomyces Cerevisiae*. Annual Review of Genetics *27*, 33-70.

Prelich, G., Kostura, M., Marshak, D.R., Mathews, M.B., and Stillman, B. (1987). The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in vitro. Nature *326*, 471-475.

Prelich, G., and Stillman, B. (1988). Coordinated leading and lagging strand synthesis during SV40 DNA replication in vitro requires PCNA. Cell *53*, 117-126.

Prochasson, P., Florens, L., Swanson, S.K., Washburn, M.P., Workman, J.L. (2005). The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. Genes & Development *19*, 2534-2539.

Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E., Kunkell, T.A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science *317*, 127-130.

Qian, Y.-W., and Lee, E.Y.H.P. (1995). Dual Retinoblastoma-binding Proteins with Properties Related to a Negative Regulator of Ras in Yeast. Journal of Biological Chemistry *270*, 25507-25513.

Qian, Y.-W., Wang, Y.-C.J., Hollingsworth, R.E., Jones, D., Ling, N., and Lee, E.Y.H.P. (1993). A retinoblastoma-binding protein related to a negative regulator of Ras in yeast. Nature *364*, 648-652.

Qian, Z., Huang, H., Hong, J.Y., Burck, C.L., Johnston, S.D., Berman, J., Carol, A., Liebman, S.W. (1998). Yeast Ty1 retrotransposition is stimulated by a synergistic interaction between mutations in chromatin assembly factor I and histone regulatory proteins. Molecular & Cellular Biology *18*, 4783-4792.

Quivy, J.-P., Grandi, P., and Almouzni, G. (2001). Dimerization of the largest subunit of chromatin assembly factor 1: importance in vitro and during *Xenopus* early development. The EMBO Journal *20*, 2015-2027.

Quivy, J.-P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y., and Almouzni, G. (2004). A CAF-1 dependent pool of HP1 during heterochromatin duplication. The EMBO Journal *23*, 3516-3526.

Rabilloud, T. (2002). Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. Proteomics *2*, 3-10.

Raghuraman, M.K., Brewer, B.J., and Fangman, W.L. (1997). Cell Cycle-Dependent Establishment of a Late Replication Program. Science *276*, 806-809.

Raghuraman, M.K., Winzeler, E.A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D.J., Davis, R.W., Brewer, B.J., and Fangman, W.L. (2001). Replication Dynamics of the Yeast Genome. Science *294*, 115-121.

Rao, H., Marahrens, Y., and Stillman, B. (1994). Functional conservation of multiple elements in yeast chromosomal replicators. Molecular & Cellular Biology *14*, 7643-7651.

Rao, H., and Stillman, B. (1995). The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proceedings of the National Academy of Sciences of the United States of America *92*, 2224-2228.

Ray-Gallet, D., Quivy, J.-P.., Scamps, C.,Martini, E.M.-D., Lipinski, M., Almouzni, G. (2003). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Molecular Cell *9*, 1091-1100.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D.*, et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Recht, J., Dunn, B., Raff, A., and Osley, M. (1996). Functional analysis of histones H2A and H2B in transcriptional repression in *Saccharomyces cerevisiae*. Molecular & Cellular Biology *16*, 2545-2553.

Reinders, J., Sickmann, A. (2005). State-of-the-art in phosphoproteomics. Proteomics *5*, 4052-4061.

Rhind, N., and Russell, P. (1998). Mitotic DNA damage and replication checkpoints in yeast. Current Opinion in Cell Biology *10*, 749-758.

Richards, E.J., Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489-500.

Ridgway, P., and Almouzni, G. (2000). CAF-1 and the inheritance of chromatin states: at the crossroads of DNA replication and repair. Journal of Cell Science *113*, 2647-2658.

Roberts, C., Sutherland, H.F., Farmer, H., Kimber, W., Halford, S., Carey, A., Brickman, J.M., Wynshaw-Boris, A., and Scambler, P.J. (2002). Targeted Mutagenesis of the Hira Gene Results in Gastrulation Defects and Patterning Abnormalities of Mesoendodermal Derivatives Prior to Early Embryonic Lethality. Molecular & Cellular Biology *22*, 2318-2328.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234*, 364-368.

Rouse, J., and Jackson, S.P. (2002). Interfaces Between the Detection, Signaling, and Repair of DNA Damage. Science *297*, 547-551.

Ruiz-García, A.B., Sendra, R., Galiana, M., Pamblanco, M., Pérez-Ortín, J.E., and Tordera, V. (1998). HAT1 and HAT2 Proteins Are Components of a Yeast Nuclear Histone Acetyltransferase Enzyme Specific for Free Histone H4. Journal of Biological Chemistry *273*, 12599-12605.

Rusche, L.N., Kirchmaier, A.L. Jasper, R. (2003). The establishment, inheritance and function of silenced chromatin in *Saccharomyces cerevisiae*. Annual Review of Biochemistry *72*, 481-516.

Salazar, C., Höfer, T. (2009). Multisite protein phosphorylation – from molecular mechanisms to kinetic models. FEBS Journal *276*, 3177-3198.

Schneider, J., Bajwa, P., Johnson, F.C., Bhaumik, S.R., and Shilatifard, A. (2006). Rtt109 Is Required for Proper H3K56 Acetylation. Journal of Biological Chemistry *281*, 37270-37274.

Schroeder, M.J., Shabanowitz, J., Schwartz, J.C., Hunt, D.F., Coon, J.J. (2004). A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. Analytical Chemistry *76*, 3590-3598.

Scigelova, M., Makarov, A. (2006). Orbitrap Mass Analyzer – Overview and Applications in Proteomics. Proteomics *6*, 16-21.

Sclafani, R.A., Holzen, T. M. (2007). Cell cycle regulation of DNA replication. Annual Review of Genetics *41*, 237-280.

Sclafani, R.A., Tecklenburg, M., and Pierce, A. (2002). The mcm5-bob1 Bypass of Cdc7p/Dbf4p in DNA Replication Depends on Both Cdk1-Independent and Cdk1-Dependent Steps in *Saccharomyces cerevisiae*. Genetics *161*, 47-57.

Shamoo, Y., Steitz, T.A. (1999). Building a replisome from interacting pieces. Cell *99*, 155-166.

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

Sharp, J.A., Franco, A.A., Osley, M.A., Kaufman, P.D. (2002). Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in *Saccharomyces cerevisiae*. Genes & Development *16*, 85-100.

Sharp, J.A., Krawitz, D.C., Gardner, K.A., Fox, C.A., Kaufman, P.D. (2003). The budding yeast silencing protein Sir1 is a functional component of centromeric chromatin. Genes & Development *17*, 2356-2361.

Shen, Y., Smith, R. D. (2002a). Proteomics based on high-efficiency capillary separations. Electrophoresis *23*, 3106-3124.

Shen, Y., Zhao, R., Berger, S.J., Anderson, G.A., Rodriguez, N., Smith, R.D. (2002b). High-Efficiency Nanoscale Liquid Chromatography Coupled On-Line with Mass Spectrometry Using Nanoelectrospray Ionization for Proteomics. Analytical Chemistry *74*, 4235-4249.

Sherwood, P.W., Osley, M. A. (1991). Histone Regulatory (hir) Mutations Suppress {delta} Insertion Alleles in *Saccharomyces cerevisiae*. Genetics *128*, 729-738.

Sherwood, P.W., Tsang, S.V., Osley, M.A. (1993). Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in *Saccharomyces cerevisiae*. Molecular & Cellular Biology *13*, 28-38.

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

Shibahara, K.-i., Verreault, A., and Stillman, B. (2000). The N-terminal domains of histones H3 and H4 are not necessary for chromatin assembly factor-1- mediated

nucleosome assembly onto replicated DNA in vitro. Proceedings of the National Academy of Sciences of the United States of America *97*, 7766-7771.

Shibahara, K., Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Sims Iii, R.J., Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends in Genetics *19*, 629-639.

Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., and Gottschling, D.E. (1998). Identification of High-Copy Disruptors of Telomeric Silencing in *Saccharomyces cerevisiae*. Genetics *150*, 613-632.

Sklenar, A., and Parthun, M. (2004). Characterization of yeast histone H3-specific type B histone acetyltransferases identifies an ADA2-independent Gcn5p activity. BMC Biochemistry *5*, 11.

Smith, J.S., Caputo, E., and Boeke, J.D. (1999a). A Genetic Screen for Ribosomal DNA Silencing Defects Identifies Multiple DNA Replication and Chromatin-Modulating Factors. Molecular & Cellular Biology *19*, 3184-3197.

Smith, S., and Stillman, B. (1991). Immunological characterization of chromatin assembly factor I, a human cell factor required for chromatin assembly during DNA replication in vitro. Journal of Biological Chemistry *266*, 12041-12047.

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

Smith, T.F., Gaitatzes, C., Saxena, K., and Neer, E.J. (1999b). The WD repeat: a common architecture for diverse functions. Trends in Biochemical Sciences *24*, 181-185.

Sobel, R.E., Cook, R.G., Perry, C.A., Annunziato, A.T., and Allis, C.D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. Proceedings of the National Academy of Sciences of the United States of America *92*, 1237-1241.

Sogo, J.M., Stahl, H., Koller, T., and Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes : The replication fork, core histone segregation and terminal structures. Journal of Molecular Biology *189*, 189-204.

Speck, C., and Stillman, B. (2007). Cdc6 ATPase Activity Regulates ORC·Cdc6 Stability and the Selection of Specific DNA Sequences as Origins of DNA Replication. Journal of Biological Chemistry *282*, 11705-11714.

Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H., Cardoso, M.C. (2002). DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. Cell *10*, 1355-1365.

Steen, H., Jebanathirajah, J.A., Rush, J., Morrice, N., Kirschner, M.W. (2006). Phosphorylation Analysis by Mass Spectrometry. Molecular & Cellular Proteomics *5*, 172-181.

Steen, H., Mann, M. (2004). The abc's (and xyz's) of peptide sequencing. Nature Reviews Molecular Cellular Biology *5*, 699-711.

Stelter, P., and Ulrich, H.D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature *425*, 188-191.

Stoimenov, I., Helleday, T. (2009). PCNA on the crossroad of cancer. Biochemical Society Transactions *037*, 605-613.

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

Sugiyama, N., Masuda, T., Shinoda, K., Nakamura, A., Tomita, M., Ishihama, Y (2007). Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. Molecular & Cellular Proteomics *6*, 1103-1109.

Suka, N., Luo, K., and Grunstein, M. (2002). Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. Nature Genetics *32*, 378-383.

Suka, N., Suka, Y., Carmen, A.A., Wu, J., Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Molecular Cell *8*, 473-479.

Swaney, D.L., Wenger, C.D., Coon, J.J. (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. Journal of Proteome Research *9*, 1323-1329.

Swarup, G. (1998). Regulation of cellular and molecular functions by protein phosphorylation. Resonance *3*, 70-78.

Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F. (2004) Peptide and protein sequence analysis by electro transfer dissociation mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America *101*, 9528-9533.

Taddei, A., Roche, D., Sibarita, J.-B., Turner, B.M., and Almouzni, G. (1999). Duplication and Maintenance of Heterochromatin Domains. The Journal of Cell Biology *147*, 1153-1166.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. Cell *116*, 51-61.

Takahashi, T.S., Wigley, D.B., and Walter, J.C. (2005). Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. Trends in Biochemical Sciences *30*, 437-444.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. Genes & Development *17*, 1153-1165.

Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Scheid, O.M., Shibahara, K.-i*., et al.* (2004). BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. Genes & Development *18*, 782-793.

Tan, B.C.-M., Chien, C.-T., Hirose, S., and Lee, S.-C. (2006). Functional cooperation between FACT and MCM helicase facilitates initiation of chromatin DNA replication. The EMBO Journal *25*, 3975-3985.

Tan, C.K., Castillo, C., So, A.G., and Downey, K.M. (1986). An auxiliary protein for DNA polymerase-delta from fetal calf thymus. Journal of Biological Chemistry *261*, 12310-12316.

Tanaka, S., and Diffley, J.F.X. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. Nature Cell Biology *4*, 198-207.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328-332.

Tang, Y., Holbert, M.A., Wurtele, H., Meeth, K., Rocha, W., Gharib, M., Jiang, E., Thibault, P., Verreault, A., Cole, P.A.*, et al.* (2008). Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP. Nature Structural & Molecular Biology *15*, 738-745.

Tanny, J.C., Dowd, G.J., Huang, J., Hilz, H., Moazed, D. (1999). An enzymatic activity in the yeast Sir2 protein that is essential for gene silencing. Cell *99*, 735-745.

Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A Mammalian Histone Deacetylase Related to the Yeast Transcriptional Regulator Rpd3p. Science *272*, 408-411.

Tchenio, T., Casella, J.-F., and Heidmann, T. (2001). A Truncated Form of the Human CAF-1 p150 Subunit Impairs the Maintenance of Transcriptional Gene Silencing in Mammalian Cells. Molecular & Cellular Biology *21*, 1953-1961.

Theis, J.F., and Newlon, C.S. (1994). Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. Molecular & Cellular Biology *14*, 7652-7659.

Thingholm, T.E., Jensen, O. N., Larsen, M. R. (2009). Analytical strategies for phosphoproteomics. Proteomics *9*, 1451-1468.

Thingholm, T.E., Jorgensen, T.J.D., Jensen, O.N., Larsen, M.R. (2006). Highly selective enrichment of phosphorylated peptides using titanium dioxide. Nature Protocols *1*, 1929-1935.

Thiru, A., Nietlispach, D., Mott, H.R., Okuwaki, M., Lyon, D., Nielsen, P.R., Hirshberg, M., Verreault, A., Murzina, N.V., and Laue, E.D. (2004). Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. The EMBO Journal *23*, 489-499.

Trinidad, J.C., Specht, C.G., Thalhammer, A., Schoepfer, R., Burlingame, A.L. (2006). Comprehensive Identification of Phosphorylation Sites in Postsynaptic Density Preparations. Molecular & Cellular Proteomics *5*, 914-922.

Tsai, C.-F., Wang, Y.-T., Chen, Y.-R., Lai, C.-Y., Lin, P.-Y., Pan, K.-T., Chen, J.-Y., Khoo, K.-H., Chen, Y.-J. (2008). Immobilized Metal Affinity Chromatography Revisited: pH/Acid Control toward High Selectivity in Phosphoproteomics. Journal of Proteome Research *7*, 4058-4069.

Tsubota, T., Berndsen, C.E., Erkmann, J.A., Smith, C.L., Yang, L., Freitas, M.A., Denu, J.M., Kaufman, P.D (2007). Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes Molecular Cell *25*, 703-712.

Tyler, J. (2002). Chromatin assembly. Cooperation between histone chaperones and ATP-dependent nucleosome remodeling machines. European Journal of Biochemistry *269*, 2268 - 2274.

Tyler, J.K., Adams, C.R., Chen, S.-R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature *402*, 555-560.

Tyler, J.K., and Kadonaga, J.T. (1999). The "Dark Side" of Chromatin Remodeling: Repressive Effects on Transcription. Cell *99*, 443-446.

Ulintz, P.J., Yocum, A.K., Bodenmiller, B., Aebersold, R., Andrews, P.C., Nesvizhskii, A.I. (2008). Comparison of MS2-only, MSA, and MS2/MS3 methodologies for phosphopeptide identification. Journal of Proteome Research *8*, 887-899.

Umar, A., Buermeyer, A.B., Simon, J.A., Thomas, D.C., Clark, A.B., Liskay, R.M., Kunkel, T.A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell *87*, 65-73.

Unwin, R.D., Griffiths, J.R., Leverentz, M.K., Grallert, A., Hagan, I.M., Whetton, A.D. (2005). Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity. Molecular & Cellular Proteomics *4*, 1134-1144.

Verreault, A., Kaufman, P.D., Kobyashi, R., Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. Cell *87*, 95-104.

Verreault, A., Kayfman, P.D., Kobayashi, R., Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. Current Biology *8*, 96-108.

Waga, S., Masuda, T., Takisawa, H., and Sugino, A. (2001). DNA polymerase is required for coordinated and efficient chromosomal DNA replication in *Xenopus* egg extracts. Proceedings of the National Academy of Sciences of the United States of America *98*, 4978-4983.

Waga, S., and Stillman, B. (1998). The DNA replication fork in eukaryotic cells. Annual Review of Biochemistry *67*, 721-751.

Wallace, J., and Orr-Weaver, T. (2005). Replication of heterochromatin: insights into mechanisms of epigenetic inheritance. Chromosoma *114*, 389-402.

Walter, J., Newport, J. (2000). Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase. Molecular Cell *5*, 617-627.

Walther, T.C., Mann, M. (2010). Mass spectrometry–based proteomics in cell biology. The Journal of Cell Biology *190*, 491-500.

Wang, J., Zhang, Y., Jiang, H., Cai, Y., Qian, X. (2006). Phosphopeptide detection using automated online IMAC-capillary LC-ESI-MS/MS. Proteomics *6*, 404-411.

Wang, J.C. (2002). Cellular roles of DNA topoisomerases: a molecular perspective. Nature Review Molecular Cellular Biology *3*, 430-440.

Wang, S.-C., Nakajima, Y., Yu, Y.-L., Xia, W., Chen, C.-T., Yang, C.-C., McIntush, E.W., Li, L.-Y., Hawke, D.H., Kobayashi, R.*, et al.* (2006). Tyrosine phosphorylation controls PCNA function through protein stability. Nature Cell Biology *8*, 1359-1368.

Wang, X., Li, X., Li, Y. (2007). A modified coomassie brilliant blue staining method at nanogram sensitivity compatible with proteomic analysis. Biotechnology Letters *29*, 1599-1603.

Warbrick, E. (1998a). PCNA binding through a conserved motif. BioEssays *20*, 195-199.

Warbrick, E., Lane, D.P., Glover, D.M., and Heatherington, W. (1998b). PCNA binding proteins in *Drosophila melanogaster*: the analysis of a conserved PCNA binding domain. Nucleic Acids Research *26*, 3925-3932.

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

Waters, L.S., Minesinger, B.K., Wiltrout, M.E., D'Souza, S., Woodruff, R.V., and Walker, G.C. (2009). Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. Microbiology & Molecular Biology Reviews *73*, 134-154.

Watson, J.T., Sparkman, O.D. (2007). Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation, 4th edition edn (Chichester, England, John Wiley & Sons).

Weinert, T.A., Hartwell, L.H. (1988). The RAD9 gene controls the cell cycle response to DNA damage in *Saccharomyces cerevisiae*. Science *15*, 317-322.

Wilm, M., Mann, M. (1996). Analytical Properties of the Nanoelectrospray Ion Source. Analytical Chemistry *68*, 1-8.

Wilming, L.G., Snoeren, C.A., van Rijswijk, A., Grosveld, F., Meijers, C. (1997). The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in human CATCH22 patients. Human Molecular Genetics *6*, 247-258.

Wold, M.S. (1997). REPLICATION PROTEIN A:A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaryotic DNA Metabolism. Annual Review of Biochemistry *66*, 61-92.

Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D.A., White, F.M. (2007). Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. Proceedings of the National Academy of Sciences *104*, 5860-5865.

Wolters, D.A., Washburn, M.P., Yates, J.R. (2001). An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. Analytical Chemistry *73*, 5683-5690.

Wyrick, J.J., Aparicio, J.G., Chen, T., Barnett, J.D., Jennings, E.G., Young, R.A., Bell, S.P., and Aparicio, O.M. (2001). Genome-Wide Distribution of ORC and MCM Proteins in *Saccharomyces cerevisiae*: High-Resolution Mapping of Replication Origins. Science *294*, 2357-2360.

Xiao, W., Chow, B.L., Broomfield, S., and Hanna, M. (2000). The Saccharomyces cerevisiae RAD6 Group Is Composed of an Error-Prone and Two Error-Free Postreplication Repair Pathways. Genetics *155*, 1633-1641.

Xu, H., Kim, U.J., Schuster, T., and Grunstein, M. (1992). Identification of a new set of cell cycle-regulatory genes that regulate S-phase transcription of histone genes in *Saccharomyces cerevisiae*. Molecular & Cellular Biology *12*, 5249-5259.

Xu, H., Zhang, P., Liu, L., and Lee, M.Y.W.T. (2001). A Novel PCNA-Binding Motif Identified by the Panning of a Random Peptide Display Library. Biochemistry *40*, 4512-4520.

Xu, W., Aparicio, J., Aparicio, O., and Tavare, S. (2006). Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in *Saccharomyces cerevisiae*. BMC Genomics *7*, 276.

Yabuki, N., Terashima, H., and Kitada, K. (2002). Mapping of early firing origins on a replication profile of budding yeast. Genes to Cells *7*, 781-789.

Yabuuchi, H., Yamada, Y., Uchida, T., Sunathvanichkul, T., Nakagawa, T., and Masukata, H. (2006). Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. The EMBO Journal *25*, 4663-4674.

Yates, J.R. (2004). Mass spectral analysis in proteomics. Annual Review of Biophysics and Biomolecular Structure *33*, 297-316.

Yates, J.R., Ruse, C.I., Nakorchevsky, A. (2009). Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. Annual Review of Biomedical Engineering *11*, 49-79.

Ye, J., Ai, X., Eugeni, E.E., Zhang, L., Carpenter, L.R., Jelinek, M.A., Freitas, M.A., Parthun, M.R. (2005). Histone H4 lysine 91 acetylation: a core domain modification associated with chromatin assembly Molecular Cell *18*, 123-130.

Ye, X., Franco, A.A., Santos, H., Nelson, D.M., Kaufman, P.D., Adams, P.D. (2001). Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. Cell *11*, 341-351.

Yocum, A.K., Chinnaiyan, A.M. (2009). Current affairs in quantitative targeted proteomics: multiple reaction monitoring–mass spectrometry Briefings in Functional Genomics and Proteomics *8*, 145-157.

Zegerman, P., and Diffley, J.F.X. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281-285.

Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., Reinberg, D. (1997). Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. Cell *89*, 357-364.

Zhang, Y., Ng, H.-H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes & Development *13*, 1924-1935.

Zhang, Z., Shibahara, K-I., Stillman, B. (2000). PCNA connects DNA replication to epigenetic inheritance in yeast. Nature *408*, 221-225.

Zharkov, D. (2008). Base excision DNA repair. Cellular and Molecular Life Sciences *65*, 1544-1565.

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

Zhu, W., Smith, J.W., Huang, C.-M. (2010). Mass Spectrometry-Based Label-Free Quantitative Proteomics. Journal of Biomedicine & Biotechnology *2010*, 6.

Zlotkin, T., Kaufmann, G., Jiang, Y., Lee, MY., Uitto, L., Syväoja, J., Dornreiter, I., Fanning, E., Nethanel, T. (1996). DNA polymerase epsilon may be dispensable for SV40-but not cellular-DNA replication. The EMBO Journal *15*, 2298-2305.

Zolnierowicz, S., Bollen, M. (2000). Protein phosphorylation and protein phosphatases De Panne, Belgium, September 19-24, 1999. The EMBO Journal *19*, 483-488.

Zou, L., and Stillman, B. (2000). Assembly of a Complex Containing Cdc45p, Replication Protein A, and Mcm2p at Replication Origins Controlled by S-Phase Cyclin-Dependent Kinases and Cdc7p-Dbf4p Kinase. Molecular & Cellular Biology *20*, 3086-3096.

2. Artifactual sulfation of silver-stained proteins: implications for the assignment of phosphorylation and sulfation sites

Marlene Gharib, Maria Marcantonio, Sylvia G. Lehmann, Mathieu Courcelles, Sylvain Meloche, Alain Verreault, Pierre Thibault. (2009)

Molecular & Cellular Proteomics, (8)3: 506-518.

2.1. Abstract

Sulfation and phosphorylation are post-translational modifications imparting an isobaric 80 Da addition on the side chain of serine, threonine, or tyrosine residues. These two post-translational modifications are often difficult to distinguish because of their similar MS fragmentation patterns. Targeted MS identification of these modifications in specific proteins commonly relies on their prior separation using gel electrophoresis and silver staining. In the present investigation, we report a potential pitfall in the interpretation of these modifications from silver-stained gels due to the artifactual sulfation of serine, threonine, and tyrosine residues by sodium thiosulfate, a commonly used reagent that catalyzes the formation of metallic silver deposits onto proteins. Detailed MS analyses of gel-separated protein standards and *Escherichia coli* cell extracts indicated that several serine, threonine, and tyrosine residues were sulfated using silver staining protocols but not following Coomassie Blue staining. Sodium thiosulfate was identified as the reagent leading to this unexpected side reaction, and the degree of sulfation was correlated with increasing concentrations of sodium thiosulfate up to 0.02%, which is typically used for silver staining. The significance of this artifact is discussed in the broader context of sulfation and phosphorylation site identification from *in vivo* and *in vitro* experiments.

2.2. Introduction

Protein phosphorylation represents one of the most important post-translational regulatory mechanisms controlling cell function and cell signaling. Over the years, significant efforts have been devoted to the identification of phosphorylation sites in a wide range of proteins. A number of important studies have identified several thousand sites in large scale phosphoproteome experiments (Olsen, 2006; Rikova, 2007; Villen, 2007). More than 500 known eukaryotic kinases phosphorylate proteins on serine, threonine, or tyrosine residues, and it is estimated that one-third of all proteins are phosphorylated at some stage of their life cycle (Mann, 2002; Walsh, 2006; Zolnierowicz, 2000). Changes in the phosphorylation status of many receptors, transcription factors, protein kinases and several other substrates regulate a diverse array of biological processes including signal transduction, gene expression, growth, motility, proliferation, apoptosis, and differentiation (Bonenfant, 2003; Hunter, 2000; Mann, 2002; McLachlin, 2001).

Sulfation is another type of post-translational modification of serine, threonine, and tyrosine residues. In contrast to protein phosphorylation, the significance of sulfation in modulating protein function and biological processes is poorly understood (Strott, 2002). Since the first report of its occurrence in fibrinogen (Moore, 2003), only 275 tyrosine sulfated proteins have been reported in the UniProt database (Monigatti, 2006). Tyrosine sulfation is mediated by two Golgi membrane-associated protein tyrosylsulfotransferases, TPST1 and TPST2, which use 3-phosphoadenosine 5′-phosphosulfate as the sulfate group donor (Monigatti, 2006; Moore, 2003; Yu, 2007). Consequently many secretory proteins and Golgi membrane-anchored proteins are sulfated on tyrosine residues, including the chemokine receptor CCR5, the adhesion molecule P-selectin glycoprotein ligand 1 (PSGL-1), and coagulation factor VIII (Kehoe, 2000; Monigatti, 2006; Yu, 2007). Generally protein sulfation appears to play a role in mediating protein-protein interaction.

Although sulfation was primarily identified on tyrosine residues, a recent report by Medzihradszky *et al.* (Medzihradszky, 2004) indicated that this modification was also found on serine and threonine residues of several eukaryotic proteins, suggesting that its cellular function and distribution are more significant than originally suspected. However, in contrast to protein phosphorylation, there is no known consensus motif recognized by sulfotransferases and no dynamic turnover has been described for sulfation (Kehoe, 2000).

Studies of the functional consequences of protein phosphorylation and sulfation along with the identification of novel substrates are important to refine our knowledge of their potential roles in disease progression. In recent years, mass spectrometry (MS) has become a powerful tool in large scale quantitative proteomics investigations because of its high throughput and high sensitivity capabilities (Baldwin, 2004; Witze, 2007). The identification of protein post-translational modifications is typically achieved following either one- or two-dimensional gel electrophoresis prior to MS analysis (Mann, 2003). Gel electrophoresis provides unparalleled separation of intact proteins, thus enabling the resolution of complex cell extracts into bands or spots of minimal complexity. Compared with gel-free analysis of complex protein extracts, gel separation enhances the accuracy of protein assignment and maximizes the dynamic range of protein identification and their modifications. Furthermore many detergents and buffer components that are often necessary to solubilize proteins but are incompatible with MS can be conveniently removed by gel electrophoresis (Shevchenko, 2006). Protein bands or spots separated by gel electrophoresis are commonly visualized by silver staining, which provides relatively high sensitivity (~1–10 ng) compared with Coomassie Blue staining (~200–300 ng) (Graham, 2005; Shevchenko, 1996). However, undesirable side reactions of silver staining reagents that result in modifications, such as lysine formylation have recently been reported (Metz, 2006; Osés-Prieto, 2007; Richert, 2004; Shevchenko, 2006). Obviously this artifact can lead to ambiguous interpretation given that formylation can be mistakenly assigned to dimethylation because of its isobaric mass.

In this study, we report another potential pitfall in the analysis of post-translational modifications in proteins extracted from silver-stained gels. Sulfation of hydroxylated amino acid residues is caused by a side reaction of sodium thiosulfate ($Na_2S_2O_3$), a commonly used agent that catalyzes the reduction of silver nitrate ($AgNO_3$) into metallic silver as a prerequisite for protein staining. Detailed analyses of both protein standards and *E. coli* protein extracts separated by gel electrophoresis revealed that several serine, threonine, and tyrosine residues exhibited a +80 Da mass shift when analyzed after staining with silver but not with Coomassie Blue. High precision mass measurements combined with tandem MS analyses were performed using in-gel digests to confirm that the silver staining-specific modification corresponded to sulfation rather than protein phosphorylation. The extent of sulfation was correlated with increasing concentrations of the sodium thiosulfate reagent.

2.3. Experimental procedures

2.3.1. Material and reagents

Enolase from the yeast *Saccharomyces cerevisiae* was purchased from Sigma-Aldrich. Bacterial proteins were extracted from BL-21(DE3) *E. coli* cells purchased from Stratagene. ERK1 samples were prepared from human HEK293 cells transfected with pcDNA3-HAERK1-GST using the calcium phosphate method (12 × 10-cm plates for each condition). The pcDNA3-HAERK1-GST vector expresses the hamster ERK1 cDNA with a hemagglutinin tag epitope at the N-terminus and a GST sequence at the C-terminus. The day after transfection, the medium was changed to a serum-free medium, and the cells were starved for 24 h. Then the cells were stimulated for 15 min with 10% serum. After stimulation, the cells were washed twice with PBS, frozen in liquid nitrogen, and stored at −80°C until analysis. Cells were lysed for 30 min at 4°C in 0.5 ml of lysis buffer (50 mM Tris, pH 7.4, 100 mM NaCl, 5 mM EDTA, 1% Triton, 40 mM β-glycerol phosphate, 50 mM NaF, 1 mM orthovanadate, 0.1 mM PMSF, 1 mM leupeptin, 1 mM pepstatin) per dish. The cellular extracts were cleared by centrifugation for 10 min at 10,000 × *g*. The supernatants were then incubated with 75 µl of glutathione-Sepharose beads for 1.5 h. The beads were washed four times in TNET buffer (50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 5 mM EDTA, 0.5% Triton X-100) and boiled twice in Laemmli buffer for 5 min at 95°C to elute the retained ERK1 protein. The concentration of the protein extract was determined using the micro-BCA protein assay kit from Pierce. ACN and methanol (MeOH) were purchased from Fisher Scientific. Acetic acid was purchased from Fluka Biochemika (Oakville, Ontario, Canada). Formic acid (FA) was purchased from EMD Chemicals Inc. (Gibbstown, NJ). Sequencing grade modified trypsin was purchased from Promega (Madison, WI). All other chemicals were purchased from Sigma.

2.3.2. Protein separation by SDS-PAGE

A total of 2 µg of yeast enolase, 1 µg of ERK1, and 1 µg of *E. coli* protein extract were subjected to electrophoresis through a 4–12% NuPAGE bis-Tris polyacrylamide gel (Invitrogen) under denaturing conditions. After electrophoresis, proteins were fixed within the polyacrylamide gel by incubating the entire gel in 5% (v/v) acetic acid in a 1:1 (v/v) water:ethanol solution. For silver staining, the gel was first sensitized for 1 min using an aqueous solution of 0.02% sodium thiosulfate ($Na_2S_2O_3$) unless otherwise indicated. Staining was performed by incubating the gel in 0.1% (v/v) silver nitrate ($AgNO_3$) in water for 25 min at 4 °C. Finally the gel was developed in 3% (w/v) sodium carbonate (Na_2CO_3, pH 11.4) containing 0.05% formalin (v/v) in water. The staining was then stopped with a solution of 5% (v/v) acetic acid in water. For Coomassie Blue staining, proteins were fixed after gel separation and stained in a one-step procedure by incubating the entire gel in 0.1% (w/v) Coomassie Brilliant Blue R-250 in a 1:8:11 (v/v/v) acetic acid:methanol:water mixture for 1 h at room temperature. The gel was finally rinsed three times in a 1:4:5 (v/v/v) acetic acid:methanol:water solution at room temperature to visualize protein bands.

2.3.3. Destaining and in-gel digestion

Protein bands were excised from the gel and destained in 200 µl of destaining solution composed of 30 mM potassium hexacyanoferrate ($K_3Fe(CN)_6$) and 100 mM $Na_2S_2O_3$ in a 1:1 ratio for all silver-stained gel bands. For Coomassie Blue-stained gels, the destaining process was performed by incubating the gel pieces in 200 µl of a 1:1 (v/v) water:ACN solution. Proteins were then reduced with 10 mM DTT in 50 mM ammonium bicarbonate (NH_4HCO_3), pH 8.5 at 56°C for 1 h and alkylated using 55 mM iodoacetamide in 50 mM ammonium bicarbonate, pH 8.0 at room temperature for 1 h in the dark. Proteins were digested with trypsin in 50 mM ammonium bicarbonate, pH 8.0 at 37 °C for 4 h. Peptides were extracted with 5% (v/v) TFA in a 1:1 (v/v) water:ACN mixture. Following

evaporation to dryness, peptides were resuspended in 30 µl of 0.2% FA in water and analyzed.

2.3.4. β-elimination

Dried enolase peptides were dissolved in 20 µl of a freshly prepared saturated barium hydroxide ($Ba(OH)_2$) solution. The mixture was incubated at room temperature for 15 min, and the reaction was terminated by adding 1 µl of 100% FA. The reaction mixture was evaporated in a SpeedVac, and the dried peptides were redissolved in 30 µl of 0.2% FA in water prior to MS analysis.

2.3.5. Mass spectrometry analysis

All MS analyses were performed using an LTQ-Orbitrap hybrid mass spectrometer with a nanoelectrospray ion source (ThermoElectron, San Jose, CA) coupled to a nano-flow LC system (Eksigent, Dublin, Ireland) equipped with a Finnigan AS autosampler (Thermo Electron, San Jose, CA). Protein digests were separated using a 10-cm-long, 150-µm-inner diameter analytical column and a 4-mm-long, 360-µm-inner diameter trap column packed in house with 3-µm C_{18} particles (Jupiter 300 Å, Phenomenex, Torrance, CA). The mobile phase consisted of 0.2% FA in water (solvent A) and 0.2% FA in ACN (solvent B). The pump flow rate was set to 0.6 µl/min, and peptide elution was achieved using a linear gradient of 5–40% B for the first 53 min followed by a rapid increase to 60% B for the next 3 min. The conventional MS spectra (survey scan) were acquired at high resolution (M/ΔM, 60,000 full-width half-maximum) over the acquisition range of m/z 400–1600. Full scan MS^2 and MS^3 spectra were analyzed in the linear trap. Peptides were analyzed in data-dependent mode where for each 1-s survey scan the three most intense precursor ions with intensity above 10,000 counts were selected for MS^2 sequencing with a total duty cycle of 2.5 s. Product ions arising from a neutral loss of 80 Da from the precursor ion were subsequently targeted for MS^3 sequencing. To prevent the reacquisition of product ion

spectra from the same precursor ion, a dynamic exclusion window of 0.5 Da was applied for 90 s. Mass calibration used either an internal lock mass (protonated $(Si(CH_3)_2O))_6$; *m*/*z* 445.12057) or external calibration using Calmix (caffeine, MRFA, and Ultramark) and typically provided mass accuracy within 5 ppm for all nano-LC-MS experiments.

2.3.6. Database searching

MS data were analyzed using the Xcalibur software (version 2.0 SR1). Peak lists were then generated using the Mascot distiller software (version 2.1.1, Matrix Science, London, UK) where MS processing was done using the LCQ_plus_zoom script. Database searches were performed using the search engine Mascot (version 2.1, Matrix Science). For all protein analyses, searches were conducted using the NCBInr database containing 3,310,354 entries (NCBInr March 3, 2006). The error window for experimental peptide mass values and MS^2 fragment ion mass values were set to ±0.02 and ±0.5 Da, respectively. The number of allowed missed cleavage sites for trypsin was set to 1, and phosphorylation (Ser, Thr, Tyr), oxidation (Met), deamidation (Asn, Gln) and carbamidomethylation (Cys) were all selected as variable modifications. No fixed modification was included in the search. Manual inspection of all MS^2 and MS^3 spectra for modified peptides was performed to validate assignments.

2.3.7. Peptide detection and clustering

Raw data files (.raw) generated from the LTQ-Orbitrap acquisition software were converted into text files representing all ions according to their corresponding *m*/*z* values, retention time, peak width, intensity, and charge state using an in-house peptide detection software (Jaitly, 2007). Intensity values above a user-defined intensity threshold (10,000 counts) were considered for further analysis. Segmentation analyses were performed across different sample sets using hierarchical clustering with criteria based on their respective

m/*z*, charge, and time within a user-defined tolerance (±0.05 *m*/*z* and ±1 min). Normalization of retention time is then performed on the initial peptide cluster list using a dynamic and nonlinear correction. A moving average time window interpolation scheme is used to compute the time shifts for each peptide across the different data sets. For replicate LC-MS injections, this alignment confines the retention time distribution to less than ±0.1 min (<0.3% relative standard deviation) on average. The generated unique list of peptide clusters allowed the direct comparison of peptide abundance between samples in different conditions to identify those showing reproducible and statistically meaningful changes in abundance.

2.4. Results

2.4.1. Evaluation of staining protocols on the occurrence of protein sulfation artifact

In a previous study, silver staining was found to introduce chemical modification artifacts on various proteins. More specifically, staining procedures requiring the use of formaldehyde resulted in ε-formylation of lysine residues (Osés-Prieto, 2007). In the course of preliminary gel electrophoresis experiments on protein standards, we encountered another chemical modification artifact associated with the gel staining protocol. To illustrate this problem, we analyzed yeast enolase using electrophoresis through an SDS-polyacrylamide gel, which was then stained with either silver or Coomassie Blue following standard protocols (see "Experimental Procedures"). Bands were excised from the gel and digested with trypsin, and the resulting peptides were analyzed by nano-LC-MS^n on an LTQ-Orbitrap mass spectrometer. The acquired data were then searched against a yeast database using the search engine Mascot to identify modified peptides. In total, Mascot identified eight peptides showing an addition of 80 Da in the silver-stained enolase sample: ^{16}GNPTVEVELTTEK^{28}, ^{33}SIVPSGASTGVHEALEMR^{50}, ^{89}AVDDFLISLDGTANK^{103}, ^{186}IGSEVYHNLK^{195}, ^{244}IGLDC(carb)ASSEFFKDGK^{258} (C(carb) represents carbamidomethylated cysteine), ^{259}YDLDFKNPNSDK^{270}, ^{313}TAGIQIVADDLTVTNPKR^{330}, and ^{359}AAQDSFAAGWGVMVSHR^{375}. This modification potentially can be assigned to either phosphorylation or sulfation. Analyses of all CID tandem mass spectra showed the presence of an abundant fragment ion corresponding to a loss of 80 Da from the precursor ion. However, because no fragment ions generated from cleavages along the peptide backbone were observed after the neutral loss, positive identification of the peptide sequence and the modification site was not possible. This is mainly explained by the fact that phospho- and sulfoester bonds are more labile than peptide bonds upon low energy collisional activation. To circumvent this

problem, a subsequent MS^3 experiment was performed by isolating and fragmenting the precursor ion $[M + 2H - 80]^{2+}$. Consequently extensive backbone fragmentation of the modified peptides was achieved, thus enabling sequence assignment of the modified peptides. However, upon inspection the fragmentation pattern of these modified peptides appeared to be similar to that of the unmodified peptide counterpart, making it impossible to confirm the modification site. The MS^2 and MS^3 spectra of the eight enolase peptides exhibiting a mass shift of 80 Da are presented in Figure A1 (Appendix A).

A manual inspection of the LC-MS chromatogram obtained for the Coomassie Blue-stained enolase sample revealed no peptide ion exhibiting an 80 Da shift. Figure 2.1 illustrates the comparison of the abundance for all eight modified peptides obtained from the silver- and Coomassie Blue-stained enolase samples. To confirm that the complete absence of these modified peptides in the Coomassie Blue staining-derived samples was not due to variations in gel loading or mass spectrometry performance, the intensity of four of the unmodified yeast enolase peptides (^{68}NVNDVIAPAFVK^{79}, ^{331}IATAIEKK^{338}, ^{339}AADALLLK^{346}, and ^{347}VNQIGTLSESIK^{358}) was compared between the two conditions (Figure 2.1). Unmodified enolase peptides were equally detected in both samples. In contrast, peptides bearing a +80 Da mass shift in silver-stained samples were not detected in Coomassie Blue-stained samples. This comparison suggests that the presence of this artifact is specific to the silver staining procedure.

Phosphorylation and sulfation of proteins both impart a mass shift of 80 Da on serine, threonine, and tyrosine residues. However, because the mass difference between these two types of modification is only 0.0095 Da, the distinction between sulfation and phosphorylation cannot be achieved using low resolution mass spectrometers. For unambiguous assignment of these modifications, we relied on the LTQ-Orbitrap MS instrument, which provides high resolution ($M/\Delta M$, 60,000) and good mass accuracy (<5 ppm).

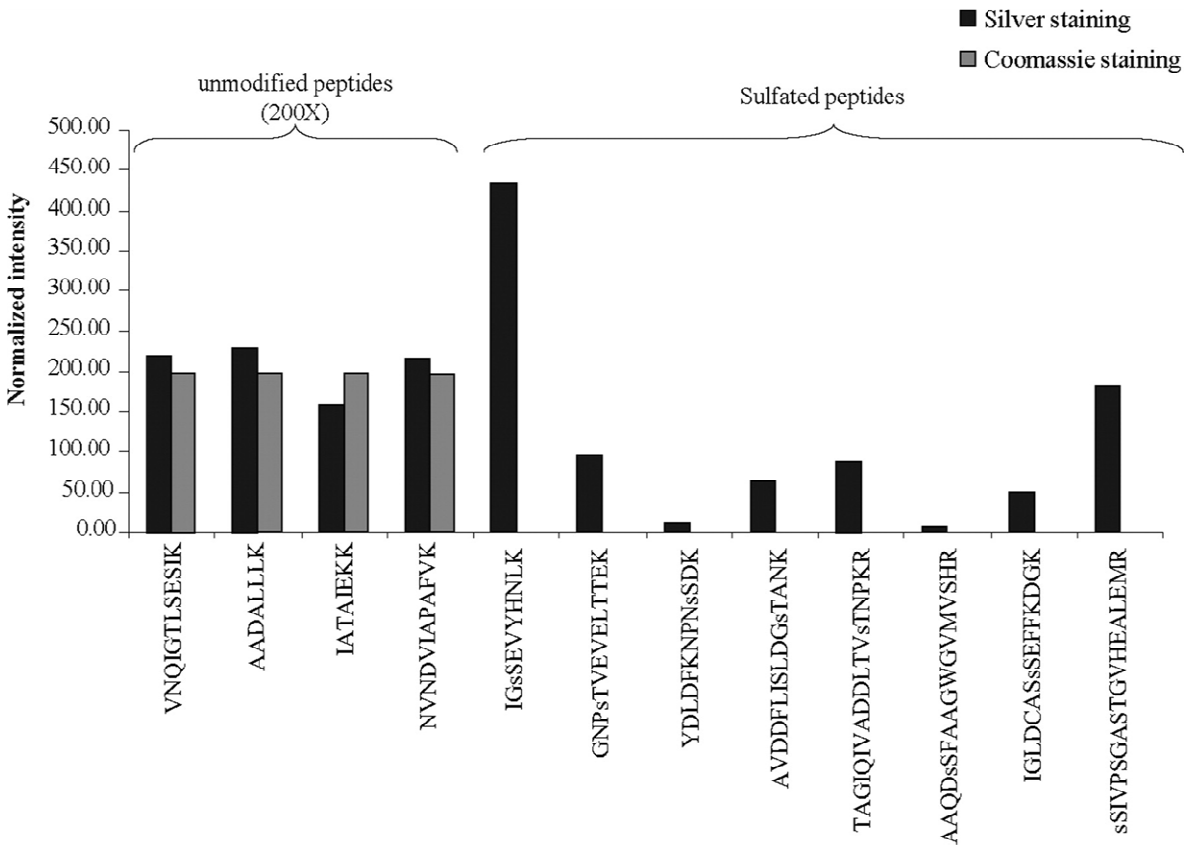


Figure 2.1 Intensity distribution of sulfopeptides and unmodified peptides from silver- and Coomassie Blue-stained yeast enolase.

Mass measurements obtained for all eight modified peptides are listed in table 2.1. The observed and calculated mass values for each peptide ion were compared for the anticipated phosphopeptide and sulfopeptide counterparts. The mass measurements obtained for the eight modified peptides (ΔM, +80 Da) were in closer agreement with those expected for sulfated peptides. For example, the exact mass value for the peptide ^{16}GNPTVEVELTTEK^{28} (+80 Da modification) eluting at 31.5 min is 1495.6718 Da. The mass difference between the observed and theoretical values for the sulfated or phosphorylated peptide is +1.1 and −5.2 ppm, respectively. For all eight modified peptides,

the average mass difference between the observed and theoretical values was 1.5 ± 0.5 ppm for sulfation compared with −4.3 ± 1.4 ppm for phosphorylation. To provide additional evidence that the +80 Da mass shift on each of these modified peptides is due to sulfation, the difference between the modified peptide and its unmodified counterpart was also calculated. In all cases, the average difference between the modified and unmodified peptides was 79.9549 ± 0.0014 Da confirming the presence of a sulfated moiety (79.9568 Da) and not its phosphorylated counterpart (79.9663 Da).

Table 2.1 Identification of sulfopeptides from yeast enolase

Peptide sequence	RT (min)	Observed mass (Da)	Calculated mass (Da)	Δmass[a] (ppm)	Modification[b] (Da)
IG**sS**EVYHNLK[c]	22.4	1238.5602	1238.5590	1.0	79.9528
IGSEV**sY**HNLK[c]	27.1	1238.5600	1238.5590	0.8	79.9526
IGSEVYHNLK	21.0	1158.6074	1158.6022	4.5	---
GNP**sT**VEVELTTEK[c]	31.5	1495.6718	1495.6702	1.1	79.9538
GNPTVEVELTTEK	28.5	1415.7180	1415.7134	3.2	---
YDLDFKNPN**sS**DK	29.1	1534.6248	1534.6236	0.8	79.9546
YDLDFKNPNSDK	27.4	1454.6702	1454.6666	2.5	---
AVDDFLISLDG**sT**ANK	43.4	1657.7514	1657.7494	1.2	79.9560
AVDDFLI**sS**LDGTANK	44.0	1657.7528	1657.7494	2.1	79.9574
AVDDFLISLDGTANK	42.7	1577.7954	1577.7926	1.8	---
TAGIQIVADDLTV**sT**NPKR	36.4	1991.0022	1990.9982	2.0	79.9542
TAGIQIVADDL**sT**VTNPKR	40.4	1991.0016	1990.9982	1.7	79.9536
TAGIQIVADDLTVTNPKR	36.7	1911.0480	1911.0414	3.5	---
AAQD**sS**FAAGWGVMVSHR[c]	40.2	1868.7964	1868.7924	2.1	79.9563
AAQDSFAAGWGVMVSHR	38.9	1788.8401	1788.8356	2.5	---
IGLDC(carb)AS**sS**EFFKDGK	36.3	1752.7354	1752.7324	1.7	79.9548
IGLDC(carb)A**sS**SEFFKDGK[c]	37.7	1752.7344	1752.7324	1.1	79.9538
IGLDC(carb)ASSEFFKDGK	35.3	1672.7806	1672.7756	3.0	---
sSIVPSGASTGVHEALEMR[c]	35.8	1919.8738	1919.8706	1.7	79.9562
SIVPSGA**sS**TGVHEALEMR	36.4	1919.8740	1919.8706	1.8	79.9564
SIVP**sS**GASTGVHEALEMR	37.0	1919.8732	1919.8706	1.4	79.9556
SIVPSGASTGVHEALEMR	32.9	1839.9176	1839.9138	2.1	---

[a]difference between observed and calculated masses,
[b] Mass difference between modified and unmodified peptide (phosphate, 79.9663; sulfate, 79.9568),
[c] Sulfate site has not been confirmed by β-eimination. Sulfated residues are in bold.
The rows with dashes correspond to the unmodified peptides. sS, sulfated serine; sT, sulfated threonine; C(carb), carbamidomethylated cysteine; RT, retention time.

2.4.2. Sulfation artifacts on serine, threonine and tyrosine residues

Manual inspection of all sulfated peptides identified in the silver-stained enolase sample (Figure 2.1) indicated the presence of closely eluting sulfopeptide isomers. For example, the reconstructed ion chromatogram for the sulfopeptide ^{33}SIVPSGASTGVHEALEMR^{50} at *m*/*z* 960.94^{+2} shows three distinct chromatographic peaks eluting within 2 min from each other (Figure 2.2a). Interestingly this peptide contains three serine residues and a threonine residue that can be potentially sulfated. This was also observed with other sulfopeptides such as ^{259}YDLDFKNPNSDK^{270} (one serine residue) and ^{244}IGLDC(carb)ASSEFFKDGK^{258} (two serine residues) where one and two chromatographic peaks appeared, respectively. It is noteworthy that peptides containing other hydroxylated residues, such as threonine and tyrosine, also showed additional peaks. For example, ^{186}IGSEVYHNLK^{195}, ^{313}TAGIQIVADDLTVTNPKR^{330}, and ^{359}AAQDSFAAGWGVMVSHR^{375} all comprised isomeric elution doublets. On the other hand, not all peptides containing more than one modified residue (serine, threonine, or tyrosine) presented multiple chromatographic peaks. This was the case for peptides ^{259}YDLDFKNPNSDK^{270} and ^{16}GNPTVEVELTTEK^{28}.

Confirmation of the identity of the sulfopeptide isomers was obtained from detailed examination of MS^2 and MS^3 spectra. For example, sulfopeptides ^{33}SIVPSGASTGVHEALEMR^{50} at *m*/*z* 960.94^{+2} eluting at 35.8, 36.4, and 37.0 min are presented in Figure 2.2a. The MS^2 spectra extracted from the three corresponding peptides all showed a prominent loss of SO_3 $[M + 2H - 80]^{2+}$ (Figure 2.2c). Similarly the MS^3

spectra of the sulfated fragment ions all showed identical fragmentation patterns confirming the presence of sulfopeptide isomers of ^{33}SIVPSGASTGVHEALEMR^{50} (Figure 2.2c).

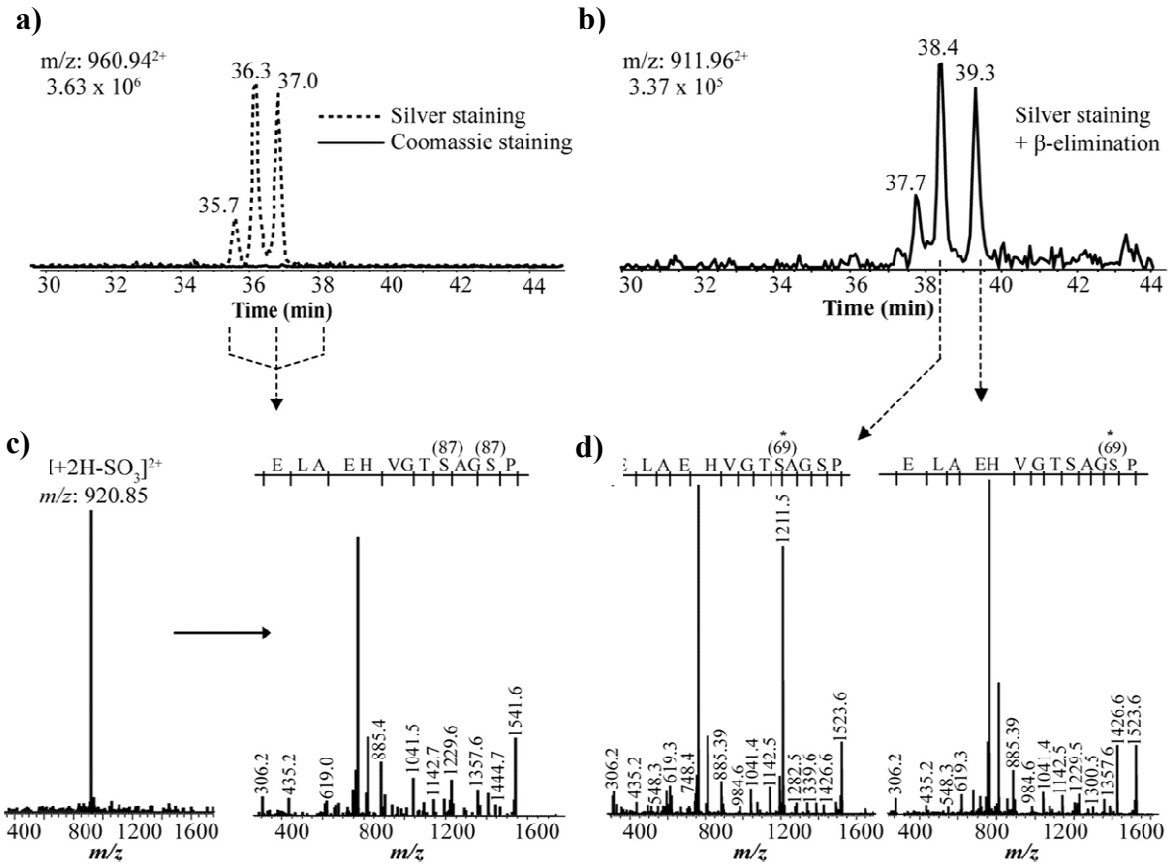


Figure 2.2 Identification of modified peptides from yeast enolase. a) Reconstructed ion chromatogram of *m/z* 960.94 from silver and Coomassie Blue stained gel bands. The ion was undetectable in the Coomassie Blue stained band. **b)** Reconstructed ion chromatogram of *m/z* 911.96 following β-elimination using $Ba(OH)_2$ of sulfopeptide shown in a. **c)** First generation (MS^2) product ion spectrum of *m/z* 960.94 (left) and second generation (MS^3) product ion spectrum of *m/z* 920.85 (right) obtained from precursor *m/z* 960.94. **d)** MS^2 product ion spectrum of two isomeric β-elimination peptides eluting at 38.4 and 39.3 min.

The prompt dissociation of the sulfate group from the modified peptide during low energy collisional activation precludes the precise localization of the modified residue(s). To identify the sites of modification, we used β-elimination and analysis of the

corresponding cleavage products using LC-MS^2 as described previously for sulfation (Medzihradszky, 2004), phosphorylation (Oda, 2001b; Thompson A.J., 2003), and *O*-glycosylation (Greis, 1996; Wells, 2002). The alkaline hydrolysis of phosphoester or sulfoester bonds leaves readily traceable β-elimination products that correspond to the dehydrated form of the modified residues (*i.e.* dehydroalanine and dehydroaminobutyrine for modified serine and threonine, respectively).

Prior to nano-LC-MS, peptides derived from a silver-stained enolase gel band were subjected to the β-elimination reaction using an aqueous solution of $Ba(OH)_2$. Data obtained for the β-elimination products of the sulfopeptide ^{33}SIVPSGASTGVHEALEMR^{50} are presented in Figure 2.2b,d. The corresponding peptide ion at *m*/*z* 911.96^{+2} eluted at three different chromatographic retention times: 37.7, 38.4, and 39.3 min (Figure 2.2b). Interestingly the elution profile of the β-elimination product peptide is similar to that of its sulfated counterpart (Figure 2.2a). However, after β-elimination, all three peptide isomers eluted 2 min later than their sulfated counterparts. This trend is consistent with the notion that loss of the sulfate group caused by β-elimination favors the formation of a more hydrophobic counterpart (Chalmers, 2004) (e.g. the conversion of a sulfated serine into dehydroalanine). The MS^2 spectra of the two main peptides eluting at 38.4 and 39.3 min yielded characteristic y-ions that enabled a straightforward identification of the modified residues (Figure 2.2d). Indeed the form of the peptide ion ^{33}SIVPSGASTGVHEALEMR^{50} eluting at 38.4 min generated a fragment ion at *m*/*z* 1211.5 (y_{11}) corresponding to the dehydroalanine conversion of a sulfated Ser-8 residue. In contrast, the peptide ion eluting at 39.3 min did not generate a fragment at *m*/*z* 1211.5. Rather a peak at *m*/*z* 1229.5, assigned to an unmodified serine at y_{11}, was detected. The exact sulfation site was assigned to Ser-5 due to the presence of a fragment ion at *m*/*z* 1426.6 corresponding to the dehydroalanine generated after β-elimination. Unfortunately the modified peptide eluting at 37.7 min did not yield an MS^2 spectrum of sufficient quality for unambiguous identification of the remaining site, thus leaving either Ser-1 or Thr-9 as possible modification sites. It is noteworthy that the β-elimination reaction did not provide results for all sulfated peptides.

This is because of the fact that β-elimination on pTyr is inefficient and does not lend itself to the identification of these modified residues (Chalmers, 2004). In addition, β-elimination of serine residues flanked by specific amino acids (such as pSer-Pro) is not favored, and pThr is 20 times less susceptible than pSer to undergo this reaction (Byford, 1991; Li, 2003). Therefore, the site of modification was inferred on other potential sulfated serine, threonine, and tyrosine residues when MS^2 spectra were not available. These results show that the silver staining-induced sulfation artifact affects serine, threonine, and tyrosine residues randomly.

2.4.3. Sodium thiosulfate gives rise to protein sulfation artifacts

$Na_2S_2O_3$ is a reducing agent used to sensitize proteins prior to metallic silver deposition on the side chains of amino acids. This pre-treatment of the gel greatly improves the silver staining method by increasing the sensitivity and the contrast of the stain (Chevallet, 2006). Because this chemical is the only potential source of sulfate used in the procedure, we investigated its effect on sulfation of serine, threonine, and tyrosine residues during silver staining. Yeast enolase was subjected to gel electrophoresis and then stained with silver using increasing concentrations of $Na_2S_2O_3$ (0–0.02%). MS analysis of the tryptic digests of yeast enolase revealed a linear relationship between the abundance of sulfopeptide artifacts and the concentration of $Na_2S_2O_3$ used during the sensitizing step (Figure 2.3).

The eight sulfopeptides described previously (Table 2.1) all showed increasing levels of sulfation that were correlated with the $Na_2S_2O_3$ concentration (Figure 2.3). On the other hand, the intensity of four unmodified yeast enolase peptides was not affected by increasing $Na_2S_2O_3$ concentration. In the complete absence of $Na_2S_2O_3$, no enolase sulfopeptide was detected. This result indicates that $Na_2S_2O_3$, which is also used later during the destaining process prior to MS analysis, does not contribute to the formation of the artifact. Taken together, these results clearly demonstrate that the sulfation artifact

introduced on serine, threonine, and tyrosine side chains during the silver staining process is caused by sodium thiosulfate. The omission of thiosulfate prevents the formation of this artifact but has a major drawback. Decreasing the amount of $Na_2S_2O_3$ markedly reduces the sensitivity of protein visualization by silver staining (not shown).

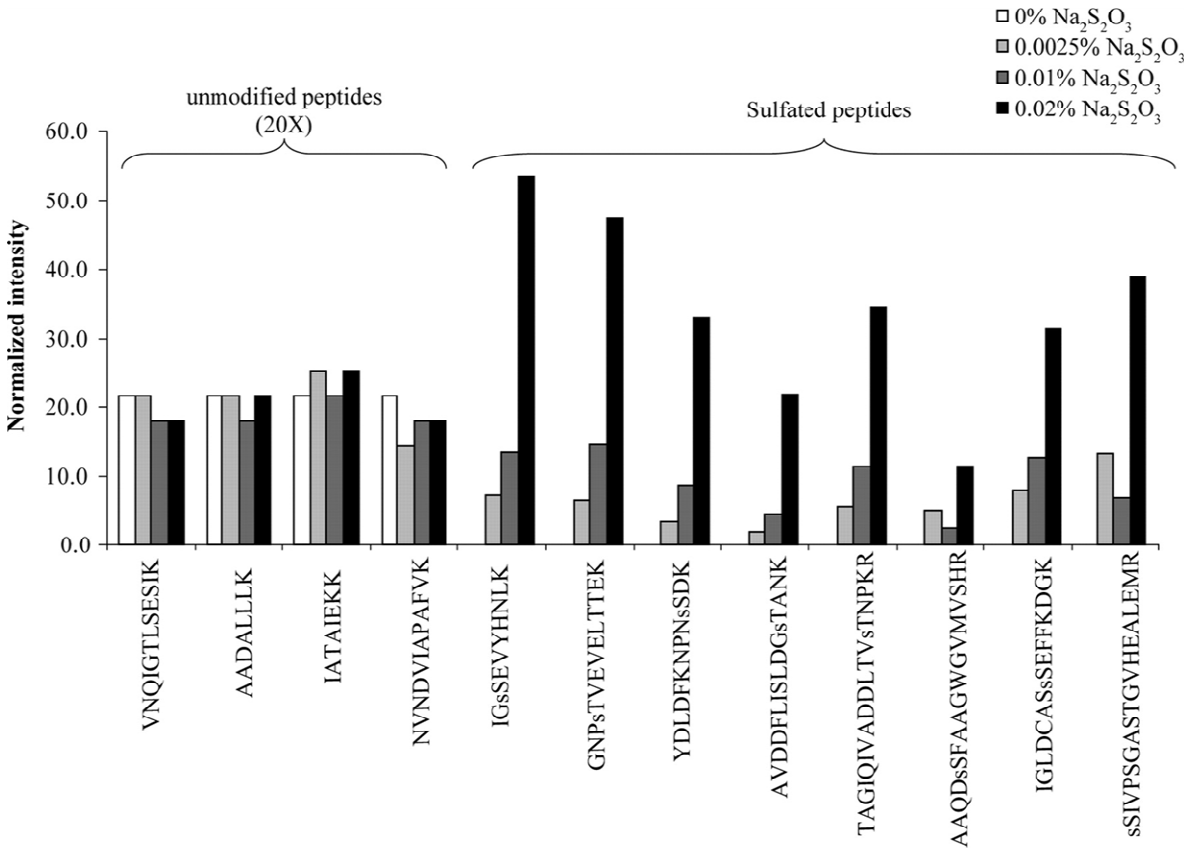


Figure 2.3 Influence of sodium thiosulfate concentration on the abundance of silver stained-induced sulfopeptides. Peptides identified from product ion spectra are shown below each bar. The letter s preceding the single letter amino acid code indicates the sulfated residue.

2.4.4. Sulfation artifact caused by silver staining of phosphorylated ERK1

We found that proteins extracted from silver-stained gels undergo artifactual sulfation that can potentially be mistaken for genuine phosphorylation. To demonstrate the impact of this artifact on the interpretation of real phosphorylation sites and to investigate further how these two types of modifications can be distinguished from each other, we analyzed ERK1 samples. ERK1 is a mitogen-activated protein kinase that is strongly activated in response to various extracellular stimuli and plays a pivotal role in mediating signal transduction (Roux, 2004). Activation of ERK1 is achieved through its phosphorylation on two conserved residues, Thr-202 and Tyr-204, that are located within its activation loop (Butch, 1996). Therefore, ERK1 provides an adequate model protein to distinguish sulfation artifacts from real phosphorylation events within a single protein. ERK1 samples isolated from HEK293 cells were separated by SDS-PAGE and then stained with either silver or Coomassie Blue. Protein bands were excised, digested with trypsin, and analyzed by MS. We obtained sequence coverage of 59% (23 peptides) and 69% (26 peptides) for silver- and Coomassie Blue-stained ERK1, respectively. Among the 23 peptides detected in the silver-stained band of ERK1, we identified seven peptides bearing a +80Da modification (Table 2.2). The precise mass of each detected peptide was measured and compared with the calculated mass values for the corresponding phospho- and sulfopeptides. The mass of five peptides (^{157}YIHSANVLHR^{166}, ^{167}DLKPSNLLINTTCDLK^{182}, ^{213}APEIMLNSK^{221}, ^{311}MLTFNPNKR^{319}, and ^{361}LKELIFQETAR^{371}) were in close agreement with the expected mass for sulfopeptides. Interestingly one additional peptide exhibiting a +80 Da mass shift, ^{192}IADPEHTGFLTEYVATR^{208}, eluted at 28.5 and 30.7 min.

These two species correspond to phosphorylated and sulfated analogs, respectively (Figure 2.4*a*). As shown in table 2.2, the peptide eluting at 28.5 min had a measured mass closer to that of the phosphorylated form (−1.7 ppm) than the sulfated form (2.5 ppm).

Conversely the peptide eluting at 30.7 min was assigned to a sulfopeptide (−1.4 ppm) and not to the corresponding phosphopeptide because of the larger mass difference (−5.7 ppm). MS/MS analyses of the peptide eluting at 28.5 min showed an abundant fragment ion at *m*/*z* 1077.76^{2+} corresponding to the loss of 98 Da ($-H_3PO_4$) from the precursor ion together with low intensity fragment ions enabling the sequence assignment and the identification of the phosphorylation site on Thr-202 (Figure 2.4b). Phosphorylation at this residue represents one of two known phosphorylation sites (Thr-202 and Tyr-204) that can be independently phosphorylated in ERK1 (Ballif, 2005). The observed fragmentation pattern, particularly the loss of 98 Da, is characteristic for phosphopeptides and is generally seen when serine and threonine residues are phosphorylated. Figure 2.4c shows the MS/MS spectrum of the sulfopeptide eluting at 30.7 min where an abundant loss of 80 Da ($-SO_3$) from the precursor ion at *m*/*z* 1086.78^{2+} is observed instead of the 98-Da loss (H_3PO_4) commonly detected for phosphopeptides. In addition, no fragment ion allowing the sequencing of the peptide was detected in striking contrast to the phosphopeptide fragmentation. Hence the sulfopeptide required a subsequent MS^3 analysis for proper sequence assignment (Figure 2.4d). The five remaining sulfopeptides also produced a neutral loss of 80 Da from the precursor ion. These are presented in Figure A2 (Appendix A). We performed β-elimination by alkaline hydrolysis using $Ba(OH)_2$ to identify the exact residues that were sulfated. The assignment of sulfation sites based on these analyses is shown in Table 2.2. Results obtained using the phosphoprotein ERK1 demonstrate that differentiation between the sulfation artifact and real phosphorylation is relatively straightforward using instruments that provide high mass accuracy. Otherwise manual validation of the MS/MS spectra along with the detection of the signature $[M + 2H - 98]^{2+}$ fragment peak is required for phosphopeptide assignment.

Table 2.2 Identification of modified peptides from ERK1

Peptide sequence	RT (min)	Observed mass (Da)	Calculated mass (Da)	Δmass[a] (ppm)	Modification[b] (Da)
APEIMLN**sS**K	28.2	1081.4799	1081.4784	1.4	79.9563
APEIMLNSK	26.5	1001.5236	1001.5215	2.1	---
ML**sT**FNPNKR	31.6	1199.5462	1199.5427	2.9	79.9587
MLTFNPNKR	27.3	1119.5875	1119.5859	1.4	---
YIH**sS**ANVLHR	23.8	1288.6003	1288.5982	1.6	79.9564
YIHSANVLHR	20.6	1208.6439	1208.6414	2.1	---
LKELIFQE**sT**AR	35.2	1426.7165	1426.7126	2.7	79.9587
LKELIFQETAR	33.2	1346.7578	1346.7558	1.5	---
DLKPSNLLINT**sT**C(carb)DLK	42.6	1923.9315	1923.9281	1.8	79.9585
DLKPSNLLINTTC(carb)DLK	35.5	1843.9730	1843.9713	0.9	---
IADPEHDHTGFL**sT**EYVATR	30.7	2250.9819	2250.9851	-1.4	79.9558
IADPEHDHTGFL**pT**EYVATR	28.5	2250.9907	2250.9946	-1.7	79.9646
IADPEHDHTGFLTEYVATR	30.9	2171.0261	2171.0283	-1.0	---

[a] Difference between experimental and calculated masses
[b] Mass difference between modified and unmodified peptide (phosphate, 79.9663; sulfate, 79.9568) Modified residues are in bold. The rows with the dashes correspond to the unmodified peptides. sS, sulfated serine; pT, phosphorylated threonine; sT, sulfated threonine; C(carb), carbamydomehtylated cysteine; RT, retention time

2.4.5. Sulfation artifacts in silver-stained bacterial proteins

To determine whether the sulfation artifact occurs in a broad range of proteins, protein extracts from *E. coli* were resolved by SDS-PAGE and stained either with silver or Coomassie Blue. The most prominent bands at 45 (band 1) and 33 kDa (band 2) were excised from silver- and Coomassie Blue-stained gels and digested with trypsin prior to LC-MS/MS analyses (Figure 2.5a). Mascot search results provided a list of all the identified proteins in each of the protein gel bands. A script was developed to detect all

MS^2 spectra containing a neutral loss of 80 Da from assigned and unassigned mass spectra. Sulfopeptides were searched for in the five most abundant proteins in the list.

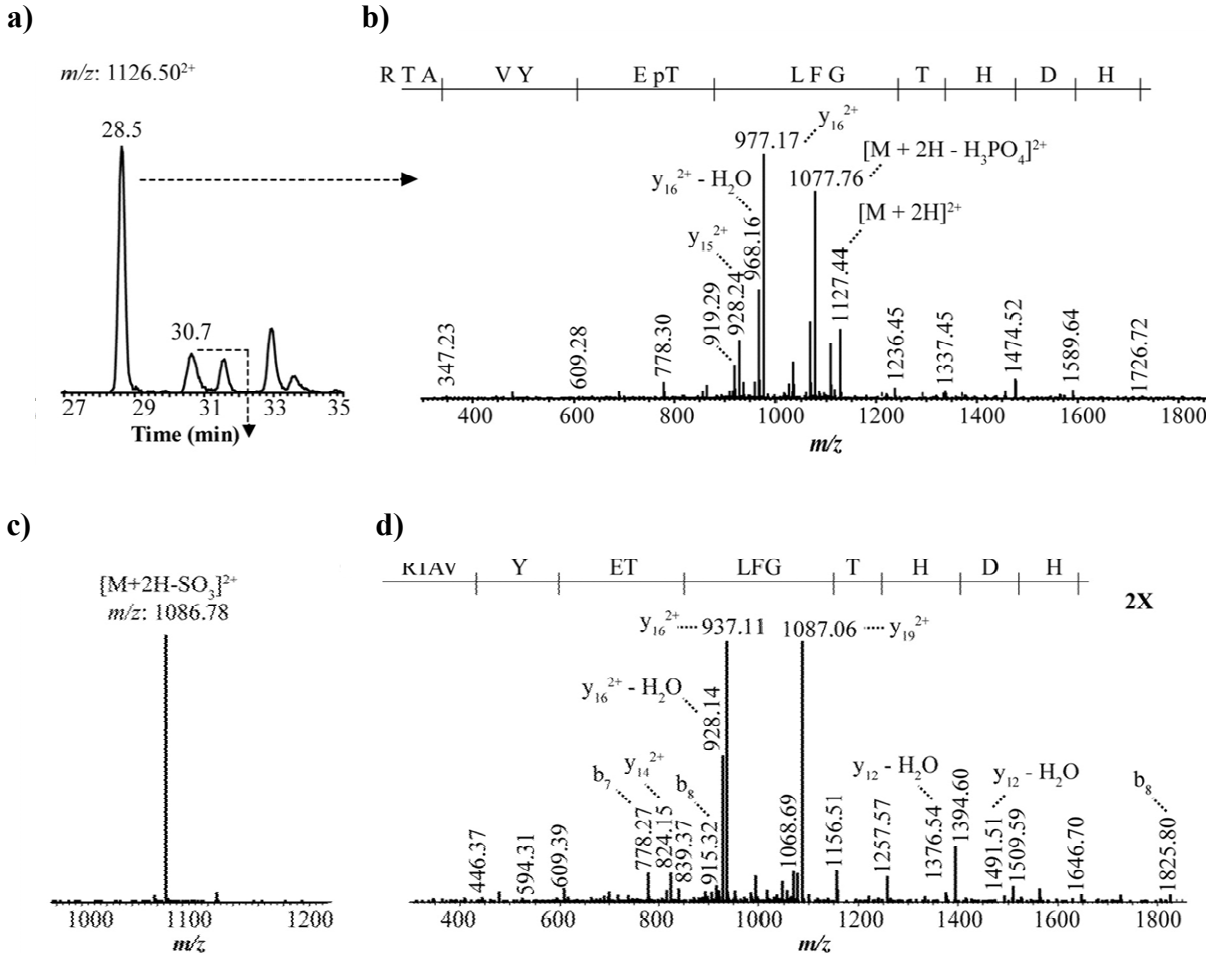


Figure 2.4 Identification of modified peptides derived from ERK1. a) Reconstructed ion chromatogram of *m/z* 1126.50 from silver-stained gel bands. **b)** First generation product ion spectrum (MS^2) of *m/z* 1126.50 eluting at 28.5 min. **c)** First generation product ion spectrum of *m/z* 1126.50 eluting at 30.7 min. **d)** Second generation product ion spectrum (MS^3) of *m/z* 1086.78 obtained from precuror 1126.50 eluting at 30.7 min.

Some of these assignments were based exclusively on peptide mass fingerprints and the presence of a neutral loss of 80 Da in the MS^2 spectra. In addition, others were obtained from their MS^3 spectra, thus enabling sequencing of the peptide and a more direct confirmation of the identity of the protein of origin (Figure A3, Appendix A). Analysis of

the silver-stained band at ~45-kDa molecular mass revealed the presence of nine sulfopeptides (Figure 2.5b-e). It is noteworthy that no sulfopeptide was detected from the Coomassie Blue-stained gel of the corresponding band. All detected sulfopeptides corresponded to elongation factor Tu (EF-Tu), an abundant 43-kDa protein that plays an essential role in protein biosynthesis (Lippmann, 1993). However, to our knowledge, no EF-Tu sulfation in bacteria has been reported thus far. To compare EF-Tu sulfation in silver- and Coomassie Blue-stained samples, all unique peptide identifications with Mascot scores above 25 were clustered together. All the observed changes are represented in a scatter plot (Figure 2.5b). The comparison of peptides from silver- and Coomassie Blue-stained bands indicated that 84% (1115) of all 1329 ion clusters (including unmodified EF-Tu ions) showed less than a 2-fold change in abundance between the two conditions, whereas 15% (205) of these same ion clusters showed, on average, a change in abundance of 56-fold. This inconsistency in peptide detection for unaffected peptides across the two conditions can be explained by the fact that reproducible cutting of gel bands at equal molecular weights cannot be fully achieved without variations because of manual handling of the gels. Nevertheless because the vast majority of peptides are reproducibly detected in both samples and unmodified EF-Tu peptides are equally present in both conditions, the abundance of EF-Tu sulfopeptides in silver- and Coomassie Blue-stained samples can be efficiently compared. A comparison of the proportion of sulfated and unmodified peptides from EF-Tu indicated that less than 1% (nine) of total peptides was identified as EF-Tu sulfopeptides in the silver-stained sample. For instance, the reconstructed ion chromatogram for the sulfopeptide VGEEVEIVGIKETQK at m/z 869.36^{+2} that eluted at 32.9 min and displayed a prominent loss of 80 Da in the silver-stained sample was completely absent in the Coomassie Blue-stained condition (Figure 2.5c-e). Accurate mass measurements further confirmed that all modified residues bearing a +80-Da shift corresponded to sulfation artifacts. Finally the analysis of silver and Coomassie Blue staining-derived samples for the gel band at 33 kDa also identified several proteins with sulfopeptides that were specifically present in the silver-stained samples (Figure 2.6).

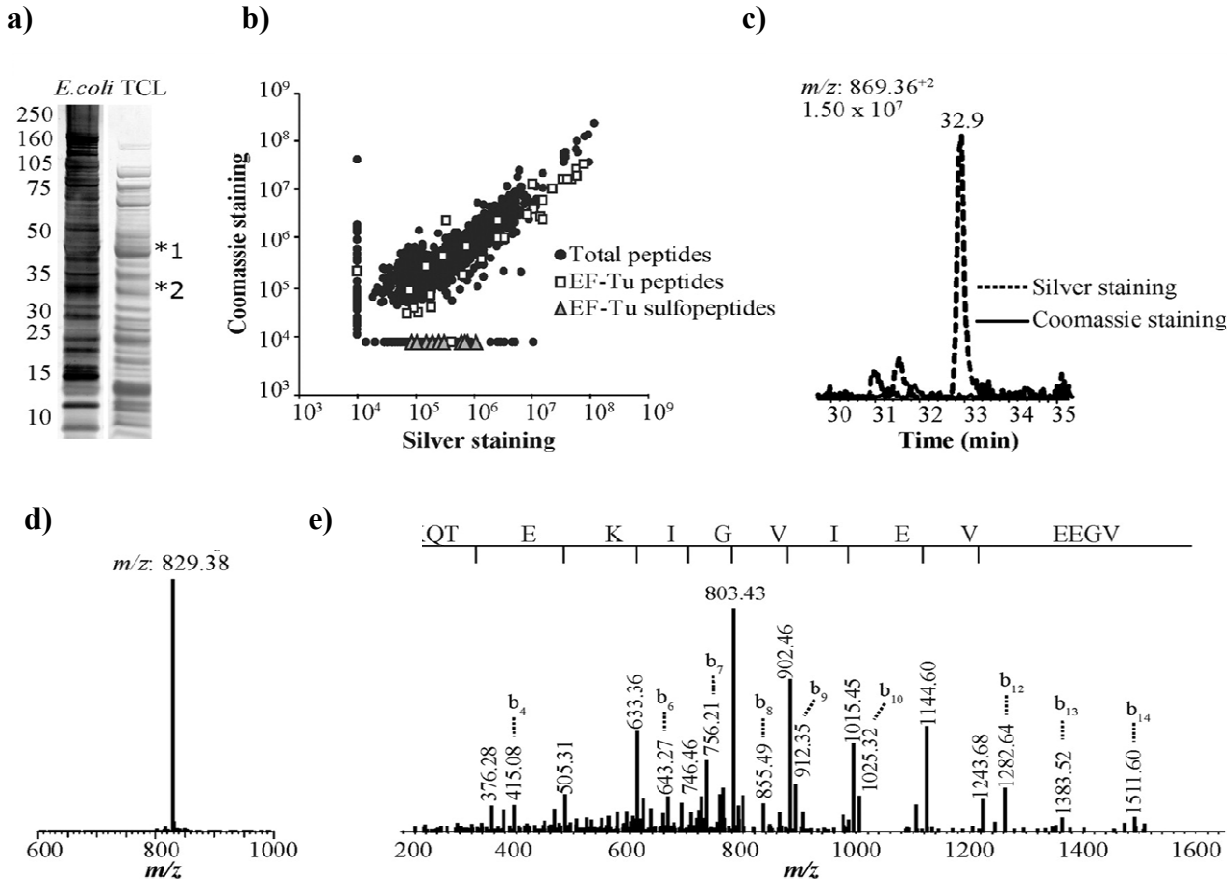


Figure 2.5 Identification of sulfopeptide artifacts from silver and Coomassie Blue stained gels of E. coli cell lysate (TCL). a) Silver (left lane) and Coomassie Blue (right lane) stained gels. **b)** Scatter plot distribution of ion intensities for the peptides identified from silver and Coomassie band 1 (shown in a). **c)** Reconstructed ion chromatogram of *m/z* 869.36 from silver and Coomassie Blue stained sample. **d)** First generation (MS^2) product ion spectrum of *m/z* 869.36. **e)** Second generation (MS^3) product ion spectrum of *m/z* 829.38 obtained from precursor *m/z* 869.36.

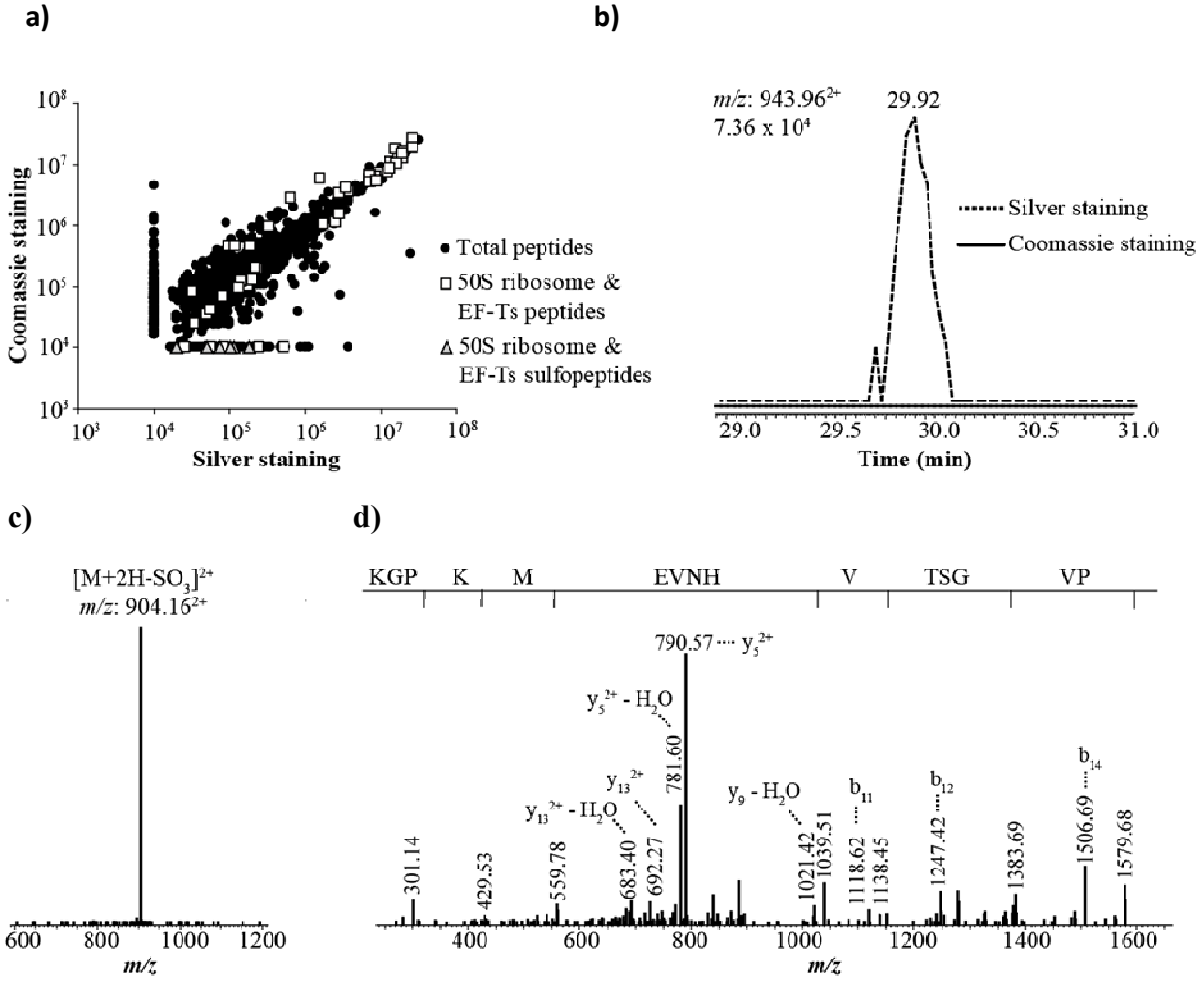


Figure 2.6 **Identification of sulfopeptide artifacts from silver and Coomassie Blue-stained gel band 2 (33 KDa identified in figure 2.5a) from *E. coli* cell lysates. b)** Scatter plot distribution of ion intensities for the peptides identified from silver and Coomassie blue stained band. **c)** Reconstructed ion chromatogram of *m/z* 943.96 from silver and Coomassie Blue stained sample. The *m/z* 943.96 was undetectable in the Coomassie-blue stained band. **d)** First generation (MS^2) product ion spectrum of *m/z* 943.96. **e)** Second generation (MS^3) product ion spectrum of *m/z* 904.16 obtained from precursor *m/z* 943.96.

2.5. Discussion

In this study, we report the identification of sulfation as a chemical modification artifact introduced during silver staining of proteins separated by polyacrylamide gel electrophoresis. The analysis of enolase, ERK1 and proteins from bacterial cell lysates derived from silver-stained gels all showed several peptides presenting a +80 Da modification on serine, threonine, and tyrosine residues. Sulfated peptides were not found in samples derived from Coomassie Blue-stained gels. The occurrence of this artifact is of particular concern for studies requiring the identification of protein phosphorylation sites using silver-stained gels and appropriate staining procedures must be used to avoid ambiguous assignment.

To identify and characterize post-translational modifications of specific proteins, most proteomics laboratories require high resolution separation methods, such as one- or two-dimensional polyacrylamide gel electrophoresis in combination with mass spectrometry. Sensitive gel staining methods enabling the detection of proteins present at low nanogram levels are also necessary for most proteomics applications (Gharahdaghi, 1999). For this purpose, silver staining is a versatile and MS-compatible method providing low nanogram detection sensitivity (Shevchenko, 1996). We have shown that sulfation of specific residues does occur during the silver staining procedure due to sodium thiosulfate, a reducing agent used to sensitize the gel prior to the addition of silver nitrate, thus enhancing the sensitivity and contrast of the staining (Rabilloud, 1990). Formaldehyde, an alternative reagent commonly used in silver staining, generates intramolecular cross-links, leading to low peptide recovery for MS analysis. Formaldehyde also gives rise to formylation artifacts on lysine residues (Osés-Prieto, 2007).

The chemistry of silver staining is well understood and is mediated through the binding of silver ions to amino acid side chains containing sulfhydryl and carboxyl groups with

subsequent reduction to free metallic silver under alkaline conditions (Granvogl, 2007; Patras, 1999; Rabilloud, 1990). In addition, the terminal sulfur atom of sodium thiosulfate has a high affinity for soft metals such as silver and forms a reactive metallic intermediate (Herrmann, 1986). Accordingly we propose that, during the staining procedure, the interaction of silver ions with sodium thiosulfate might activate this reducing reagent prior to subsequent nucleophilic reactions with free hydroxyl groups of serine, threonine, and tyrosine residues. A proposed mechanism is presented in Figure 2.7. As a result, this reaction would promote the chemical modification of residues through sulfation while forming a stable Ag_2S complex necessary for visualizing gel-separated protein bands.

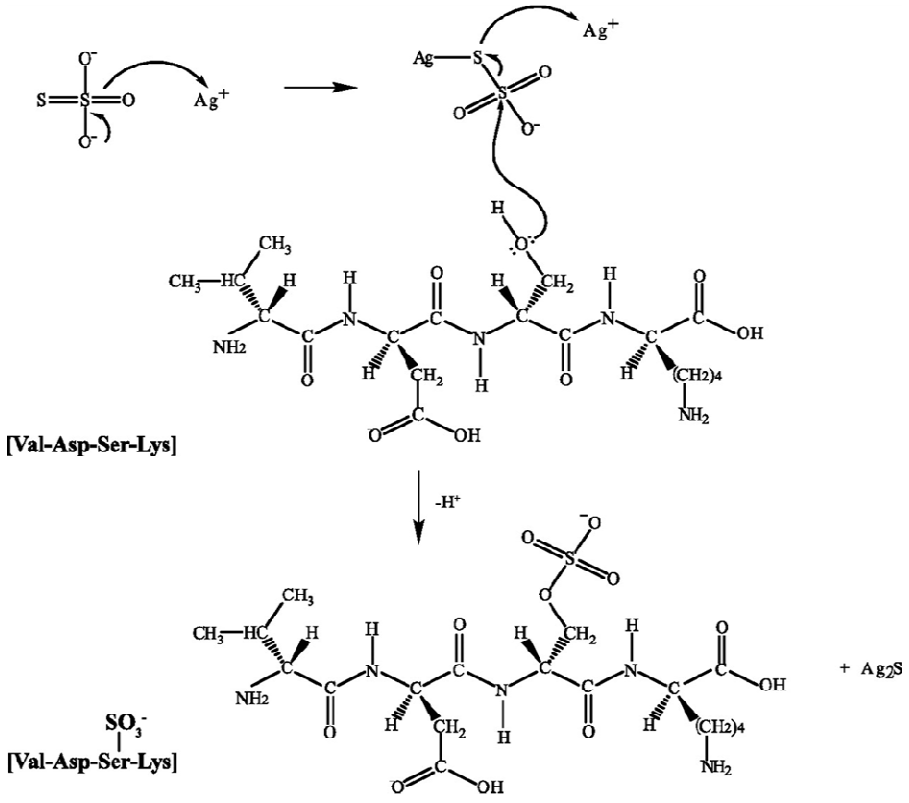


Figure 2.7 Proposed mechanism for sulfation of hydroxylated amino acids in silver-stained gels. Sodium thiosulfate reduces silver ions to form a reactive metallic intermediate that is then subject to nucleophilic attack by the lone electron pairs of hydroxylated residues (shown here for Ser). The thiosulfate intermediate can reduce a second silver ion proximal to acidic residues to produce Ag_2S necessary for visualizing the gel-separated protein and liberate a sulfated peptide.

It is noteworthy that sulfation appears to preferentially modify specific hydroxylated amino acids among all available Ser, Thr, and Tyr residues. For example, yeast enolase contains a total of 64 hydroxyamino acids (32 serine, 19 threonine, and 13 tyrosine residues), but only 14 (eight serine, five threonine, and one tyrosine residues) were detectably sulfated following silver staining (Table 2.1). This could be explained in part by the local environment influencing the nucleophilic substituents and hence the reactivity of the hydroxyl group of serine, threonine, and tyrosine side chains such as observed previously for ε-methyl-Lys residues of calmodulin (Zhang, 1993). Moreover the occurrence of protein sulfation from silver staining is likely influenced by the binding of silver ions to individual amino acids. Indeed previous reports indicated that silver ions bind predominantly to sulfhydryl and carboxylate moieties (Granvogl, 2007; Rabilloud, 1990), suggesting that Ser, Thr, and Tyr proximal to Asp, Glu, Met, and Cys residues may have a higher propensity to be sulfated. We examined the distribution of six residues on either side of the sulfated amino acid identified in enolase and ERK1 using WebLogo (Crooks, 2004). These analyses revealed that a higher frequency of acidic residues was found proximal to the modified site (Figure 2.8). Obviously these consensus motif analyses are limited to the linear sequence immediately next to the modified residues and additional neighboring amino acids might contribute to sulfation depending on the local protein conformation, notwithstanding the fact that silver staining is usually performed after electrophoresis in the presence of SDS. The requirement of proximal sulfhydryl or carboxylate moieties for the sulfation of neighboring hydroxylated residues also suggests that proteins rich in basic amino acids would be less susceptible to undergo protein sulfation.

Sulfation also changes the polarity of hydroxylated residues, rendering them more susceptible to form ionic interactions with nearby ionizable moieties. In most cases that we examined, we noticed that sulfopeptides generally eluted 1–5 min later than their unmodified counterparts (Table 2.1 and Table 2.2). This observation suggests that sulfopeptides are generally more hydrophobic than their unmodified peptides under the present elution conditions (aqueous ACN, 0.2% FA, pH 2.5). This could possibly be

explained by the formation of salt bridges between the sulfonyl group and adjacent basic moieties (N-terminal NH_2; internal His, Lys, or Arg; and/or C-terminal Lys or Arg residues) similar to those observed previously for phosphopeptides (Marcantonio, 2008). The extent of the retention time shift could also be influenced by secondary structural changes on the conformation of the modified peptide.

Figure 2.8 Distribution of 6 residues on either side of the sulfated amino acids identified in enolase and ERK1 using weblogo.

We also evaluated the degree of modification on yeast enolase by comparing the relative abundance of the sulfopeptides with their corresponding unmodified counterparts. We determined that the population of modified peptides represents ~2% of the unmodified peptides based on average intensity ratios, suggesting that protein sulfation is a relatively low yield side reaction under the present silver staining conditions (0.02% sodium thiosulfate). Analyses of *E. coli* extracts also revealed that only the most abundant proteins identified by Mascot showed peptides bearing this modification. Taken together, these results indicate that the extent of sulfation artifacts detectable from silver-stained gels depends not only on the amino acid distribution but also on the relative proportion and abundance of a given protein within a band or spot. Consequently protein sulfation should be taken into consideration when analyzing silver-stained bands or spots derived from partially purified samples obtained by immunoaffinity or conventional protein purification.

Protein sulfation caused by silver staining can also be problematic when trying to identify low abundance phosphopeptides from gel-isolated bands. Indeed the transient

nature of protein phosphorylation *in vivo* can result in purified samples where less than 1–2% of the protein is phosphorylated (Reinders, 2005). Consequently protein sulfation caused by silver staining can possibly interfere with the identification of genuine phosphorylation sites. Phosphopeptide isolation methods using IMAC or TiO_2 media are also likely to enrich indiscriminately isobaric sulfopeptides and phosphopeptides from in-gel tryptic digests, thus rendering their distinction even more challenging. A case in point is the identification of modified residues from ERK1 where we identified seven peptides bearing the +80 Da modification, but only one (Thr-202) was a genuine phosphorylation site located within the activation loop of the kinase. The observation of these low abundance modified peptides could easily have been erroneously assigned to phosphorylation in the absence of further structural validation. The possibility of obtaining accurate mass measurements together with first and second generation fragment ion spectra within the same LC-MS analysis was essential for the correct assignment of modified peptides. The use of a lock mass on the Orbitrap instrument enabled mass measurements within 5 ppm of the calculated values thus making it possible to distinguish between sulfation and phosphorylation as shown here for two isobaric ERK1 modified peptides. Obviously identification based on mass measurements would not be possible with lower resolution and mass accuracy instruments, and proper differentiation between these two isobaric modifications would rely only on distinct fragmentation patterns.

The fragmentation behavior of phosphopeptides and sulfopeptides has been studied extensively over the past years (DeGnore, 1998; Medzihradszky, 2004; Medzihradszky, 2007; Nemeth-Cawley, 2001; Syka, 2004). Cleavage of the phosphoester bond leading to a loss of H_3PO_4 through β-elimination is observed under collisional activation except for phosphorylated Tyr residues that typically retain this modification. Characteristic fragment ions displaying a +80 or a −18 Da shift compared with the unmodified peptide serve as hallmark ions to identify the modification sites of phosphopeptides containing serine and threonine residues. In contrast, sulfopeptides examined as part of this study all exhibited an abundant loss of SO_3 (−80 Da) from the precursor ion with very few sequence-specific

fragment ions. Although this fragmentation feature can be used to identify potential sulfopeptides, the sequence of the corresponding peptide could only be revealed following second generation product ion spectra (MS^3). However, the prompt loss of SO_3 from the precursor ion leaves an intact serine or threonine residue with fragmentation features identical to that of the unmodified peptide (data not shown). Although collisional activation dissipates sufficient vibrational energy in the product ions to favor dissociation of the labile SO_3 group, a recent study indicated that sulfopeptides are also more susceptible to undergo this neutral loss during electron capture dissociation (ECD) compared with their corresponding phosphopeptide analogs (Medzihradszky, 2007). This loss was less prominent using electron transfer dissociation (ETD), although several fragments ions were shown to eliminate SO_3 (Medzihradszky, 2007; Syka, 2004). Interestingly the replacement of the sulfate proton with an alkali ion such as Na^+ appeared to stabilize the prompt loss of SO_3, thereby favoring the identification of the modification site in ECD or ETD experiments (Medzihradszky, 2007). However, ECD or ETD product ion spectra of modified peptides display c- and z-fragment ions bearing the intact sulfated or phosphorylated moieties and could lead to misassignment of phosphopeptides during the analysis of in-gel digests of silver-stained bands.

Finally the mistaken assignment of artifact sulfopeptides as real phosphorylated peptides is greatly dependent on the type of MS system used to perform the analysis. It is noteworthy that collisional activation of sulfopeptides performed using the LTQ-Orbitrap and Q-TOF instruments generated product ion spectra that greatly differ in their peptide fragmentation patterns. For example, MS^2 spectra acquired on the LTQ-Orbitrap are dominated by an abundant neutral loss of SO_3 (−80 Da). Sequence assignment of the modified peptide therefore requires an additional step of isolation and fragmentation (MS^3). On the other hand, MS^2 spectra of sulfopeptides obtained from a Q-TOF instrument display extensive fragmentation along the peptide backbone in addition to the neutral loss of SO_3 Figure A4 (Appendix A). Search engines such as Mascot use probability scores to match peptide masses as well as MS^2 fragment ion masses to theoretical protein sequences from

databases. Consequently sulfopeptides analyzed on the ion trap will have very low scores and will remain unassigned after database searching. The possibility of identifying silver stain-induced sulfopeptide artifacts and misassigning them as phosphopeptides is greatly reduced when analyses are performed on the LTQ-Orbitrap.

In view of the frequent occurrence of sulfation artifacts and the wide use of colloidal silver as a stain compatible with MS identification of gel-separated proteins, users must be aware of this specific problem to avoid misinterpretation of results. This is particularly true for the identification of low abundance phosphopeptides or genuine sulfopeptides that can be isolated from in-gel digests of silver-stained bands where the sulfation artifacts can be easily misidentified as the more common isobaric modification. Alternate stains such as Coomassie Blue or SYPRO dyes that do not require thiosulfate reagents could be used to minimize this artifact when characterizing phosphopeptides or sulfopeptides from in-gel digests. The accessibility of MS equipment providing high accuracy mass measurements (<5 ppm) combined with sensitive multistage fragment ion spectra is critical to obtain unambiguous assignment.

2.6. References

Baldwin, M.A. (2004). Protein Identification by mass spectrometry. Molecular & Cellular Proteomics *3*, 1-9.

Ballif, B.A., Roux, P.P., Gerber, S.A., MacKeigan, J.P., Blenis, J., Gygi, S.P. (2005). Quantitative phosphorylation profiling of the ERK/p90 ribosomal S6 kinase-signaling cassette and its targets, the tuberous sclerosis tumor suppressors. Proceedings of the National Academy of Sciences of the United States of America *102*, 667-672.

Bonenfant, D., Schmelzle, T., Jacinto, E., Crespo, J.L., Mini, T., Hall, M.N., Jenoe, P. (2003). Quantitation of changes in protein phosphorylation: A simple method based on stable isotope labeling and mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America *100*, 880-885.

Butch, E.R., Guan, K-L. (1996). Characterization of ERK1 activation site mutants and the effect on recognition by MEK1 and MEK2. Journal of Biological Chemistry *271*, 4230-4235.

Byford, M.F. (1991). Rapid and selective modification of phosphoserine residues catalysed by Ba2+ ions for their detection during peptide microsequencing. Biochemical Journal *280*, 261-265.

Chalmers, M.J., Kolch, W., Emmett, M.R., Marshall, A.G., Mischak, H. (2004). Identification and analysis of phosphopeptides. Journal of Chromatography B *803*, 111-120.

Chevallet, M., Luche, S., Rabilloud, T. (2006). Silver staining of proteins in polyacrylamide gels. Nature Protocols *1*, 1852-1858.

Crooks, G.E., Hon, G., Chandonia, J-M., Brenner, S.E. (2004). WebLogo: A Sequence Logo Generator. Genome Research *14*, 1188-1190.

DeGnore, J.P., Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. Journal of the American Society for Mass Spectrometry *9*, 1175-1188.

Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S., Mische, S.M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: A method for the removal of silver ions to enhance sensitivity. Electrophoresis *20*, 601-605.

Graham, D.R.M., Elliott, S.T., van Eyk, J.E. (2005). Broad-based proteomic strategies: a practical guide to proteomics and functional screening. The Journal of Physiology *563*, 1-9.

Granvogl, B., Plöscher, M., Eichacker, L. (2007). Sample preparation by in-gel digestion for mass spectrometry-based proteomics. Analytical and Bioanalytical Chemistry *389*, 991-1002.

Greis, K.D., Hayes, B.K., Comer, F.I., Kirk, M., Barnes, S., Lowary, T.L., Hart, G.W. (1996). Selective Detection and Site-Analysis of O-GlcNAc-Modified Glycopeptides by [beta]-Elimination and Tandem Electrospray Mass Spectrometry. Analytical Biochemistry *234*, 38-49.

Herrmann, W.A. (1986). Multiple Bonds between Transition Metals and ldquoBarerdquo Main Group Elements: Links between Inorganic Solid State Chemistry and Organometallic Chemistry. Angewandte Chemie International Edition in English *25*, 56-76.

Hunter, T. (2000). Signaling--2000 and Beyond. Cell *100*, 113-127.

Jaitly, G., Bonneil, E., Jaitly, N., Eng, K., Pomiès, C., Thibault, P. (2007). Comprehensive profiling of unlabeled peptide ions from large-scale proteomics experiments using one and two dimensional nanoLC-MS/MS. Paper presented at: Proceedings of the 55th ASMS Conference on Mass Spectrometry (Indianapolis, IN).

Kehoe, J.W., Bertozzi, C.R. (2000). Tyrosine sulfation: a modulator of extracellular protein-protein interactions. Chemistry & Biology *7*, R57-R61.

Li, W., Backlund, P.S., Boykins, R.A., Wang, G., Chen, H-C. (2003). Susceptibility of the hydroxyl groups in serine and threonine to [beta]-elimination/Michael addition under commonly used moderately high-temperature conditions. Analytical Biochemistry *323*, 94-102.

Lippmann, C., Lindschau, C., Vijgenboom, E., Schröder, W., Bosch, L., Erdmann, V.A. (1993). Prokaryotic elongation factor Tu is phosphorylated in vivo. Journal of Biological Chemistry *268*, 601-607.

Mann, M., Jensen, O.N. (2003). Proteomic analysis of post-translational modifications. Nature Biotechnology *21*, 255-261.

Mann, M., Ong, S-E., Grønborg, M., Steen, H., Jensen, O. N., Pandey, A. (2002). Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends in Biotechnology *20*, 261-268.

Marcantonio, M., Trost, M., Courcelles, M., Desjardins, M., Thibault, P. (2008). Combined Enzymatic and Data Mining Approaches for Comprehensive Phosphoproteome Analyses. Molecular & Cellular Proteomics *7*, 645-660.

McLachlin, D.T., Chait, B.T. (2001). Analysis of phosphorylated proteins and peptides by mass spectrometry. Current Opinion in Chemical Biology *5*, 591-602.

Medzihradszky, K.F., Darula, Z., Perlson, E., Fainzilber, M., Chalkley, R.J., Ball, H., Greenbaum, D., Bogyo, M., Tyson, D.R., Bradshaw, R.A., Burlingame, A.L. (2004). O-Sulfonation of Serine and Threonine. Molecular & Cellular Proteomics *3*, 429-440.

Medzihradszky, K.F., Guan, S., Maltby, D.A., Burlingame, A.L. (2007). Sulfopeptide Fragmentation in Electron-Capture and Electron-Transfer Dissociation. Journal of the American Society for Mass Spectrometry *18*, 1617-1624.

Metz, B., Kersten, G.F.A., Baart, G.J.E., de Jong, A., Meiring, H., ten Hove, J., van Steenbergen, M.J., Hennink, W.E., Crommelin, D.J.A., Jiskoot, W. (2006). Identification of Formaldehyde-Induced Modifications in Proteins: Reactions with Insulin. Bioconjugate Chemistry *17*, 815-822.

Monigatti, F., Hekking, B., Steen, H. (2006). Protein sulfation analysis--A primer. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics *1764*, 1904-1913.

Moore, K.L. (2003). The Biology and Enzymology of Protein Tyrosine O-Sulfation. Journal of Biological Chemistry *278*, 24243-24246.

Nemeth-Cawley, J.F., Karnik, S., Rousse, J.C. (2001). Analysis of sulfated peptides using positive electrospray ionization tandem mass spectrometry. Journal of Mass Spectrometry *36*, 1301-1311.

Oda, Y., Nagasu, T., Chait, B.T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nature Biotechnology *19*, 379-382.

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

Osés-Prieto, J.A., Zhang, X., Burlingame, A.L. (2007). Formation of ε-Formyllysine on Silver-stained Proteins. Molecular & Cellular Proteomics *6*, 181-192.

Patras, G., Qiao, G.G., Solomon, D.H. (1999). On the mechanism of background silver staining during sodium dodecyl sulphate-polyacrylamide gel electrophoresis. Electrophoresis *20*, 2039-2045.

Rabilloud, T. (1990). Mechanisms of protein silver staining in polyacrylamide gels: A 10-year synthesis. Electrophoresis *11*, 785-794.

Reinders, J., Sickmann, A. (2005). State-of-the-art in phosphoproteomics. Proteomics *5*, 4052-4061.

Richert, S., Luche, S., Chevallet, M., van Dorsselaer, A., Leize-Wagner, E., Rabilloud, T. (2004). About the mechanism of interference of silver staining with peptide mass spectrometry. Proteomics *4*, 909-916.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., Hu, Y., Tan, Z., Stokes, M., Sullivan, L., Mitchell, J., Wetzel, R., Macneill, J., Ren, J.M., Yuan, J., Bakalarski, C.E., Villen, J., Kornhauser, J.M., Smith, B., Li, D., Zhou, X., Gygi, S.P., Gu, T.L., Polakiewicz, R.D., Rush, J., Comb, M.J. (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell *131*, 1190-1203.

Roux, P.P., Blenis, J. (2004). ERK and p38 MAPK-Activated Protein Kinases: a Family of Protein Kinases with Diverse Biological Functions. Microbiol & Molecular Biology Reviews *68*, 320-344.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols *1*, 2856-2860.

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

Strott, C.A. (2002). Sulfonation and Molecular Action. Endocrine Reviews *23*, 703-732.

Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America *101*, 9528-9533.

Thompson A.J., H.S.R., Franz C., Barnouin K., Ridley A., Cramer R. (2003). Characterization of protein phosphorylation by mass spectrometry using immobilized metal ion affinity chromatography with on-resin beta-elimination and Michael addition. Analytical Chemistry *75*, 3232-3243.

Villen, J., Beausoleil, S.A., Gerber, S.A., Gygi, S.P. (2007). Large-scale phosphorylation analysis of mouse liver. Proceedings of the National Academy of Sciences of the United States of America *104*, 1488-1493.

Walsh, C.T. (2006). Post-translational modification of proteins: expanding nature`s inventory, 1st edn (Greenwood Village, CO, Roberts and Co. publishers).

Wells, L., Vosseller, K., Cole, R.N., Cronshaw, J.M., Matunis, M.J., Hart, G.W. (2002). Mapping Sites of O-GlcNAc Modification Using Affinity Tags for Serine and Threonine Post-translational Modifications. Molecular & Cellular Proteomics *1*, 791-804.

Witze, E.S., Old, W.M., Resing, K.A., Ahn, N.G. (2007). Mapping protein post-translational modifications with mass spectrometry. Nature Methods *4*, 798-806.

Yu, Y., Hoffhines, A.J., Moore, K.L., Leary, J.A. (2007). Determination of the sites of tyrosine O-sulfation in peptides and proteins. Nature Methods *4*, 583-588.

Zhang, M., Vogel, H.J. (1993). Determination of the side chain pKa values of the lysine residues in calmodulin. Journal of Biological Chemistry *268*, 22420-22428.

Zolnierowicz, S., Bollen, M. (2000). Protein phosphorylation and protein phosphatases De Panne, Belgium, September 19-24, 1999. The EMBO Journal *19*, 483-488.

3. Phosphorylation of CAF-1 by Cdc7/Dbf4 promotes telomeric silencing in *Saccharomyces cerevisiae*

Marlene Gharib, Michael Weinreich, Pierre Thibault, Alain Verreault,

PNAS (To be submitted)

3.1. Abstract

Chromatin Assembly Factor 1 (CAF-1), is an evolutionarily conserved protein that plays an important role in *de novo* nucleosome assembly by directly depositing newly synthesized histones H3 and H4 behind DNA replication forks. CAF-1 is composed of three polypeptide subunits termed Cac1, Cac2 and Cac3 in yeast. Although phosphorylation of some CAF-1 subunits has been previously reported, the sites of modification and their physiological significance are not known. This study reports a comprehensive identification of phosphorylation sites in CAF-1 purified from *Saccharomyces cerevisiae*. We identified five phosphorylation sites in the Cac1 subunit, including Ser-503 that is evolutionary conserved and is phosphorylated by Cdc7/Dbf4 *in vitro*. Site-directed mutagenesis experiments indicated that Ser-503 phosphorylation is important for heterochromatin-mediated silencing of reporter genes integrated into sub-telomeric regions. In contrast, a lack of Cac1 Ser-503 phosphorylation has no effect on other known functions of CAF-1 such as DNA damage resistance or the role of pericentric chromatin in mitotic chromosome segregation. These results demonstrate that phosphorylation of Cac1 Ser-503 separates its function in heterochromatin-mediated silencing from its other functions.

3.2. Introduction

In eukaryotic cells, genomic DNA is tightly packaged within the nucleus into a structure known as chromatin. Chromatin assembly largely depends on histone chaperones involved in depositing histones onto DNA to form nucleosomes. Each nucleosome is composed of 147 base pairs of DNA wrapped nearly twice around an octamer of histone proteins containing two molecules each of the histones H2A, H2B, H3 and H4 (Luger, 1997). Formation of this nucleoprotein complex has a profound influence on a variety of biological processes important for genome function, including gene expression, DNA damage resistance and the maintenance of genomic integrity. Proliferating cells are faced with the dual challenge of not only of correctly replicating their DNA, but also of faithfully duplicating chromatin structures that need to be transmitted to daughter cells. The process of chromatin assembly has profound implications for the transmission of epigenetic marks, such as histone modifications, from parental cells to daughter cells. Duplication of chromatin structure is carried out as replication of DNA is proceeding during S-phase of the cell cycle (Krude, 2001). A crucial protein whose function is tightly linked to DNA replication is Chromatin Assembly Factor 1 (CAF-1), a protein composed of three evolutionarily conserved subunits termed p150, p60 and p48 in human cells and Cac1, Cac2 and Cac3 in yeast. CAF-1 plays a unique role in *de novo* nucleosome assembly by directly depositing newly synthesized histones H3 and H4 containing specific acetylation patterns onto nascent DNA behind replication forks (Kaufman, 1995; Krude, 1995; Marheineke, 1998; Verreault, 1996). *In vitro* and *in vivo* studies demonstrated that CAF-1 plays an important role in replication-coupled nucleosome assembly and is essential for viability of human cells (Hoek, 2003; Nabatiyan, 2004). Unlike in human cells, CAF-1 is dispensable for viability in budding yeast. Hence, *Saccharomyces cerevisiae* is a valuable model organism to study CAF-1 mutations without compromising cell viability. In *Saccharomyces cerevisiae*, disruption of genes encoding the CAF-1 subunits results in sensitivity to DNA damaging agents and transcriptional silencing defects at

heterochromatic regions, notably sub-telomeric regions, rDNA repeats and the silent mating type loci (Enomoto, 1998, 1997; Game, 1999; Kaufman, 1997; Linger, 2005; Monson, 1997). A synergistic reduction in silencing of reporter genes integrated near telomeres (hereafter termed telomeric silencing), increased sensitivity to MMS and proliferation defects are observed when *CAC* gene mutations are combined with mutation in any of the genes encoding the four subunits of another nucleosome assembly factor known as the Hir protein complex. This suggests that CAF-1 and the Hir protein complex act in redundant *de novo* assembly pathways (Kaufman, 1998; Qian, 1998; Sharp, 2002). Consistent with this, defects in chromatin structure are more prominent in cells lacking both CAF-1 and Hir proteins than in single mutants.

In both yeast and human cells, CAF-1 preferentially promotes nucleosome assembly onto newly replicated DNA, as opposed to non-replicating DNA. This specificity is achieved through a direct interaction between human p150 (yeast Cac1) and a replication factor known as Proliferating Cell Nuclear Antigen (PCNA) (Krawitz, 2002; Shibahara, 1999). PCNA is a homo-trimeric ring-shaped protein that is essential for tethering DNA polymerases to sites of DNA synthesis during replication and repair (Moldovan, 2009). Several biochemical and genetic studies in yeast and human cells have firmly established that the interaction between the largest subunit of CAF-1 and PCNA is necessary for its role in replication-coupled nucleosome assembly. In addition, critical surfaces of p150 and PCNA have been identified as essential for CAF-1 function in DNA replication and DNA repair-coupled nucleosome assembly (Moggs, 2000). For example, specific PCNA mutations that reduce heterochromatin-mediated silencing in yeast also weaken PCNA binding to CAF-1 *in vitro* and significantly decrease the association of Cac1 with chromatin *in vivo* (Zhang, 2000). Furthermore, a recent study has found that the p150 subunit of CAF-1 contains two distinct PCNA interacting peptides termed PIP1 and PIP2 (Ben-Shahar, 2009). PIP2 presents a single amino acid substitution from the PIP consensus sequence present in many other proteins. The highly conserved glutamine residue present in consensus PIPs is replaced by a lysine in vertebrate CAF-1 p150 PIP2. As a result of this

single amino acid difference from the consensus, PIP2 binds weakly to PCNA *in vitro* but, in spite of this, PIP2 is essential for CAF-1 mediated replication-dependent nucleosome assembly *in vitro* and for targeting CAF-1 to DNA replication foci *in vivo*. This suggests the existence of another level of regulation that controls the affinity of CAF-1 PIPs for PCNA and, ultimately, the function of CAF-1 *in vivo*.

Human p150 is directly phosphorylated *in vitro* by Cdc7/Dbf4, a conserved protein kinase that is essential for initiation of DNA replication. Phosphorylation of p150 increases its affinity for PCNA and thereby promotes replication-dependent nucleosome assembly *in vitro* (Gerard, 2006). In addition, several lines of evidence support the notion that reversible protein phosphorylation regulates the function of CAF-1 both *in vivo* and *in vitro*. First, based on metabolic labelling with [^{32}P] orthophosphate, the p150 and p60 subunits of CAF-1 are phosphorylated *in vivo*. Second, the activity of S-phase specific kinases (Cdc7/Dbf4 and cyclin A/E-dependent kinases) and phosphatases promote the nucleosome assembly activity of CAF-1 *in vitro* (Gerard, 2006; Keller, 2000). Moreover, phosphorylation of CAF-1 is induced in response to UV irradiation of human cells (Martini, 1998). Although these studies are interesting, neither the sites of phosphorylation nor their *in vivo* function were established.

In order to obtain a more complete understanding of how phosphorylation controls CAF-1 activity *in vivo*, we set out to identify CAF-1 phosphorylation sites in a genetically tractable organism, the budding yeast *Saccharomyces cerevisiae*. We have mapped multiple phosphorylation sites on the Cac1 subunit of CAF-1 by mass spectrometry. We found that some of these sites contain consensus sequences phosphorylated by either cyclin-dependent kinases (CDKs) or Cdc7/Dbf4 (hereafter abbreviated DDK for Dbf4-dependent kinase). We demonstrate that DDK associates with Cac1 *in vivo* and directly phosphorylates Cac1 Ser-503 *in vitro*. Mutational analysis revealed that the phosphorylation of a specific DDK site, Cac1 Ser-503, by substituting it into an alanine residue significantly reduces epigenetic silencing of a sub-telomeric reporter gene. Interestingly, based on genetic assays, we

demonstrate that phosphorylation of Cac1 Ser-503 is not needed for other CAF-1 functions. We therefore propose that phosphorylation of Cac1 Ser-503 by DDK contributes to sub-telomeric hetechromatin without affecting other CAF-1 functions.

3.3. Material and methods

3.3.1. Strains and plasmids

To create the pRS315 plasmid containing the wild-type and mutant alleles of *CAC1*, a 2.1Kb fragment containing the *CAC1* promoter, full-length ORF and a FLAG3 tag was amplified from genomic DNA of a W303 yeast strain using the primers (5'-AAACTGCAGCGGCTTGCATATATATGCTCACG-3') and (5'-TAAGAATGCGGCCGCCGCCCCGGGTTATCGCGATTTATC-3'), in which the endogenous *CAC1* gene was previously tagged at its C-terminus with three FLAG epitopes. The PCR product was digested with *Not*1 and *Pst*1 and cloned into the vector pRS315 (*ARS CEN6 LEU2*). The entire *CAC1* gene was sequenced to verify that there were no PCR-induced mutations. All the pRS315/*cac1-FLAG3* mutant plasmids were generated using the QuickChange site-directed mutagenesis kit (Stratagene) following the manufacturer's instructions. Plasmids for *E. coli* expression of His-tagged Cac1 and Hst3 fragments were generated as follows. A PCR product encoding residues 226 to 607 of Cac1 was inserted between the *Nco*I and *Not*I sites in-frame with the C-terminal His6 tag of the pET28a(+). A PCR product encoding residues 46 to 377 of Hst3 was inserted between the *Nco*I and *Not*I sites in-frame with the C-terminal His6 tag of the pET28a(+).

3.3.2. Yeast strains

The PKY618 strain (W303 MATα *cac1Δ::hisG hir1Δ::HIS3 ADE2-VR URA3-VIIL*) used in this study was kindly provided by Dr. Paul Kaufman (University of Massachusetts) (Krawitz, 2002). The S288C strains expressing Cac2-TAP or Cdc7-TAP were retrieved from the yeast TAP fusion library available from Open Biosystems. Strains expressing *cac1-FLAG3* phosphorylation site mutants were generated by introducing pRS315 plasmids containing the wild-type and mutant *CAC1* alleles into PKY618. The strain co-expressing wild-type Cac1-FLAG3 and Cdc7-TAP was created by introducing the pRS315 plasmid

containing the wild-type *CAC1-FLAG3* allele into the *CDC7-TAP* Open Biosystems strain. Standard yeast manipulations were performed as described (Amberg, 2005).

3.3.3. Purification of yeast CAF-1

For Cac2-TAP purification, exponentially growing yeast cells (4L at approximately 1.8 x 10^7 cells/ml) were harvested by centrifugation and washed twice with ice-cold water. All further steps were carried out on ice. The pellet was resuspended in an equal volume of lysis buffer (25mM Tris-HCl pH 7.5, 25mM NaCl, 0.01% NP-40, 1mM EDTA) with EDTA-free protease inhibitor cocktail (Roche), 1mM Pefabloc and 1X phosphatase inhibitors (2mM imidazole, 1.15 mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate) The resuspended cells were frozen in liquid nitrogen and lysed in a freezer mill using 4 cycles with 2 min of grinding and 2 min of cooling per cycle. The resulting cell lysate powder was thawed out on ice, incubated with 250 U of benzonase on ice for 30 min and the extract was clarified by centrifugation at 20,000 rpm for 30 min. The supernatant was then incubated with 250 µl of IgG-Sepharose beads for 2 h. The beads were recovered by centrifugation, extensively washed with lysis buffer followed by several washes with TEV cleavage buffer (10mM Tris-HCl, pH 7.5, 25mM NaCl, 0.01% NP-40, 1mM DTT) containing 1X phosphatase inhibitors (2mM imidazole, 1.15 mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate). In order to elute the proteins, the beads were incubated with 150 U of TEV protease, at 16^{o}C for 2 h in TEV cleavage buffer. Eluted proteins were subsequently incubated with calmodulin binding buffer (10mM Tris-HCl pH 7.5, 25mM NaCl, 0.01% NP-40, 1mM magnesium acetate, 1mM imidazole, 2mM calcium chloride and 10mM 2-mercaptoethanol) supplemented with EDTA-free protease inhibitor cocktail (Roche) containing calmodulin-agarose beads for 2 h at 4^{o}C. After incubation, the beads were extensively washed with calmodulin binding buffer and eluted with SDS-PAGE sample buffer. Purified proteins were resolved by SDS-PAGE and visualized by Coomassie Blue staining.

3.3.4. Recombinant proteins

Recombinant *Saccharomyces cerevisiae* Cdc7 and Dbf4 were co-expressed in baculovirus-infected Sf9 cells and the holoenzyme was purified as previously described (Weinreich, 1999). Cac1-His6 (residues 226-607) and Hst3-His6 (residues 46-377) were expressed in Arctic Express (DE3)RIL to enhance solubility. Expression plasmids were transformed into Arctic Express (DE3)RIL and colonies that formed on LB medium plates containing 25 μg/ml kanamycin and 20 μg/ml gentamycin were selected. Liquid cultures (100ml LB containing 25 μg/ml kanamycin and 20 μg/ml gentamycin) were inoculated from single colonies and grown up to early exponential phase (OD_{600} ~0.3) at 30°C. Cultures were transferred to 16°C for about an hour. When the temperature of the culture had reached 16°C, protein expression was induced by addition of 1mM IPTG. Incubation at 16°C was resumed for 48 h and cells were harvested by centrifugation. Recombinant His-tagged proteins were purified using nickel-NTA spin columns (Qiagen) according to the manufacturer's instructions, except that 50 mM imidazole was present in the wash buffers.

3.3.5. In-gel trypsin digestion

Coomassie Blue stained protein bands were excised from the gel and incubated with shaking in a 50% acetonitrile (ACN) destaining solution. Proteins were then reduced with DTT and alkylated with iodoacetamide prior to in-gel digestion with 200 ng of trypsin for 4 h at 37^{o}C, as previously described (Gharib, 2009). The resulting peptides were then extracted from the gel using a solution containing 5% TFA and 50% ACN, dried in a Speed-Vac and redissolved in 0.2% formic acid (FA) for nano-LC-MS/MS analysis.

3.3.6. In-solution trypsin digestion and phosphopeptide enrichment

Proteins were denatured using 1% SDS. Ammonium bicarbonate was added to the denatured proteins to a final concentration of 50 mM. Proteins were reduced with 0.5 mM TCEP for 20 min at 37^{o}C, alkylated using 50 mM iodoacetamide for 20 min at 37^{o}C, and excess iodoacetamide was quenched with 5 mM DTT. A solution of 50 mM ammonium bicarbonate was added to dilute SDS to 0.1% and the mixture was subjected to digestion with 200 ng of trypsin for 4 h at 37^{o}C. After proteolysis, the digest was acidified with TFA and evaporated to dryness in a Speed-Vac. Phosphopeptide enrichment was performed using TiO_2 micro-columns containing 1.25 mg of material (GL Sciences, Tokyo, Japan). Sample loading, washing and elution steps were performed by centrifugation at 5000 rpm and 4^{o}C. Briefly, micro-columns were first equilibrated with 10 µL of 3% TFA/70% ACN. The peptide mixture was redissolved in 50 µL of 250 mM lactic acid in 3% TFA/70% ACN and loaded onto the TiO_2 column. The column was washed with 10 µL of 250 mM lactic acid in 3% TFA / 70% ACN followed by a second washing step with 30µl of 3% TFA / 70% ACN. Phosphopeptides were eluted from the TiO_2 column with 30 µL of 1% NH_4OH, pH 10. The eluate was acidified with 1 µL of TFA prior to MS analysis.

3.3.7. LC-MS/MS analysis

All MS analyses were conducted using an LTQ-Orbitrap XL hybrid mass spectrometer equipped with a nanoelectrospray ion source (Thermo Fisher Scientific, San Jose, CA) and coupled to a nano-flow LC system (Eksigent, Dublin, Ireland) and a Finnigan AS autosampler (Thermo Fisher Scientific). The chromatographic separation was performed using a trapping column (4 mm length, 360 µm i.d.) and an analytical column (10 cm length, 150 µm i.d.) packed in house with 3 µm C_{18} particles (Jupiter 300Å, Phenomenex, Torrance, CA). Peptides were eluted using a linear gradient of 5-40% ACN containing 0.2% FA for 53 min, followed by a rapid increase to 60% ACN in 3 min. The flow rate was set to 0.6 µL/min. Peptides were analyzed in data-dependent mode. For each

high resolution MS scan, the three most abundant ions with an intensity above 10,000 counts were selected for MS^2 sequencing by the LTQ-Orbitrap XL (60,000 full-width half-maximal resolution; acquisition range of 400-1600 *m/z*). Product ions corresponding to the loss of a phosphate group (*i.e.* a neutral loss of 98 Da) from the precursor peptide were subsequently selected for MS^3 fragmentation. For all nano-LC-MS experiments, an internal mass lock (protonated $(Si(CH_3)_2O))_6$ with *m/z* 445.120025) or an external calibration mixture (Caffeine, MRFA and Ultramark) were used for mass calibration and provided mass accuracy within 5ppm. MS data were analyzed using the Xcalibur software (version 2.0 SR1). Peak lists were then generated using the Mascot distiller software (version 2.1.1, Matrix science) where MS processing was performed using the LCQ_plus_zoom script. Searches of the NCBInr database containing 3,310,354 entries (NCBInr March 3 2006) were conducted using the Mascot search engine (version 2.1, Matrix Science, London, U.K.). The database searches were narrowed to yeast and allowed up to 2 missed cleavage sites for trypsin. The mass tolerance threshold for experimental MS precursor ions and MS^2 fragment ions was set to ±0.02 and ±0.5 Da, respectively. All searches were conducted to allow the following variable modifications: +80 for phosphorylation (Ser, Thr or Tyr), +16 oxidation (Met), +1 for deamidation (Asn and Gln), and + 57 for carbamidomethylation (Cys). Finally, all MS^2 and MS^3 spectra ascribed to phosphopeptides by the computer search algorithm were manually verified and validated.

3.3.8. Proliferation, telomeric silencing and DNA damage sensitivity assays

Yeast strains were grown to saturation and then diluted to a concentration of 1×10^7 cells/ml. Cells were spotted on plates in 5-fold serial dilutions. For viability tests, cells were plated on rich medium (YPD) and minimal medium lacking leucine (to select for plasmids encoding *CAC1-FLAG3*) and then placed at either 18°C or 30°C for 2-5 days. To assay telomeric silencing, the yeast strain PKY618, in which the *URA3* gene is integrated

near the telomere on the left arm of chromosome VII was used as described (Gottschling, 1990). Serial dilutions were spotted on media lacking leucine and supplemented with 5-fluoroorotic acid (5-FOA) and then grown at 30°C for 4 days. For DNA damage sensitivity experiments, cells were spotted on rich media containing 0.01% methyl methanesulfonate (MMS) or plated on YPD and then irradiated with 100J/m^2 (254 nm) in a UV crosslinker (Stratalinker 2400, Stratagene). Colony formation was monitored after 2-5 days at 30°C. For cell cycle analysis, asynchronous yeast cells were grown to saturation at either 18 or 30°C. Approximately 1 x 10^7 cells were collected and prepared to measure DNA content by flow cytometry (Hasse, 2002).

3.3.9. Co-immunoprecipitation of Cdc7-TAP and Cac1-FLAG3

Exponentially growing yeast cells (1L) were harvested by centrifugation and washed twice with ice-cold water. Extracts were prepared and TAP purification conducted as described above, except that the TEV protease cleavage step was omitted. The Lysis/IP buffer was composed of 100mM Hepes-KOH pH 7.5, 110mM potassium acetate, 5% glycerol, 0.1% Tween-20, 10mM 2-mercaptoethanol, 1mM MG132 and EDTA-free protease inhibitor cocktail (Roche), 1mM Pefabloc and phosphatase inhibitors (2mM imidazole, 1.15 mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate). To immunoprecipitate Cdc7-TAP, IgG-Sepharose beads were added to the cleared lysate and incubated for 4 h at 4°C. Beads were collected by centrifugation, washed five times with Lysis/IP buffer and resuspended in SDS-PAGE sample buffer for immunoblotting. Mouse anti-FLAG and rabbit anti-CBP antibodies were used at a 1/2500 dilution to detect Cac1-FLAG3 and Cdc7-TAP.

3.3.10. In vitro kinase assay

Extracts from bacterial cells expressing either Cac1-His6 (226-607) or Hst3-His6 (46-337) were prepared in 1X kinase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1mM DTT). Kinase assays were performed with yeast HA-Cdc7-Dbf4 in 1X kinase

buffer supplemented with 50μM ATP and radiolabelled [γ-^{32}P] ATP (10μCi). Reactions were incubated for 30 min at 37^{o}C with gentle agitation. Reaction products were resolved by SDS-PAGE, Coomassie stained for visualisation of protein bands and exposed to autoradiographic film. For MS analysis of Cac1 phosphorylation by DDK, kinase assays were performed using purified yeast Cac1-His6 (226-607) produced in *E. coli*. The Ni-NTA beads containing Cac1-His6 (226-607) were washed 3 times in 1X kinase buffer supplemented with 50μM ATP only. The phosphorylation reaction was carried out for 30 min at 37^{o}C with gentle agitation. Reaction products were then processed for in-solution digestion, phosphopeptide enrichment and MS analysis as described above.

3.3.11. Chromatin binding assay

Chromatin fractionation of control (*cac1*Δ *hir1*Δ) cells or FLAG3-tagged *CAC1* wild-type, *cac1-S503A* and *cac1-S503D* cells was performed as followed. Exponentially growing cells were harvested by centrifugation and washed twice with ice-cold water. Cells were resuspended in DTT/Tris buffer (0.1 mM Tris-HCl pH 9.4, 10 mM DTT) and incubated with shaking at 30^{o}C for 15 min. Cells were harvested at maximal speed in a tabletop centrifuge and resuspended in sorbitol/HEPES buffer (1.2M Sorbitol, 20mM HEPES pH 7.4) and 0.5 mg of zymolyase was added per gram of yeast cells. Cell wall digestion was carried out at 30^{o}C with gentle shaking. To monitor cell wall digestion, 10 μl aliquots were removed and mixed with 990 μl of 1% SDS. The percentage of lysed cells was monitored by measuring absorbance at OD_{600}. When cell lysis reached approximately 95%, the digestion was terminated by adding cold Sorbitol/PIPES/$MgCl_2$ buffer (1.2M sorbitol, 20mM PIPES pH 6.8, 1mM $MgCl_2$). Spheroplasts were collected by centrifugation in a tabletop centrifuge at maximum speed for 5 min, resuspended in ice-cold NIB buffer (0.25 M sucrose, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 15 mM MES pH 6.6, 0.8% Triton-X100, 1mM Pefabloc) and incubated in ice-cold water for 20 min for lysis. Lysed spheroplasts were harvested by centrifugation at 4000 x *g* for 5 min and subsequently washed twice with NIB buffer. Cell pellets were washed twice in buffer A

(10mM Tris-HCl pH 8.0, 0.4 M NaCl, 30 mM NaButyrate, 1mM Pefabloc), incubated on ice for 15 min and total nuclear extracts were fractionated into soluble and insoluble fractions by centrifugation 4000 x *g* for 5 min. The supernatant and the insoluble chromatin-enriched pellet were boiled in an equal volume of 2X SDS-PAGE sample buffer. The association of Cac1-FLAG3 with chromatin was monitored by probing immunoblots with an antibody against the FLAG epitope tag.

3.4. Results and discussion

3.4.1. Purification of yeast CAF-1 and identification of in vivo phosphorylation sites by mass spectrometry

To investigate the regulation of CAF-1 by phosphorylation in *Saccharomyces cerevisiae*, we purified CAF-1 from asynchronous cells and determined sites of modification by mass spectrometry. CAF-1 was purified by tandem affinity purification (TAP) from a strain expressing a Cac2-TAP fusion protein. The Cac2-TAP fusion protein was previously shown to be functional. The cell lysate was first incubated using IgG-Sepharose beads, and retained proteins were subsequently released from the resin by cleavage with TEV protease prior to a second round of affinity purification using calmodulin-agarose beads. Purified proteins were resolved through an SDS-polyacrylamide gel followed by Coomassie Blue staining (Figure 3.1a). Rather than silver staining, Coomassie Blue was used to avoid artifactual sulfatation of serine and threonine residues that can be mistakenly assigned to phosphorylation (Gharib, 2009). Cac1, Cac3 and PCNA, a known interaction partner of CAF-1, were co-purified with Cac2-CBP, confirming that the whole complex was effectively isolated. In order to identify CAF-1 phosphorylation sites, bands corresponding to Cac1, Cac2-CBP and Cac3 were excised, digested with trypsin and analyzed by LC-MS/MS using an LTQ-Orbitrap XL mass spectrometer. The high resolution and mass accuracy (<5 ppm) of this instrument, combined with fmole level peptide sequencing capability, facilitated unambiguous identification of post-translational modifications. In total, 76% of the total amino acid sequence was covered and five phosphorylation sites were identified on the Cac1 subunit (Figure 3.1b). Similar MS analyses performed on the Cac2 subunit provided 64% of the total amino acid sequence coverage (Figure 3.2a). We did not identify any phosphorylation site on Cac2. This is unexpected given that its human homologue, p60, is known to be phosphorylated, although the sites of phosphorylation were never identified (Marheineke,

1998). LC-MS/MS analyses of the Cac3 subunit provided 55 % amino acid sequence coverage and demonstrated phosphorylation of Ser-13 (Figure 3.2b). For the first time, our analyses established that the Cac1 and Cac3 subunits of *Saccharomyces cerevisiae* CAF-1 are phosphorylated *in vivo* on at least six different residues, five of which reside in the Cac1 subunit.

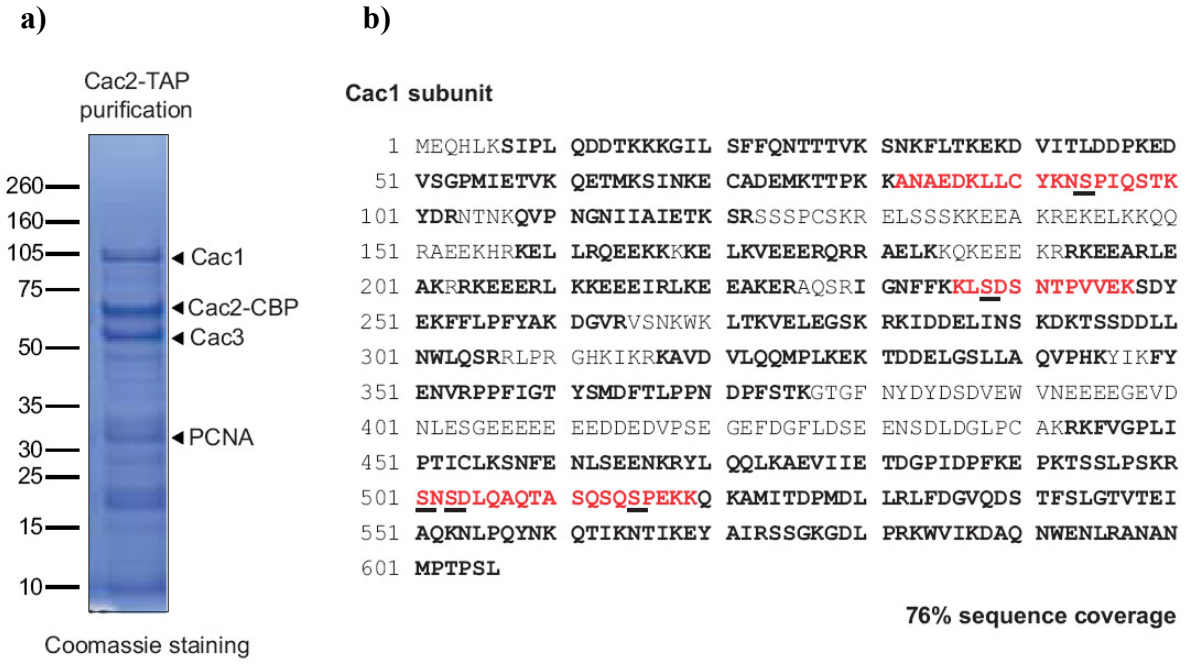


Figure 3.1 Cac1 is phosphorylated on Ser-94, Ser-238, Ser-501, Ser-503 and Ser-515 *in vivo*. a) Coomassie Blue-stained gel of CAF-1 affinity-purified from of asynchronous cells expressing Cac2-TAP. The protein bands corresponding to the Cac1, Cac2-CBP and Cac3 subunits are indicated. **b)** Amino acid sequence coverage obtained for Cac1. Tryptic peptides identified by MS/MS are shown in bold and phosphopeptides are in red. Phosphorylation sites are underlined.

a)

Cac2 subunit

```
  1 MEASHLQIYW HDSQPVYSLT FQKNSANDKL FTAGGDNKVR IWKLNRDENG
 51 QNGGVRKIES LDFLGSLTHH EQAINVIRFN SKGDVLASAG DDGQVLLWKQ
101 EDPNTQQESV VRPFGMDAET SEADENKEKW VVWKRLRGGS GATAAAEIYD
151 LAWSPDNRNI VVACMDNSIR LFDVGAGMLV CGQSDHGHYV QGVAWDPLNQ
201 FILSQSADRS LHVYGVILSS AGVVTGLKLR SKIAKAELPC PGDVLRTNYL
251 FHNETLPSFF RRCSISPCGG LVVIPSGVYK VAGDEVANCV YVYTRSGILN
301 SAGGVKNRPA IRIPSLKKPA LMAAFSPVFY ETCQKSVLKL PYKLVFAIAT
351 TNEVLVYDTD VLEPLCVVGN IHYSPITDLA WSEDGSTLLI SSTDGFCSYK
401 SIDTETQFGS RIEPPAMHAE PLDTDESAVA AKNQREAGGI VNMLPVKKIP
451 CNSSDSKKRR IHPTPVDL
```

64% sequence coverage

b)

Cac3 subunit

```
  1 MNQCAKDITH EASSIPIDLQ ERYSHWKKNT KLLYDYLNTN STKWPSLTCQ
 51 FFPDLDTTSD EHRILLSSFT SSQKPEDETI YISKISTLGH IKWSSLNNFD
101 MDEMEFKPEN STRFPSKHLV NDISIFFPNG ECNRARYLPQ NPDIIAGASS
151 DGAIYIFDRT KHGSTRIRQS KISHPFETKL FGSHGVIQDV EAMDTSSADI
201 NEATSLAWNL QQEALLLSSH SNGQVQVWDI KQYSHENPII DLPLVSINSD
251 GTAVNDVTWM PTHDSLFAAC TEGNAVSLLD LRTKKEKLQS NREKHDGGVN
301 SCRFNYKNSL ILASADSNGR LNLWDIRNMN KSPIATMEHG TSVSTLEWSP
351 NFDTVLATAG QEDGLVKLWD TSCEETIFTH GGHMLGVNDI SWDAHDPWLM
401 CSVANDNSVH IWKPAGNLVG HS
```

55% sequence coverage

Figure 3.2 Cac3 Ser-13 is phosphorylated *in vivo*. Amino acid sequence coverage obtained for **a)** Cac2 and **b)** Cac3. Peptides extracted from gel bands were analyzed by reversed-phase HPLC nanoelectrospray tandem mass spectrometry on an LTQ-Orbitrap hybrid mass spectrometer. Identified peptides are marked in bold. Observed phosphopeptides are marked in red. The identified phosphorylation site is underlined on Cac3.

Interestingly, MS/MS analysis revealed that phosphorylation of Cac1 Ser-238 (Figure 3.3a) and Ser-503 (Figure 3.3b) occur in the previously established DDK consensus sequence [S/T][D/E] (Cho, 2006). MS/MS also showed that phosphorylation of Cac1 Ser-515 (Figure 3.4a) occurs at a canonical CDK phosphorylation motif [S/T]PX[R/K] (Brown, 1999), whereas Cac1 Ser-94 (Figure 3.4c) may represent a partial CDK consensus sequence [S/T]P. We also identified phosphorylation of Cac1 Ser-501 (Figure 3.4b), a site that is located within a sequence context that has not been attributed to any known kinase. However, the fact that Cac1 Ser-501 is adjacent to Ser-503 raises the possibility that Ser-501 may also be a DDK-dependent site. Inded, *in vitro* studies have shown that, possibly by mimicking an acidic residue, a phospho-Ser/Thr can prime the phosphorylation of a neighbouring serine or threonine by DDK (Cho, 2006; Montagnoli, 2006). Taken together, our data constitute the first evidence that yeast Cac1 is a potential *in vivo* target of at least two types of kinases that regulate important cell cycle events: CDKs and DDK.

For this study, we decided to focus on the DDK-like phosphorylation sites that we identified. Both Ser-238 and Ser-503 are conserved throughout the *Saccharomyces* genus (Figure 3.5a). On the other hand, alignment of *Saccharomyces cerevisiae* Cac1 with homologous proteins from other related fungi proved more difficult because of rapid evolutionary divergence in the amino acid sequences of Cac1 proteins. In spite of this, by aligning the conserved K/E/R and E/D domains of the proteins, we were able to locate regions that correspond to *Saccharomyces cerevisiae* Cac1 Ser-238 and Ser-503. These two phosphorylation sites appear to be conserved in species such as *Candida glabrata*, *Candida albicans* and *Schizosaccharomyces pombe* (Figure 3.5b). In most cases, the corresponding serine or threonine residues are also followed by an aspartate or a glutamate, consistent with the DDK consensus sequence. Although serine is the preferred phosphorylation target, recent studies have shown that threonine can also be phosphorylated by DDK (Baker, 2010; Cho, 2006). In species where the phosphorylated Ser-238 and Ser-503 is not conserved at the same position, a serine or threonine residue, either preceded or followed by an acidic amino acid, is found in close proximity. These observations suggest that CAF-1

phosphorylation on DDK-like sites is evolutionary conserved and are consistent with a previous report showing that the p150 subunit of human CAF-1 is directly phosphorylated by DDK (Gerard, 2006).

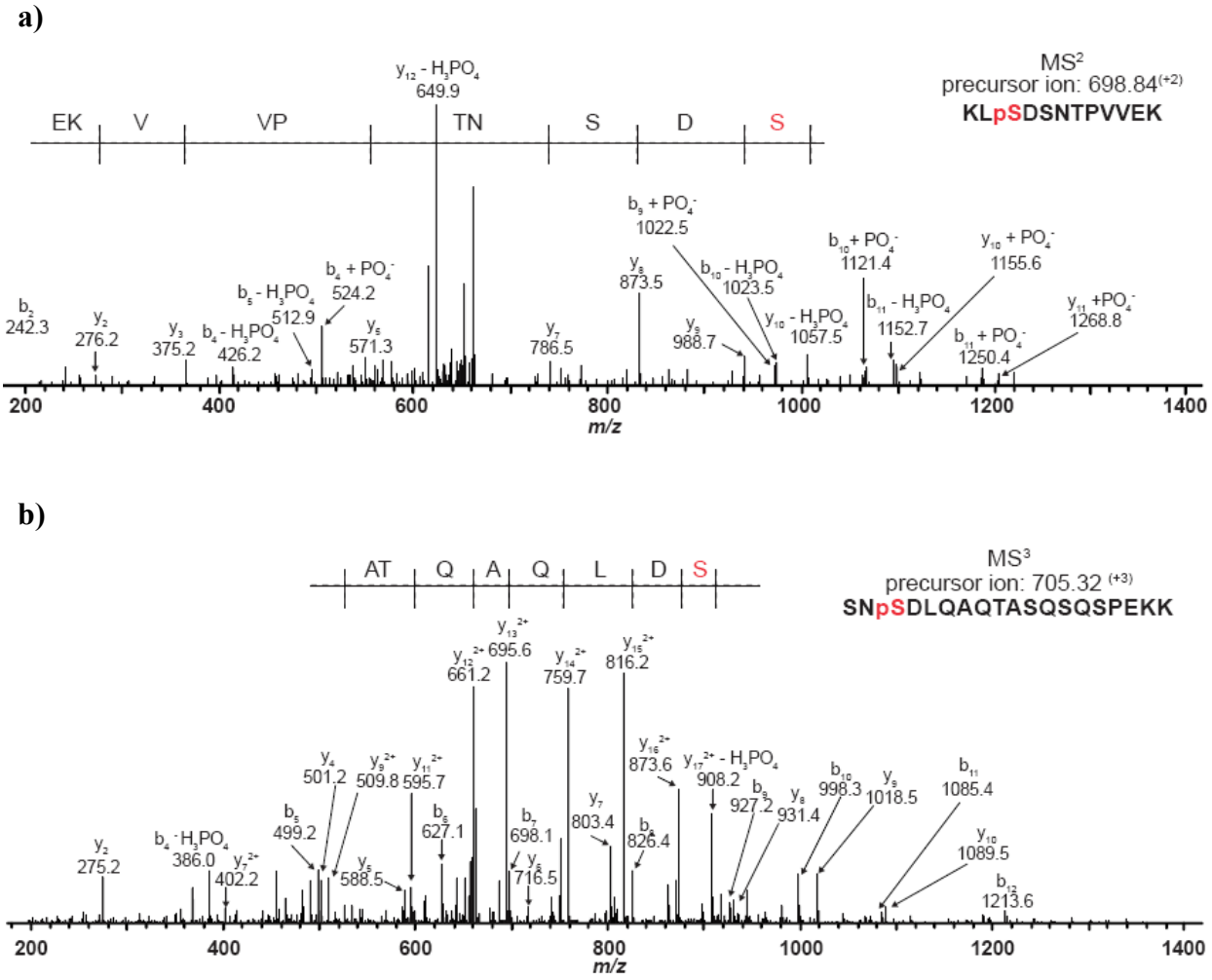


Figure 3.3 Cac1 is phosphorylated on DDK-like consensus sites. Identification of *in vivo* phosphorylation sites on Cac1 by mass spectrometry: **a)** Ser-238 and **b)** Ser-503. Residues highlighted in red correspond to the sites of phosphorylation.

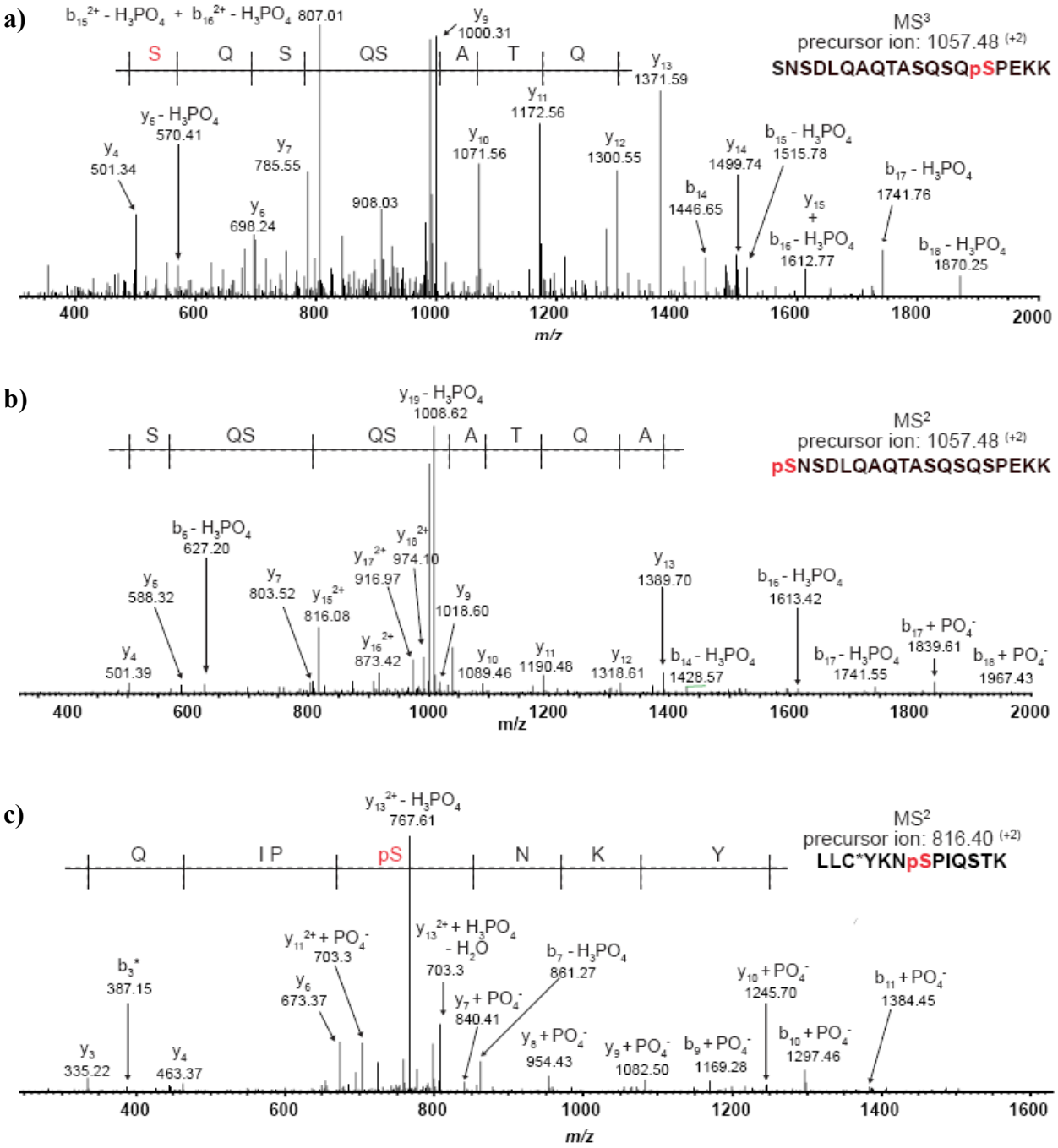
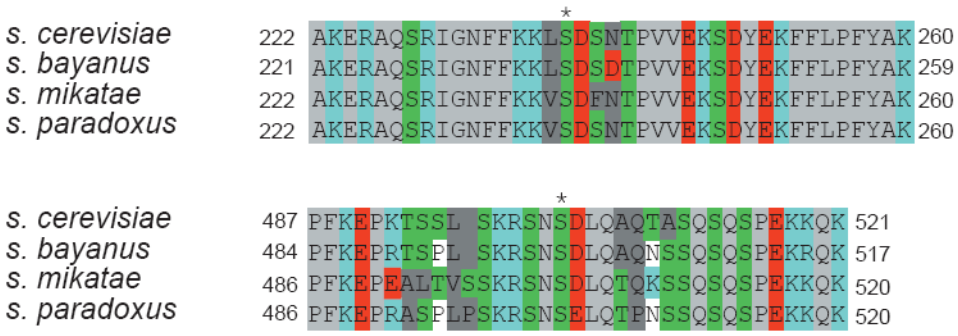


Figure 3.4 Other Cac1 phosphorylation sites identified by mass spectrometry. Mass spectra showing Cac1 phosphorylation sites: **a)** Ser-515, **b)** Ser-501 **c)** and Ser-94. Residues highlighted in red correspond to the sites of phosphorylation.

a)



b)

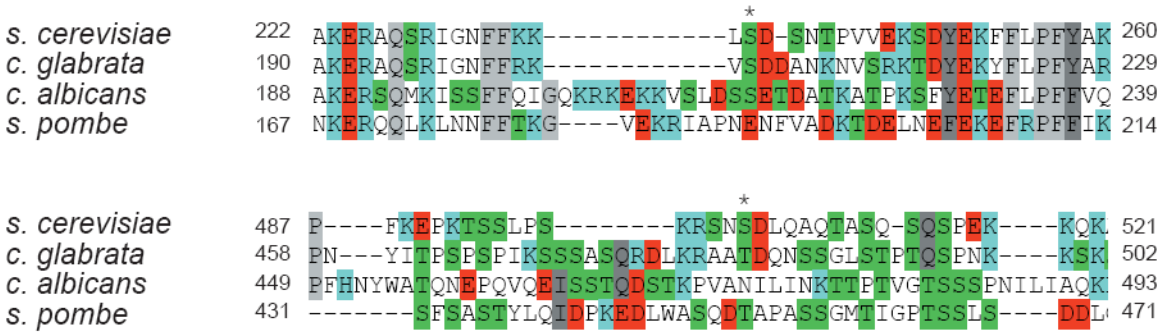


Figure 3.5 Conservation of Cac1 Ser-238 and Ser-503 throughout the *Saccharomyces* genus and other related fungi. a) Alignment of Cac1 homologues from *Saccharomyces cerevisiae*, *S. bayanus*, *S. mikatae* and *S. paradoxus* generated using the ClustalW software. **b**) Alignment of Cac1 from *Saccharomyces cerevisiae* (Q12495), *C. glabrata* (Q6FN62), *C. albicans* (Q5AFV4) and *Schizosaccharomyces pombe* (Q1MTN9) using the ClustalW software. Residues are defined as follows: identical (light grey), conserved (dark grey), basic (blue) and acidic (red). All phosphorylatable residues (Ser, Thr, Tyr) are shaded in green. The phosphorylation sites identified by MS (*Saccharomyces cerevisiae* Cac1 Ser-238 and Ser-503) are indicated by an asterisk (*)

3.4.2. Cac1 Ser-503 phosphorylation contributes to telomeric silencing, but not other functions of CAF-1

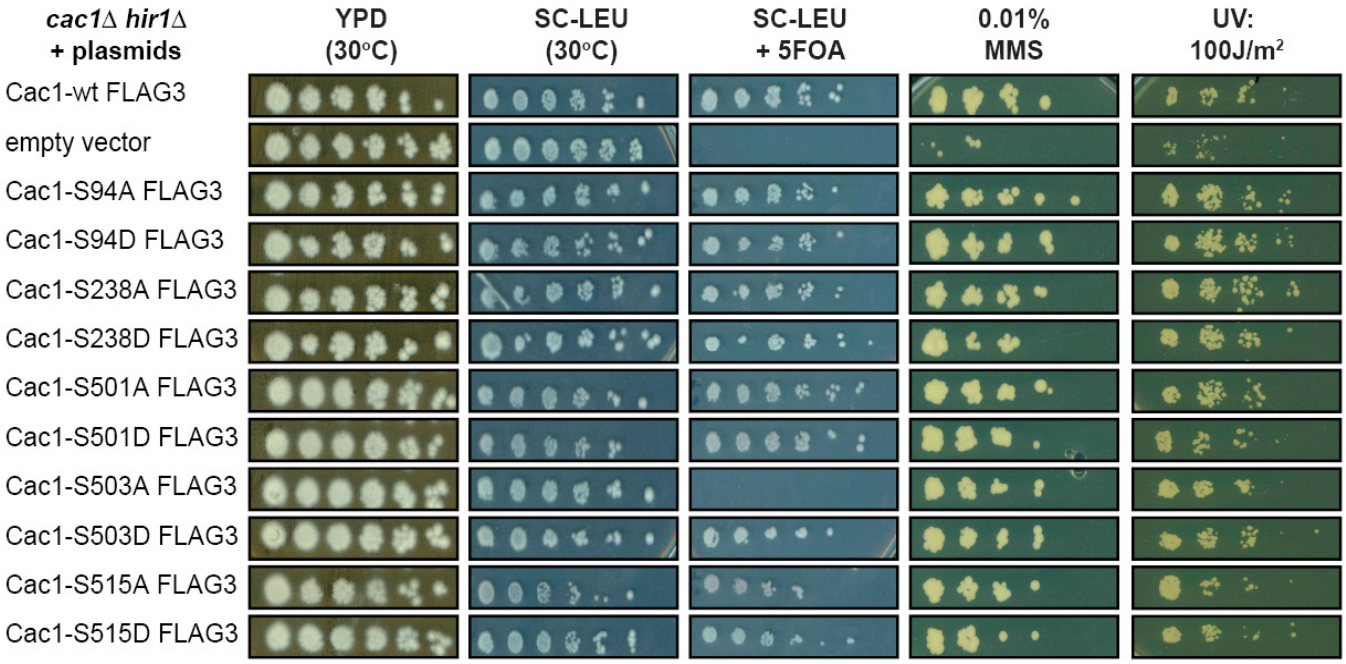
Mutations in Cac1 cause defects in telomeric silencing as well as sensitivity to DNA damaging agents such as ultraviolet radiation (UV) and the alkylating agent methyl methane sulphonate (MMS) (Enomoto, 1998, 1997; Game, 1999; Kaufman, 1997; Linger, 2005; Monson, 1997). We therefore sought to determine whether loss of Cac1 phosphorylation of any of the identified five phosphorylation sites contributes to telomeric silencing or DNA damage resistance. In order to address this, we mutated phosphorylated residues into alanine or aspartic acid to block phosphorylation or mimic constitutive phosphorylation, respectively. The mutations were generated into a low copy (*ARS CEN LEU2*) plasmid carrying the *CAC1* gene expressed from the *CAC1* promoter. Each mutant plasmid was then introduced into a strain where the chromosomal copies of *CAC1* and *HIR1* were both deleted (*cac1*Δ *hir1*Δ cells). The rationale behind the use of that strain is that the diversity and severity of phenotypes resulting from inactivation of CAF-1 are more pronounced in cells lacking any of the subunits of the HIR protein complex than in wild-type cells (Kaufman, 1998). Our experiments indicated that none of the mutations that we tested conferred an increased sensitivity to either UV radiation or MMS (Figure 3.6a). Next, we evaluated the effect of Cac1 mutations on telomeric silencing using a strain that carried a *URA3* gene integrated proximal to the telomere on the left arm of chromosome VII (*TEL VII-L*). This strain has been used extensively in studies of telomeric silencing (Gottschling, 1990; Kaufman, 1998). In this assay, a defect in silencing leads to de-repression of *URA3* near *TELVII-L* and causes cell death in the presence of 5-fluoro-orotic acid (5-FOA), a chemical that kills cells expressing the *URA3* gene. On the other hand, cell survival in the presence of 5-FOA is an indication that the *URA3* gene is properly silenced.

As expected, *cac1*Δ *hir1*Δ cells carrying an empty vector did not grow on 5-FOA plates, indicating de-repression of *URA3* at *TELVII-L* (Figure 3.6a). We then compared

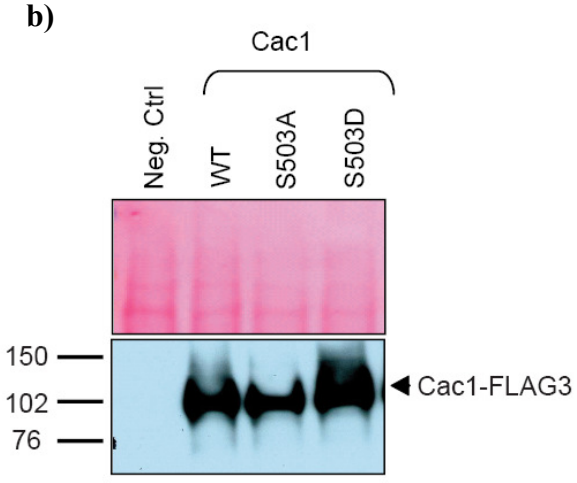
telomeric silencing in our five phosphorylation site mutants with the degree of silencing observed in *cac1*Δ *hir1*Δ cells expressing wild-type Cac1 from a plasmid. We found that only the Cac1 S503A mutant reduced telomeric silencing to the level observed in *cac1*Δ *hir1*Δ double mutants. In contrast, none of our remaining four Cac1 alanine or aspartic acid mutants showed any detectable defect compared to the degree of telomeric silencing typical of *cac1*Δ *hir1*Δ cells expressing wild-type Cac1. We tried to establish whether the disruption of telomeric silencing was due to the loss of phosphorylation on Cac1 S503 by measuring the silencing of *URA3 TELVII-L* in *hir1*Δ cells expressing a Cac1 S503D phosphomimetic mutant. Remarkably, cells expressing wild-type Cac1 and Cac1 S503D grew equally in the presence of 5-FOA, showing that the Cac1 S503D mutant protein was proficient in silencing *URA3* near *TELVII-L* (Figure 3.6a). The fact that the Cac1 S503D mutant promotes normal telomeric silencing argues that the silencing defect imparted by the Cac1 S503A mutation is due to a loss of phosphorylation on that residue rather than a structural change due to the serine-to-alanine mutation. The Cac1 S503A mutant was expressed at the same level as wild- type Cac1, showing that the silencing defect caused by alanine mutation is not due to altered expression levels of Cac1 (Figure 3.6b).

To address whether the Cac1 S503A mutation caused other phenotypes, we examined another known function of CAF-1. CAF-1 and Hir proteins both contribute to the formation of peri-centric chromatin which, in turn, plays an important role in kinetochore-microtubule attachment and mitotic chromosome segregation. As a result of checkpoint activation, due to perturbations in microtubule attachment, *cac1*Δ *hir1*Δ cells display a severe proliferation defect that is largely due to a delay in progression through G2/M phase (Sharp, 2002). This phenotype is exacerbated at low temperature, which impairs microtubule dynamics. We investigated whether *cac1-S503A hir1*Δ mutant cells displayed a slow proliferation phenotype at both physiological (30^{o}C) and cold (18^{o}C) temperature.

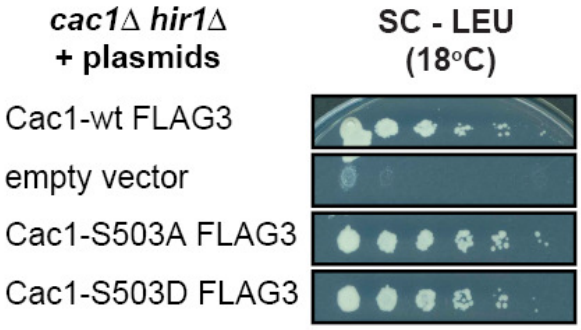
a)


cac1Δ hir1Δ + plasmids
YPD (30°C)
SC-LEU (30°C)
SC-LEU + 5FOA
0.01% MMS
UV: 100J/m²
Cac1-wt FLAG3
empty vector
Cac1-S94A FLAG3
Cac1-S94D FLAG3
Cac1-S238A FLAG3
Cac1-S238D FLAG3
Cac1-S501A FLAG3
Cac1-S501D FLAG3
Cac1-S503A FLAG3
Cac1-S503D FLAG3
Cac1-S515A FLAG3
Cac1-S515D FLAG3

b)


Cac1
Neg. Ctrl
WT
S503A
S503D
150
102
76
Cac1-FLAG3

c)


cac1Δ hir1Δ + plasmids
SC - LEU (18°C)
Cac1-wt FLAG3
empty vector
Cac1-S503A FLAG3
Cac1-S503D FLAG3

d)

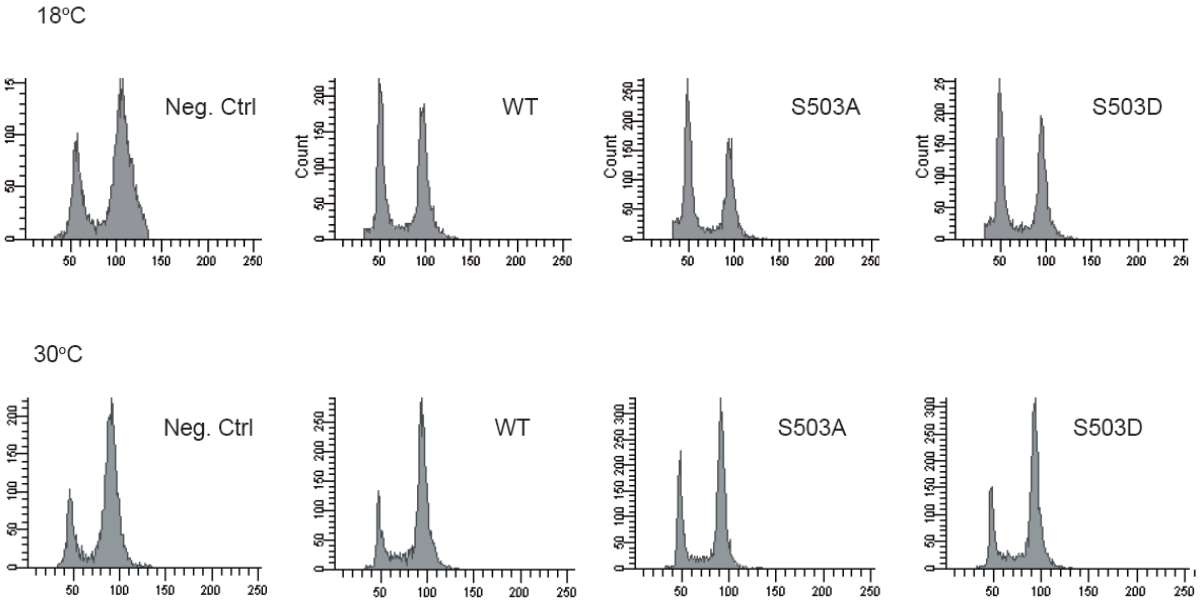


Figure 3.6 Phosphorylation of Cac1 Ser-503 is needed for telomeric silencing but not other functions of CAF-1. a) The Cac1 S503A mutation cripples telomeric silencing but does not impair cell proliferation or DNA damage resistance. Five fold serial dilutions of *cac1 hir1*Δ mutant cells were plated on rich medium (YPD) or minimal medium lacking leucine to monitor cell proliferation and –LEU+5FOA medium to assay silencing of *URA3* near *TEL VII-L*. Sensitivity to the alkylating agent methyl methane sulphonate (MMS) was monitored on YPD plates containing 0.01% MMS. Sensitivity to DNA damage induced by UV radiation was assayed by irradiating cells plated on YPD with UV at 100J/m^2. All the plates were incubated for 2-4 days at 30°C. **b)** Mutation of Ser-503 into alanine does not alter Cac1 expression. Protein extracts were prepared under denaturing conditions from 5 x 10^6 negative control cells (*cac1*Δ *hir1*Δ) or cells expressing FLAG3-tagged wild-type Cac1, Cac1-S503A and Cac1-S503D. Immunoblots were probed using an antibody against the FLAG epitope tag. **c)** *cac1-S503A hir1*Δ mutant cells do not grow slowly at 18°C. Cells were grown to exponential phase, and five-fold serial dilutions plated on minimal medium lacking leucine for 4 days at 18°C. **d)** *cac1-S503A hir1*Δ and *cac1-S503D hir1*Δ cells progress normally through the cell cycle at both 18°C and 30°C. Asynchronous cells were grown to exponential phase at 18°C and 30°C and DNA content was measured by flow cytometry. *CAC1* wild-type *hir1*Δ and *cac1*Δ *hir1*Δ cells are shown as control.

In contrast to *cac1*Δ *hir1*Δ cells, *cac1-S503A hir1*Δ mutants formed colonies at 18°C (Figure 3.6c). To further investigate whether *cac1-S503A hir1*Δ mutant cells were defective

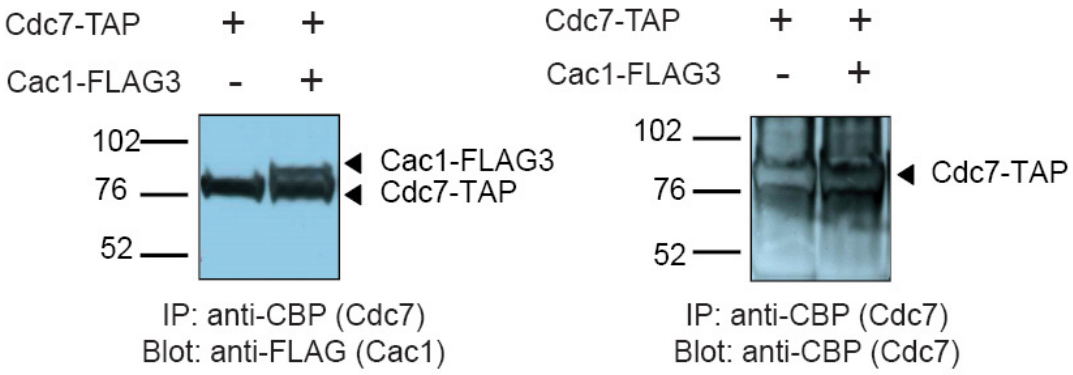
in cell cycle progression, we measured DNA content by flow cytometry on asynchronous cells cultured at 30°C and 18°C. At 30°C, the cell cycle profiles of *hir1*Δ cells expressing wild-type Cac1, Cac1 S503A or Cac1 S503D were indistinguishable (Figure 3.6d, bottom panel). As previously reported, *cac1*Δ *hir1*Δ mutants showed an accumulation of cells with G2/M DNA contents at 18°C (Figure 3.6d, top panel). However, even at 18°C, *cac1-S503A hir1*Δ and *cac1-S503A hir1*Δ mutants showed FACS profiles similar to those observed in wild-type *CAC1 hir1*Δ cells without any delay in progression through G2/M (Figure 3.6d, top panel). This result demonstrates that phosphorylation of Cac1 Ser-503 is important for telomeric silencing, but not for proliferation at cold temperature (18°C). This result also implies that the Cac1 S503A mutation is not a null mutation. Thus, a lack of Cac1 Ser-503 phosphorylation does not disrupt all functions of CAF-1, but specifically cripples telomeric silencing.

3.4.3. DDK directly phosphorylates Cac1 *in vitro*

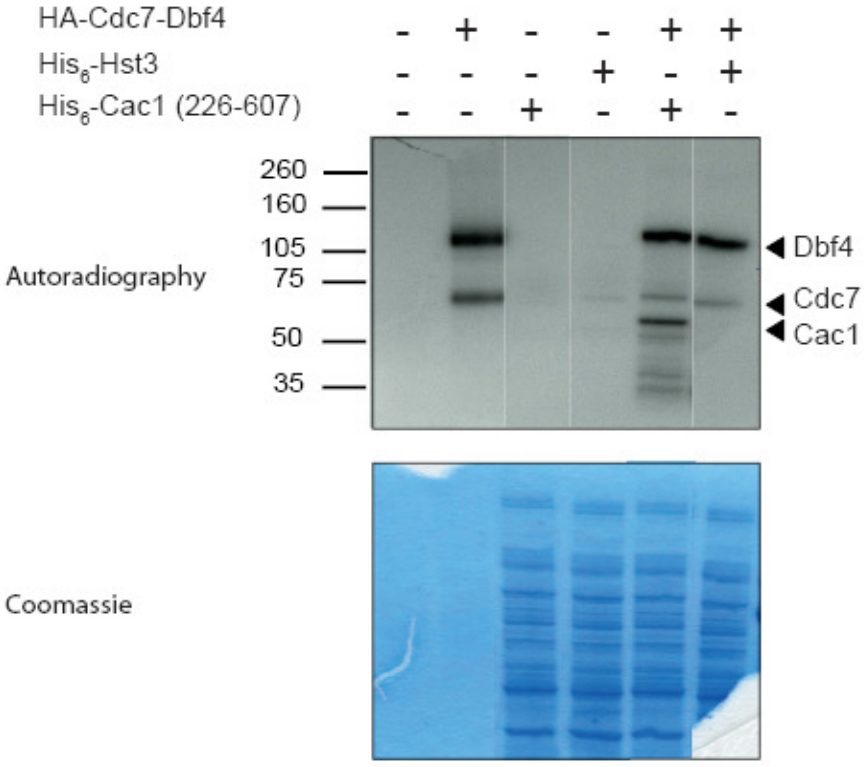
Since Ser-503 is located within a potential phosphorylation site for DDK, we sought to determine whether Cac1 might be a direct substrate of DDK. We first examined whether an interaction occurred between Cac1 and DDK *in vivo* by performing a co-immunoprecipitation experiment on whole cell extracts prepared from cells co-expressing wild-type Cac1-FLAG3 and Cdc7-TAP. A one-step (no TEV protease cleavage) immunoprecipitation was carried out using IgG-Sepharose beads to pull-down Cdc7-TAP. An immunoblot using anti-FLAG revealed that the immunoprecipitate contained Cac1-FLAG3, indicating that DDK and Cac1 form a complex *in vivo* (Figure 3.7a, left panel). The specificity of Cac1-FLAG3 detection with the FLAG monoclonal antibody was confirmed by performing a co-immunoprecipitation experiment on whole-cell extracts prepared from control cells that express Cdc7-TAP alone. Without TEV protease cleavage both the CBP and the protein A domains remain fused to Cdc7-TAP. It is important to note that this is why the FLAG antibody was also able to detect immunoprecipitated Cdc7-TAP

in the same sample. As we observed with many other TAP-tagged proteins, the protein A moiety of the TAP tag sufficiently refolds after SDS-PAGE and immunoblotting to give rise to a signal with the FLAG monoclonal antibody. This gives rise to bands denoted as Cdc7-TAP in both strains (Figure 3.7a, left panel). Consistent with this explanation, we observed a single band associated with Cdc7-TAP when blotting with an α-CBP antibody (Figure 3.7a, right panel).

a)



b)



c)

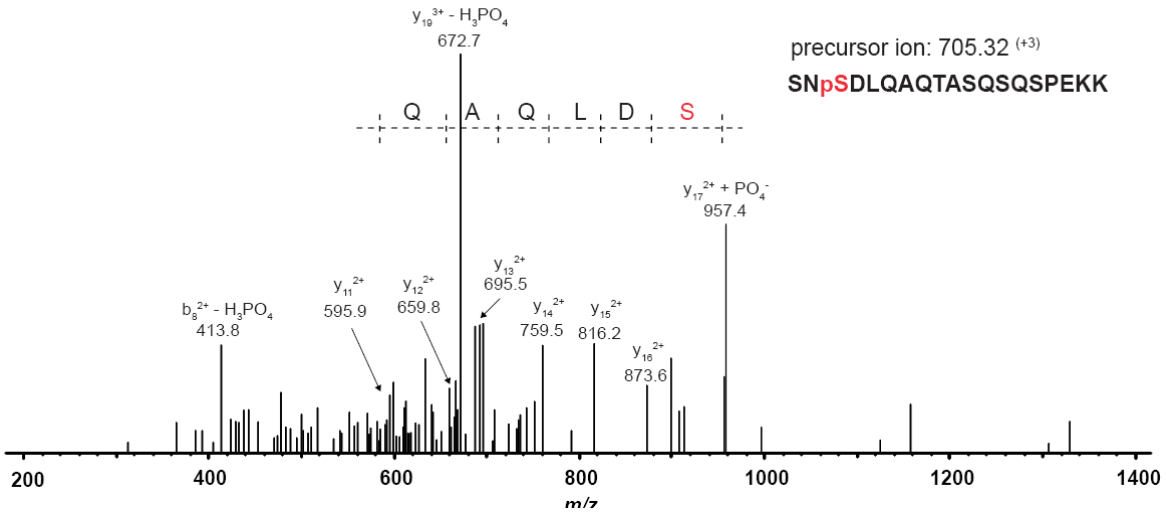


Figure 3.7 DDK interacts with Cac1 *in vivo* and directly phosphorylates Cac1 Ser-503 *in vitro*. a) Cdc7-TAP physically interacts with Cac1-FLAG3. Whole cell extracts from yeast expressing the indicated tagged proteins were incubated with IgG–Sepharose beads to pull-down Cdc7-TAP. The precipitates were then probed using antibodies against FLAG (left) and CBP (right). The top arrow indicates the band that corresponds to Cac1-FLAG3. **b)** *In vitro* kinase assay show specific phosphorylation of Cac1 by DDK. Lysates from *E. coli* strains expressing Cac1-His6 (226-606) or Hst3-His6 (46-377) as a control were incubated with (+) or without (-) HA-Cdc7/Dbf4 in the presence of radiolabeled [γ-^{32}P] ATP. Reaction products were resolved through an SDS-polyacrylamide gel, stained with Coomassie Blue and visualized by autoradiography. **c)** HA-Cdc7/Dbf4 directly phosphorylates Ser-503 *in vitro*. Recombinant Cac1-His6 (226-606) was incubated with HA-Cdc7/Dbf4 and ATP. The reaction product was digested with trypsin and phosphopeptides were enriched using a TiO_2 column. The phosphopeptide enriched sample was then analyzed by MS/MS.

Given that Cac1 and DDK physically interact with each other *in vivo*, combined with the fact that a DDK-like phosphorylation site on Cac1 is essential for telomeric silencing, prompted us to investigate whether Cac1 might be a direct substrate of DDK *in vitro*. A truncated form of Cac1 was expressed in *E. coli* and used as a substrate for the recombinant yeast DDK in the presence of radiolabelled [γ-^{32}P] ATP. Results indicated that DDK was able to phosphorylate Cac1-His6 (226-607) in a bacterial whole-cell lysate (Figure 3.7b). As a negative control, we examined the *in vitro* phosphorylation of Hst3-His6, an NAD^+-dependent protein deacetylase. No incorporation of [γ-^{32}P] was observed at

the expected molecular weight for Hst3-His6. These phosphorylation assays were performed using whole bacterial cell lysates that express histidine-tagged Cac1 or Hst3. In such conditions, the minimal [S/T][D/E] sequence required for phosphorylation by Cdc7/Dbf4 would be present in many bacterial proteins. Therefore, under these condtions, one might expect the appearance of additional radiolabelled bands resulting from phosphorylation of bacterial proteins that contain [S/T][D/E] sequences. Moreover, because of their polar and charge nature, [S/T][D/E] dipeptides would frequently be present on the surface of proteins. Nonetheless, our results clearly indicate that DDK phosphorylates Cac1 with no striking phosphorylation of bacterial proteins. Indeed, only three major distinct and sharp bands corresponding to phosphorylated Cac1-His6 (226-607) and autophosphorylated Cdc7 and Dbf4 are observed. The three minor bands below Cac1 were confirmed by MS to be Cac1 degradation products. These results demonstrate that, *in vitro*, Cac1 is phosphorylated by DDK with a considerable degree of specificity, which is consistent with Cac1 being a physiologically relevant substrate of DDK *in vivo*. We addressed whether Cac1 Ser-503 is directly phosphorylated by repeating the kinase assay using recombinant DDK, Cac1-His6 and cold ATP. The reaction products were digested with trypsin and phosphopeptides enriched using titanium dioxide (TiO_2) chromatography. The phosphopeptide enriched fraction was analysed by tandem mass spectrometry to identify phosphorylation sites. Our analysis detected a single phosphopeptide corresponding to the sequence ^{501}SNSDLQAQTASQSQSPEKK^{518} (Figure 3.7c). Careful analysis of the MS/MS spectrum revealed the presence of a neutral loss of 98 Da, corresponding to H_3PO_4 from the phosphorylated residue. The spectrum also showed a characteristic +80 Da shift on fragment ion y_{17}^{2+}, confirming that Ser-503 is phosphorylated. These results clearly showed that Cac1 Ser-503 is directly phosphorylated by DDK *in vitro*. Our analysis has also identified Ser-248 as being a direct target of Cdc7-Dbf4 *in vitro*. Although this site was not identified *in vivo* in our analyses, future studies should focus on determining whether Cac1 is phosphorylated on that specific residue as well understanding the importance of Cac1 Ser-248 phosphorylation on CAF-1 function.

In the present study, we report the unambiguous identification of multiple *in vivo* phosphorylation sites on the Cac1 subunit of *Saccharomyces cerevisiae* CAF-1. Two of these phosphorylation sites, Ser-238 and Ser-503, are conserved across the *Saccharomyces* genus and are located within consensus sequences recognized by DDK, a kinase that plays a key role in the initiation of DNA replication. For the first time, we show clear evidence that yeast CAF-1 interacts with DDK *in vivo* and is specifically phosphorylated by this kinase *in vitro*. Importantly, we demonstrate the DDK phosphorylates Cac1 Ser-503, a residue whose phosphorylation is essential for CAF-1-dependent telomeric silencing *in vivo*. In contrast a mutation that abolishes Cac1 Ser-503 phosphorylation had no effect on DNA damage resistance or cell cycle progression and cell viability at 18˚C. These results indicate that phosphorylation of Cac1 Ser-503 is not as crucial for the role of CAF-1 in peri-centric chromatin as it is in telomeric heterochromatin.

These findings raise the important question of how Cac1 phosphorylation contributes to epigenetic silencing in yeast. Although this question remains largely unanswered, several possibilities can be proposed. First, phosphorylation might mediate protein-protein interaction between CAF-1 and factors that coordinate nucleosome assembly. In human cells, phosphorylation of p150 by DDK favours CAF-1 interaction with PCNA. We found that blocking phosphorylation of Ser-503 does not significantly reduce the association of CAF-1 with chromatin (Figure 3.8). In this assay, the association of CAF-1 with chromatin partially depends upon PCNA (Zhang, 2000). Therefore, our result suggests that phosphorylation of Cac1 Ser-503 is not essential for CAF-1 to bind PCNA. However, given that telomeric silencing is sensitive to subtle changes in heterochromatin structure (Kaufman, 1998), we cannot exclude the possibility that the interaction between Cac1 S503A and PCNA is sufficiently weakened to disrupt telomeric silencing without having deleterious effects on other CAF-1 functions. Unfortunately, our attempts to conduct chromatin immunoprecipitation (ChIP) of CAF-1 at the sub-telomeric *URA3* reporter gene to determine whether CAF-1 recruitment to sub-telomeric regions

requires Cac1 Ser-503 phosphorylation were fraught with technical difficulties and inconclusive.

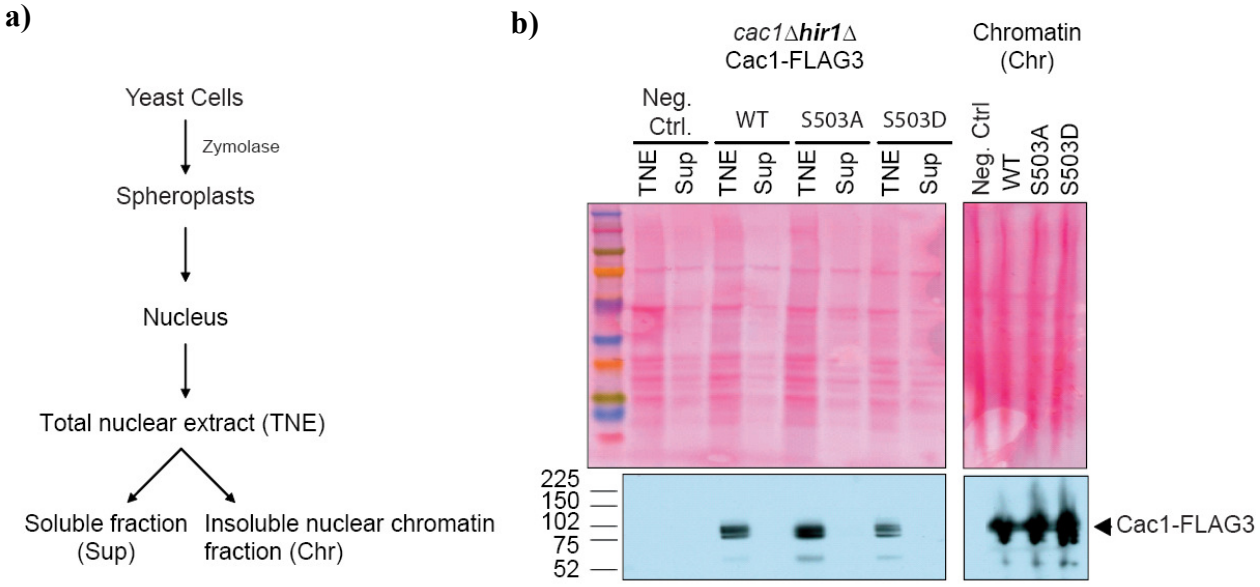


Figure 3.8 Mutation of Cac1 Ser-503 does not impair the association of CAF-1 with chromatin. a) Schematic representation of the chromatin fractionation assay. **b)** Chromatin association of CAF-1. Spheroplasts were prepared from control (*cac1*Δ *hir1*Δ) cells or cells expressing FLAG3-tagged wild type Cac1, Cac1-S503A and Cac1-S503D. Aliquots of total nuclear extract (TNE), supernatant (Sup) and Chromatin associated proteins (Chr) were resolved through an SDS 4-12% polyacrylamide gel and immunoblots were probed with an antibody against the FLAG epitope.

A second possibility is that Cac1 Ser-503 phosphorylation might promote the incorporation and/or retention of Sir proteins (Sir2, Sir3 and/or Sir4) which are key components of heterochromatin at telomeres and other loci. In *Saccharomyces cerevisiae*, the role of CAF-1 in telomeric silencing is partially redundant not only with Hir proteins,

but also with Rtt106. Rtt106 is another yeast chaperone that binds to newly synthesized histones H3/H4 (Huang, 2005). In fact, Rtt106 directly binds to Cac1 (Huang, 2005), an interaction that may help coordinate the actions of CAF-1 and Rtt106 in heterochromatin assembly. Based on ChIP assays and fluorescence microscopy, the association of Sir proteins with telomeric heterochromatin is severely perturbed in cells lacking CAF-1 and Rtt106. It is conceivable that Cac1 Ser-503 phosphorylation may promote telomeric silencing by regulating the interaction between CAF-1 and Rtt106. Regardless of the underlying molecular mechanism, our study has uncovered a novel role for DDK in heterochromatin assembly and/or stability. Future studies will be necessary to provide additional insights regarding the molecular function of DDK-mediated CAF-1 phosphorylation, the spatiotemporal regulation of this phosphorylation, and how these events promote the assembly and/or maintenance of heterochromatin structure.

3.5. References

Amberg, D.C., Burke, D.J., Strathern, J.N. (2005). Methods in yeast genetics: a cold spring harbor laboratory course manual, 2005 edn (Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press).

Baker, S.P., Phillips, J., Anderson, S., Qiu, Q., Shabanowitz, J., Smith, M.M., Yates, J.R., Hunt, D.F., Grant, P.A. (2010). Histone H3 Thr 45 phosphorylation is a replication-associated post-translational modification in *Saccharomyces cerevisiae*. Nature Cell Biology *12*, 294-298.

Ben-Shahar, T.R., Castillo, A.G., Osborne, M.J., Borden, K.L.B., Kornblatt, J. and Verreault, A. (2009). Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. Molecular & Cellular Biology *29*, 6353-6365.

Brown, N.R., Noble, M.E.M., Endicott, J.A., Johnson, L.N. (1999). The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. Nature Cell Biology *1*, 438-443.

Cho, W.-H., Lee, Y-J., Kong, S-I., Hurwitz, J., Lee, J-K. (2006). CDC7 kinase phosphorylates serine residues adjacent to acidic amino acids in the minichromosome maintenance 2 protein. Proceedings of the National Academy of Sciences of the United States of America *103*, 11521-11526.

Enomoto, S., Berman, J. (1998). Chromatin assembly factor I contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. Genes & Development *12*, 219-232.

Enomoto, S., McCune-Zierath, P.D., Gerami-Nejad, M., Sanders, M.A., Berman, J. (1997). RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function in vivo. Genes & Development *11*, 358-370.

Game, J.C., Kaufman, P.D. (1999). Role of *Saccharomyces cerevisiae* chromatin assembly factor-I in repair of ultraviolet radiation damage in vivo. Genetics *151*, 485-497.

Gerard, A., Koundrioukoff, S., Ramillon, V., Sergere, J-C., Mailand, N., Quivy, J-P., Almouzni, G. (2006). The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. EMBO Reports *7*, 817-823.

Gharib, M., Marcantonio, M., Lehmann, S.G., Courcelles, M., Meloche, S., Verreault, A., Thibault, P. (2009). Artifactual Sulfation of Silver-stained Proteins. Molecular & Cellular Proteomics *8*, 506-518.

Gottschling, D.E., Aparicio, O.M., Billington, B.L., Zakian, V.A. (1990). Position effect at *Saccharomyces cerevisiae* telomeres: Reversible repression of PolII transcription. Cell *63*, 751-762.

Hoek, M., Stillman, B. (2003). Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. Proceedings of the National Academy of Sciences of the United States of America *100*, 12183-12188.

Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E., Zhang, Z. (2005). Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proceedings of the National Academy of Sciences of the United States of America *102*, 13410-13415.

Kaufman, P.D., Cohen, J.L., Osley, M.A. (1998). Hir Proteins Are Required for Position-Dependent Gene Silencing in *Saccharomyces cerevisiae* in the Absence of Chromatin Assembly Factor I. Molecular & Cellular Biology *18*, 4793-4806.

Kaufman, P.D., Kobayashi, R., Naama, K., Stillman, B. (1995). The p150 and p60 subunits of chromatin assembly factor 1: A molecular link between newly synthesized histones and DNA replication. Cell *81*, 1105-1114.

Kaufman, P.D., Kobayashi, R., Stillman, B (1997). Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. Genes & Development *11*, 345-357.

Keller, C., Krude, T. (2000). Requirement of Cyclin/Cdk2 and Protein Phosphatase 1 Activity for Chromatin Assembly Factor 1-dependent Chromatin Assembly during DNA Synthesis. Journal of Biological Chemistry *275*, 35512-35521.

Krawitz, D.C., Kama, T., Kaufman, P.D. (2002). Chromatin Assembly Factor I Mutants Defective for PCNA Binding Require Asf1/Hir Proteins for Silencing. Molecular & Cellular Biology *22*, 614-625.

Krude, T. (1995). Chromatin Assembly Factor 1 (CAF-1) Colocalizes with Replication Foci in HeLa Cell Nuclei. Experimental Cell Research *220*, 304-311.

Krude, T., Keller, C. (2001). Chromatin assembly during S phase: contributions from histone deposition, DNA replication and the cell division cycle. Cellular and Molecular Life Sciences *58*, 665-672.

Linger, J., Tyler, J.K. (2005). The Yeast Histone Chaperone Chromatin Assembly Factor 1 Protects Against Double-Strand DNA-Damaging Agents. Genetics *171*, 1513-1522.

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

Marheineke, K., Krude, T. (1998). Nucleosome Assembly Activity and Intracellular Localization of Human CAF-1 Changes during the Cell Division Cycle. Journal of Biological Chemistry *273*, 15279-15286.

Martini, E., Roche, D.M.J., Marheineke, K., Verreault, A., Almouzni, G. (1998). Recruitment of Phosphorylated Chromatin Assembly Factor 1 to Chromatin after UV Irradiation of Human Cells. The Journal of Cell Biology *143*, 563-575.

Moggs, J.G., Grandi, P., Quivy, J-P., Jonsson, Z.O., Hubscher, U., Becker, P.B., Almouzni, G. (2000). A CAF-1-PCNA-Mediated Chromatin Assembly Pathway Triggered by Sensing DNA Damage. Molecular & Cellular Biology *20*, 1206-1218.

Moldovan, G.-L., Pfander, B., Jentsch, S. (2009). PCNA, the maestro of the replication fork. Cell *129*, 665-679.

Monson, E.K., de Bruin, D., Zakian, V.A. (1997). The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. Proceedings of the National Academy of Sciences of the United States of America *94*, 13081-13086.

Montagnoli, A., Valsasina, B., Brotherton, D., Troiani, S., Rainoldi, S., Tenca, P., Molinari, A., Santocanale, C. (2006). Identification of Mcm2 Phosphorylation Sites by S-phase-regulating Kinases. Journal of Biological Chemistry *281*, 10281-10290.

Nabatiyan, A., Krude, T. (2004). Silencing of Chromatin Assembly Factor 1 in Human Cells Leads to Cell Death and Loss of Chromatin Assembly during DNA Synthesis. Molecular & Cellular Biology *24*, 2853-2862.

Qian, Z., Huang, H., Hong, J.Y., Burck, C.L., Johnston, S.D., Berman, J., Carol, A., Liebman, S.W. (1998). Yeast Ty1 retrotransposition is stimulated by a synergistic

interaction between mutations in chromatin assembly factor I and histone regulatory proteins. Molecular & Cellular Biology *18*, 4783-4792.

Sharp, J.A., Franco, A.A., Osley, M.A., Kaufman, P.D. (2002). Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in *Saccharomyces cerevisiae*. Genes & Development *16*, 85-100.

Shibahara, K., Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Verreault, A., Kaufman, P.D., Kobyashi, R., Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. Cell *87*, 95-104.

Weinreich, M., Stillman, B. (1999). Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. The EMBO Journal *18*, 5334-5346.

Zhang, Z., Shibahara, K-I., Stillman, B. (2000). PCNA connects DNA replication to epigenetic inheritance in yeast. Nature *408*, 221-225.

4. Preliminary phosphorylation analysis of the yeast Hir protein complex during normal cell cycle progression and in response to DNA damage

4.1. Introduction

Histone proteins play a crucial role in the packaging of DNA and chromatin structure contributes to efficient chromosome segregation. As mentioned in chapter 1, the majority of histone synthesis is restricted to the S-phase of the cell cycle and is tightly coordinated with the rate of DNA replication (Osley, 1991). Both the time and the level of histone biosynthesis are critical during this phase of the cell cycle because new histones must be rapidly assembled into nucleosomes behind the replication fork, whereas excess histones can have harmful consequences for cells. Consistent with the notion that the stoichiometry of histones and DNA need to be carefully controlled, over-expression of the yeast gene pair encoding either histones H2A/H2B or H3/H4 results in non-specific binding of these highly basic proteins to chromatin and leads to frequent mitotic chromosome loss (Meeks-Wagner, 1986). Alternatively, failure to synthesize core histones during DNA replication causes cells to arrest in G2, which is accompanied by a loss of cell viability (Han, 1987; Kim, 1988). In *Drosophila melanogaster*, improper production of histone mRNAs results in abnormal oogenesis and developmental arrest during the early embryonic stages (Berloco, 2001; Sullivan, 2001).

Cells have evolved multiple mechanisms to control the level of histone transcripts and prevent the harmful effects of excess histone proteins (Osley, 1991). As cells enter S-phase, histone mRNA levels increase approximately 35-fold and return to baseline levels at the end of S-phase (Harris, 1991). In metazoans, the main regulatory system that restricts histone synthesis to S-phase acts at the post-transcriptional level and includes the regulation of histone mRNA processing and mRNA stability (Harris, 1991). In higher eukaryotes, replication-dependent histone mRNAs uniquely contain a highly conserved stem-loop structure at their 3'-end that is recognized by the stem-loop binding protein (SLBP). The cleavage of the pre-mRNA by a U7 snRNP-guided endonuclease is dependent on SLBP and is required to enhance histone mRNA translation as cells progress from G1 to S-phase

(Dominski, 2002, 1999; Sanchez, 2002). At the end of S-phase, and in response to genotoxic agents that interfere with and slow down replication, a dedicated exonuclease (3'hExo) recognizes the 3' stem-loop and then triggers rapid degradation of excess histone mRNAs (Dominski, 2003). In the budding yeast *Saccharomyces cerevisiae*, histone mRNA 3'-ends contain a poly(A) tail instead of a stem-loop (Butler, 1990; Fahrner, 1980). Although both SLBP and the exonuclease are absent, the stability of histone mRNAs in yeast is strongly regulated as histone mRNAs clearly increase during S-phase of the cell cycle even when they are transcribed from a constitutive promoter (Lycan, 1987; Xu, 1990). In yeast, a downstream sequence element, referred to as DDE, has been shown to influence the 3`-end processing and cell cycle regulation of histone mRNAs. This sequence lies approximately 110 nts downstream of the 3`-end of the histone coding sequence and appears to influence the accumulation of histones through the binding of specific protein factors (Campbell, 2002). Furthermore, genetic and biochemical evidences have shown that polyadenylation of histone mRNAs is mediated by the PolyA polymerase Trf4/Trf5 and are subsequently degraded by the nuclear exosome RRP6 (Canavan, 2007; Reis, 2007). These enzymes control the regulation of histone mRNA during the yeast cell cycle.

In addition to post-transcriptional regulation, both yeast and human cells have evolved mechanisms to control histone genes at the transcriptional level (Osley, 1991). Core histone genes are transcriptionally repressed under conditions where the presence of excess histones would be deleterious, i.e., during G1, G2 and following replication arrest during S-phase. Yeast genetic screens designed to identify repressors of histone gene transcription outside of S-phase, and during delayed replication caused by hydroxyurea, have identified four functionally related proteins, known as Hir1, Hir2, Hir3 and Hpc2 (Osley, 1987; Sherwood, 1993; Xu, 1992). The Hir proteins promote histone gene repression outside of S-phase by acting through a negative (*NEG*) *cis*-acting sequence found in three of the four divergent histone gene promoters, namely *HTA1-HTB1*, *HHT1-HHF1* and *HHT2-HHF2* (Osley, 1986; Osley, 1987). The role of Hir proteins in histone gene regulation is evolutionarily conserved. In humans, the yeast homologue of Hir1 and

Hir2 is known as HIRA. The N-terminal domain of HIRA shares homology with *Saccharomyces cerevisiae* Hir1 while the C-terminal domain of HIRA is most similar to yeast Hir2 (Kirov, 1998; Lamour, 1995). The structural similarity between HIRA and Hir1/Hir2 suggests that HIRA is a functional homologue of both Hir1 and Hir2 in higher eukaryotes. Interestingly, over-expression of HIRA inhibits DNA synthesis and causes cell cycle arrest in S-phase (Hall, 2001). Furthermore, over-expression of HIRA also represses replication-dependent histone mRNAs in human cells (Nelson, 2002). These results are consistent with the notion that, like Hir1 and Hir2 in yeast, HIRA is a repressor of histone gene transcription in humans.

The ability of Hir proteins to repress replication-dependent histone genes needs to be suppressed in S-phase because cells that replicate their DNA need an abundant supply of newly synthesized histones to ensure their rapid and efficient deposition behind replication forks. On the other hand, the ability of Hir proteins to repress histone genes needs to be re-established in G2, M and G1 in order to prevent accumulation of excess histones (Gunjan, 2005). The molecular mechanisms that disable and restore Hir-dependent histone gene repression during normal cell cycle progression or upon replication arrest induced by genotoxic agents are poorly understood. Importantly, Hir proteins do not appear to be selectively targeted for degradation during DNA replication in S-phase. Several studies have shown that the abundance of Hir proteins remains fairly constant throughout the cell cycle (Blackwell, 2004; De Lucia, 2001; Sherwood, 1993). These results suggest that the function of Hir proteins may be regulated by post-translational modifications and/or by their association with other proteins.

Cyclin-dependent kinases (CDKs) are a family of protein kinases that play a pivotal role in the cell cycle control system. As the cell progresses through the cell cycle, the kinase activity of CDKs cyclically fluctuates and directly affects the phosphorylation status of various intracellular proteins that regulate and initiate major events of the cell cycle, including DNA replication in S-phase (Morgan, 2007). Interestingly, a link between the cell

cycle control machinery and histone gene transcription in S-phase has been established in mammalian cells (Ma, 2000; Zhao, 2000). The transcription activator of histone genes NPAT co-localizes with the G1/S-phase specific CDK, cyclinE-Cdk2, at histone gene clusters *in vivo* (Ma, 2000; Zhao, 2000). Importantly, the phosphorylation of NPAT by cyclinE-Cdk2 is crucial for activation of replication-dependent histone gene expression at the G1/S phase transition (Gao, 2003; Ma, 2000; Ye, 2003; Zhao, 1998, 2000). Also, phosphorylation of NPAT by cyclinE-Cdk2 and its association with histone gene clusters is inhibited in a p21- and p53-dependent manner in response to DNA damage (Su, 2004). As a result, histone mRNA levels are down-regulated and the deleterious effects of synthesizing excess histones in the absence of DNA replication are thereby prevented. Human HIRA was also found to be an *in vivo* substrate of S-phase specific CDKs, such as cyclinA-Cdk2 and cyclinE-Cdk2 (Hall, 2001). These findings suggest that the cyclin-dependent kinases which promote passage through S-phase may exert a dual function in the control of histone gene transcription. G1/S-phase CDKs can activate transcription factors that promote histone gene transcription, such NPAT, but they can also block the action of histone gene repressors such as HIRA. Clearly, cyclin-dependent kinases appear to play a pivotal role in regulating histone gene transcription during normal cell cycle progression from G1 into S-phase.

Upon DNA damage or delayed replication caused by genotoxic agents, the checkpoint response pathway is activated, thereby arresting cell cycle progression and providing additional time for cells to efficiently repair damaged lesions (Paulovich, 1995). DNA damage results in an abrupt decrease in DNA synthesis and, consequently, the demand for histone synthesis and nucleosome assembly is drastically reduced (Gunjan, 2003). In the yeast *Saccharomyces cerevisiae*, the DNA damage checkpoint kinase Rad53 plays a dual role in preventing the accumulation of excess histones following DNA damage during S-phase. First, Rad53 associate with and triggers the degradation of excess histones that accumulate in response to the sudden decline in rates of DNA synthesis that follows DNA damage (Gunjan, 2003; Singh, 2009). Second, Rad53 also represses the transcription

of histone genes when S-phase cells are treated with genotoxic agents that interfere with DNA replication, such as MMS and HU (Paik, In preparation). A plausible hypothesis is that, following a sudden decline in rates of DNA synthesis, the accumulation of excess histones may provide a stress signal that histone genes should be repressed by Rad53 (Paik, In preparation). An attractive possibility is that the ability of Hir proteins to repress histone genes following DNA damage and replication arrest in S-phase is re-established by Rad53-dependent phosphorylation of Hir proteins.

In this chapter, I was principally interested in understanding how phosphorylation might regulate the function of Hir proteins during normal cell cycle progression and upon replication arrest following DNA damage. Since, to our knowledge, Hir protein phosphorylation was never reported in yeast, we decided to use a mass spectrometry-based approach to identify phosphorylation sites on the four subunits of the Hir complex.

4.2. Material and methods

4.2.1. Purification of Hir3-TAP

The S288C strain expressing the *HIR3*-TAP tagged gene was retrieved from the yeast TAP-Fusion library available from Open Biosystems**.** Exponentially growing yeast cells (4L) were harvested by centrifugation and washed twice with ice-cold water. Hir3-TAP purification was then performed as described in section 3.3.3 for Cac2-TAP with the exception of the composition of the buffers used, which are different for Hir3-TAP.

- Lysis buffer: (20mM Tris, pH 7.0, 300mM NaCl, 10% glycerol, 0.1% NP-40, 1µM β-mercaptoethanol) supplemented with EDTA-free protease inhibitor cocktail (Roche), 1mM Pefabloc and 1X phosphatase inhibitors (2mM imidazole, 1.15mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate).
- TEV cleavage buffer: (10mM Tris-HCl, pH 7.0, 100mM NaCl, 10% glycerol, 0.1% NP-40, 1µM β-mercaptoethanol) supplemented with 1X phosphatase inhibitors (2mM imidazole, 1.15 mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate).
- Calmodulin binding buffer: (10mM Tris-HCl, pH 7.0, 100mM NaCl, 10% glycerol, 0.1% NP-40, 1mM magnesium acetate, 1mM imidazole, 2mM $CaCl_2$, 10mM β-mercaptoethanol) supplemented with EDTA-free protease inhibitor cocktail (Roche).

The protein complex was eluted from the calmodulin beads with SDS-PAGE sample buffer, resolved by SDS-PAGE and visualized by Coomassie Blue staining.

4.2.2. Purification of Hpc2-TAP from asynchronous, MMS and MMS + caffeine treated cells

The S288C strain expressing the *HPC2*-TAP tagged gene was retrieved from the yeast TAP-Fusion library available from Open Biosystems**.** Overnight cultures were diluted in 4L YPD to OD_{660} ~ 0.5 and incubated at 30^{o}C until cells reached an OD_{660} ~ 1.0, corresponding approximately 2 x 10^7 cells/ml. For asynchronous cultures, cells were harvested. For MMS treated cultures, exponentially growing cells were then treated with 0.035% MMS and incubated at 30^{o}C for 2 h. This concentration of MMS was selected because it activates the DNA damage response in all cells, but is not cytotoxic. Following exposure to MMS, half of the cultures were subsequently incubated with 0.5% caffeine for 30 min at 30^{o}C to inhibit the checkpoint kinases. All cultures were then harvested, washed with ice-cold water and cell pellets were stored at -80^{o}C. All further steps were carried out on ice. Cell pellets were thawed out, resuspended in an equal volume of lysis buffer (20mM Tris, pH 7.0, 300mM NaCl, 10% glycerol, 0.1% NP-40, 1µM β-mercaptoethanol) supplemented with EDTA-free protease inhibitor cocktail (Roche), 1mM Pefabloc and 1X phosphatase inhibitors (2mM imidazole, 1.15 mM sodium molybdate, 1 mM sodium orthovanadate, 4 mM sodium tartrate dihydrate) and frozen as popcorn in liquid nitrogen. Cells were then lysed by mechanical disruption in a Freezer Mill using 4 cycles with 2 min of grinding and 2 min of cooling per cycle. The resulting cell lysate powders were then thawed on ice, incubated with 250 U benzonase (an enzyme that hydrolyses nucleic acids even at very low temperature) on ice for 30 min and clarified by centrifugation at 20 000 rpm and 4°C for 30 min. The supernatants (40 ml/condition) were then incubated with 250 µl of IgG-Sepharose beads for 3 h. The beads were recovered by centrifugation, extensively washed with lysis buffer, and proteins were directly eluted in SDS-PAGE sample buffer. The isolated proteins were then resolved by SDS-PAGE and stained with Coomassie Blue.

4.2.3. In-gel trypsin digestion

Proteins were digested as described in section 3.3.5.

4.2.4. LC-MS/MS analysis

All MS analyses performed on the LTQ-Orbitrap XL hybrid mass spectrometer were conducted as described in section 3.3.7.

4.2.5. MRM analysis

The Multiple Reaction Monitoring (MRM) assay of Hpc2 tryptic peptides was performed on a 4000 Q-Trap mass spectrometer (AB/MSD Analytical Technologies, Thornhill, ON, Canada), equipped with a Nanospray II interface. The chromatographic separation was performed using a trapping column (4 mm length, 360 µm i.d.) and an analytical column (10 cm length, 150 µm i.d.) packed in-house with 3-µm C_{18} particles (Jupiter 300Å, Phenomenex, Torrance, CA). The flow rate was set to 0.6µL/min and peptides were eluted using a linear gradient from 5-40% ACN containing 0.2% FA for 53 min, followed by a rapid increase to 80% for 3 min. The 784.90 →735.90 MRM transition was monitored with a 25 ms dwell-time to detect the phosphorylated ^{328}SSSASAILPKPTTTK^{342} Hpc2 peptide in asynchronous, MMS and MMS + caffeine treated cells. The sequence of the peptide and the site of phosphorylation were confirmed by MRM by triggering an enhance product ion (EPI) scan in data-dependent mode.

4.3. Results

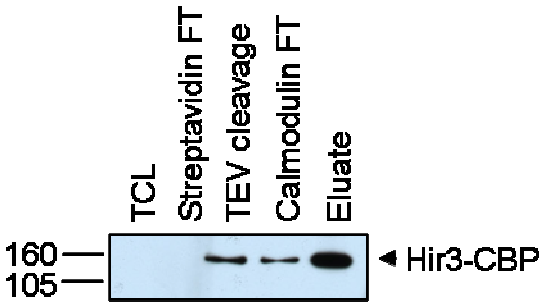
4.3.1. Purification of the Hir protein complex from yeast and identification of *in vivo* phosphorylation sites by mass spectrometry

In previous studies, HIRA, the human homologue of yeast Hir1 and Hir2 was found to be a cyclin-Cdk2 substrate that is specifically phosphorylated during S-phase of the cell cycle (Hall, 2001). In order to better understand how phosphorylation might regulate Hir proteins in yeast, we purified the Hir protein complex from asynchronous yeast cells for subsequent mapping of phosphorylation sites by mass spectrometry. The Hir3-TAP complex was purified following the standard two-step tandem-affinity purification (TAP) strategy (Puig, 2001). In order to follow the purification steps, the nitrocellulose membrane was probed with a CBP antibody that detects Hir3-TAP. Figure 4.1a shows that TAP-tagged Hir3 was efficiently isolated and enriched following IgG and Calmodulin affinity purification. The eluate was then analyzed by SDS polyacrylamide gel electrophoresis and stained with Coomassie Blue to visualize the purified Hir protein complex. Previous reports have shown that Hir1 and Hir2 co-migrate together during SDS-PAGE (Green, 2005; Prochasson, 2005). Therefore, as expected, 3 bands corresponding to the approximate molecular weights of Hir1/Hir2, Hir3 and Hpc2 were detected (Figure 4.1b).

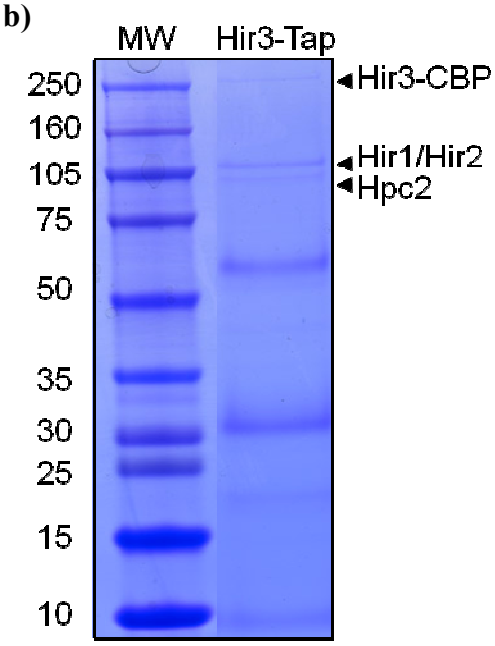
To further examine phosphorylation sites present on Hir proteins, bands corresponding to Hir1/Hir2, Hir3 and Hpc2 were excised, digested with trypsin and analyzed by LC-MS/MS using the LTQ-Orbitrap XL mass spectrometer. Mass spectrometry confirmed that Hir1 and Hpc2 co-purified with Hir3-TAP (Figure 4.1c). However, unexpectedly, we note that very low sequence coverage (18%) was retrieved for Hir1 and that peptides from the Hir2 subunit were not detectable at all in our LC-MS/MS analysis (Figure 4.1c). This may be because the TAP purification was performed at high salt concentration. The high ionic strength of the purification buffers might have

destabilized the complex resulting in low recovery of some of the subunits. Nonetheless, LC-MS/MS analysis on the Hpc2 subunit provided 64% sequence coverage and confirmed the phosphorylation of four novel sites, notably Ser-47, Ser-265, Ser-307 and Ser-435. MS/MS spectra are shown in figure 4.2a-d and the identified phosphopeptides are listed in table 4.1. These results provided the first evidence that Hpc2 is phosphorylated *in vivo*.

a)



b)

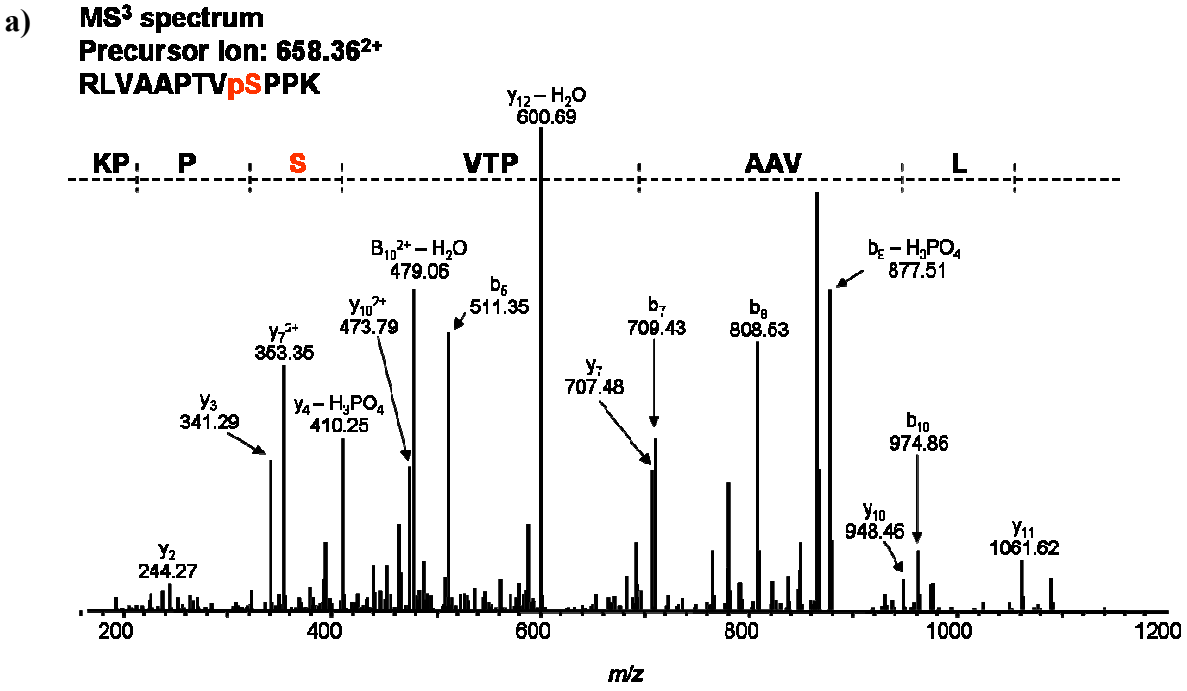


c)

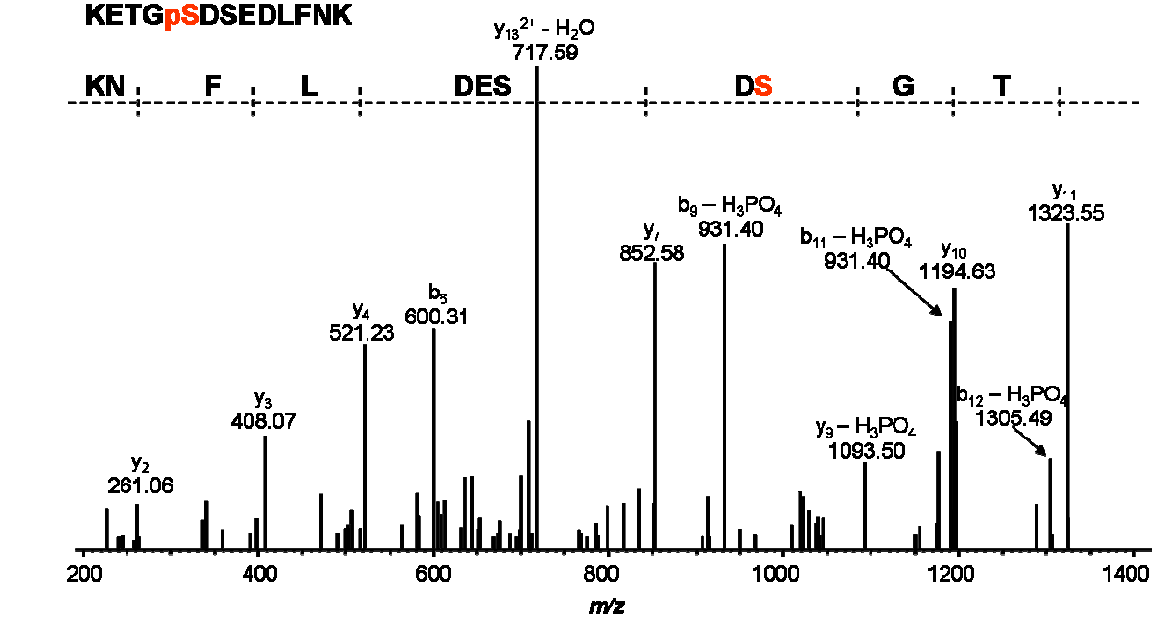
Protein ID	% coverage
Hir1	18%
Hir2	---
Hir3	39%
Hpc2	64%

Figure 4.1 Purification and detection of the Hir protein complex using the tandem-affinity purification strategy. a) Western blot probed with a CBP antibody to monitor the purification of Hir3-CBP (CBP is a portion of the TAP tag) FT: flowthrough; TCL: total cell lysate. b) Purified proteins were analyzed by SDS-PAGE and stained with Coomassie. The protein bands corresponding to Hir1, Hir2, Hir3-CBP and Hpc2 are indicated. c) Amino acid sequence coverage obtained for each Hir protein is indicated in the table.

a)


MS3 spectrum
Precursor Ion: 658.36 2+
RLVAAPTVpSPPK
KP P S VTP AAV L
y12 – H2O 600.69
B10 2+ – H2O 479.06
b6 511.35
y10 2+ 473.79
y7 2+ 353.35
b7 709.43
b8 808.63
bE – H3PO4 877.51
y7 707.48
y3 341.29
y4 – H3PO4 410.25
b10 974.86
y10 948.46
y11 1061.62
y2 244.27
200 400 600 800 1000 1200
m/z

b)


MS3 spectrum
Precursor ion: 775.33 2+
KETGpSDSEDLFNK
KN F L DES DS G T
y13 2+ - H2O 717.59
y.1 1323.55
b9 – H3PO4 931.40
y7 852.58
b11 – H3PO4 931.40
y10 1194.63
y4 521.23
b5 600.31
y3 408.07
b12 – H3PO4 1305.49
y9 – H3PO4 1093.50
y2 261.06
200 400 600 800 1000 1200 1400
m/z

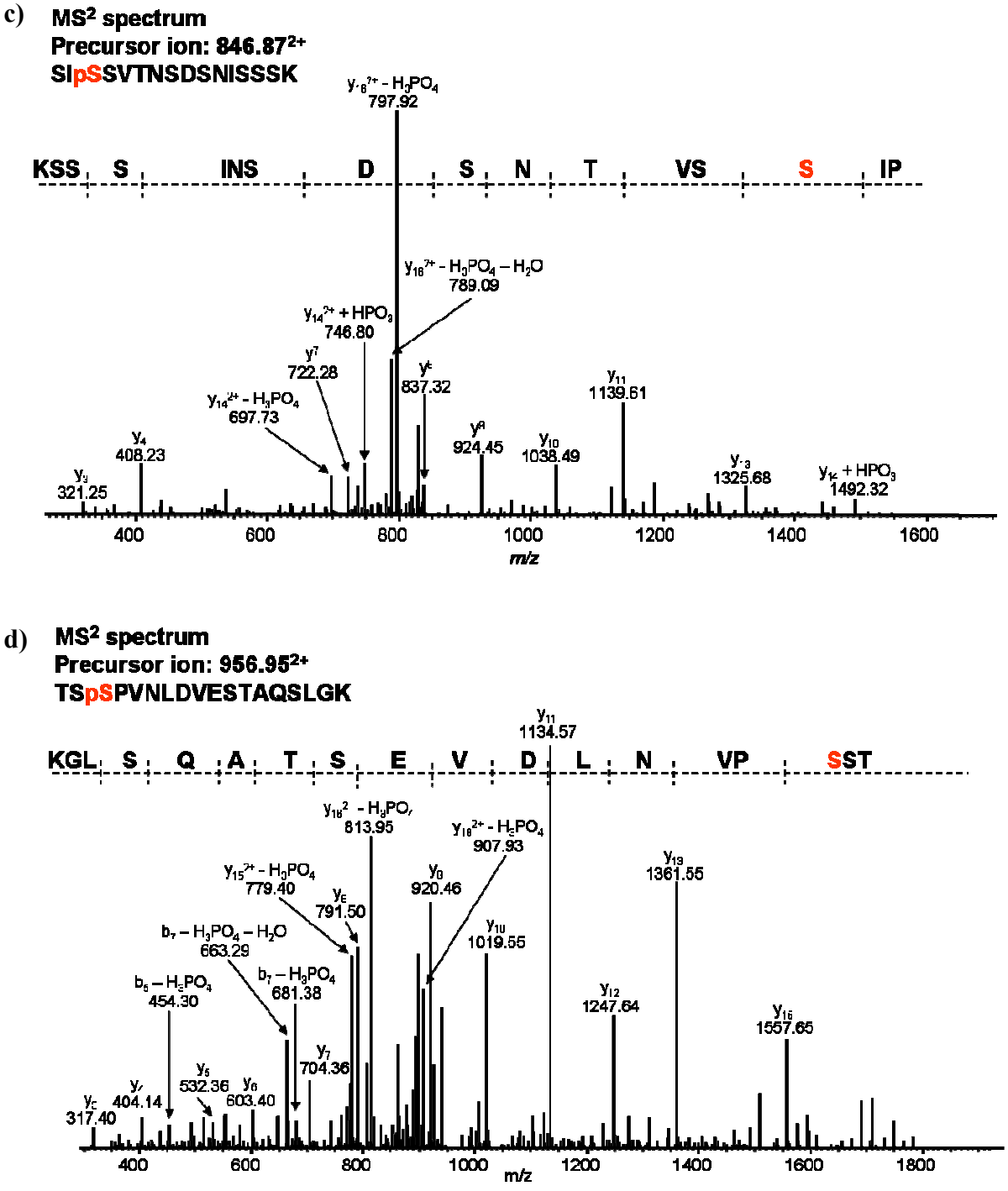


Figure 4.2 Yeast Hpc2 is phosphorylated *in vivo*. **a)** MS^3 spectrum showing phosphorylation on Ser-435. **b)** MS^3 spectrum showing phosphorylation on Ser-47. **c)** MS^2 spectrum showing phosphorylation on Ser-307. **d)** MS^2 spectrum showing phosphorylation

on Ser-265. The sequence of the identified peptide is indicated. Phosphoserine (pS) is shown in bold red. All detected *b*- and *y*-fragment ions, including loss of H_3PO_4 and/or addition of HPO_3 are denoted.

Importantly, sequence analysis revealed that Ser-265 and Ser-435 match the minimal [S/T][P] and full [S/T][P][X][K/R] consensus motif for CDK phosphorylation, respectively (Table 4.1). Furthermore, Ser-47 matches the consensus motif that is recognized and phosphorylated by the Cdc7/Dbf4 kinase ([S/T][D/E]) (Table 4.1).

Table 4.1 Summary of Hpc2 phosphorylation sites

Hpc2 phosphorylation sites*
CDK-like
^{263}TS**S**PVNLDVESTAQSLGK^{280}
^{427}RLVAAPTV**S**PPK^{438}
DDK-like
^{43}KETG**S**DSEDLFNK^{59}
Others
^{305}SI**S**SVTNSDSNISSSK^{320}

*Observed Hpc2 phosphopeptides are indicated. Identified phosphorylation sites are marked in bold red.

Our phosphorylation analysis clearly suggests that CDKs and Cdc7-Dbf4 are potentially important regulators of Hir protein function and that the phosphorylation of Hpc2 may specifically regulate the Hir protein complex and confers its functionality. This makes sense because CDKs and Cdc7-Dbf4 are known to be important regulators of major cell cycle events, including DNA replication and chromatin assembly (Morgan, 1997; Sclafani, 2000). Thus, our results suggest that histone gene regulation, also a cell cycle-

regulated process, is controlled through CDKs and/or Cdc7-Dbf4`s ability to phosphorylate the Hir protein complex.

4.3.2. Phosphorylation of Hpc2 on Ser-330 is modulated in response to DNA damage and is a target of the DNA damage checkpoint.

DNA damage that occurs during S-phase results in a rapid decrease in the rate of DNA synthesis that is accompanied by a dramatic reduction in histone transcript levels (Gunjan, 2003). Although this response plays an important role in preventing the accumulation of excess histones, its regulatory mechanisms are poorly understood. When S-phase cells are treated with genotoxic agents, the intra S-phase DNA damage checkpoint response is activated. Work performed in our laboratory, has demonstrated that Mec1/Tel1 and Rad53, key kinases involved in the DNA damage response in *Saccharomyces cerevisiae*, is essential to trigger transcriptional repression of the four core histone genes (Paik, In preparation). Recently, a proteome-wide study aimed at identifying *in vivo* targets of the DNA damage checkpoint kinases reported that Hpc2 is phosphorylated in a Mec1/Tel1- and Rad53-dependent manner (Smolka, 2007). More specifically, phosphorylation of the Hpc2 peptide ^{328}SSSASAILPKPTTTK^{342} was induced in wild type, but not in *mec1*Δ *tel1*Δ or *rad53*Δ cells in response to DNA damage (Smolka, 2007). In order to demonstrate that Hir protein phosphorylation is affected in response to DNA damage in our study, we decided to employ multiple reaction monitoring (MRM), a mass spectrometry (MS)-based quantitative approach, to confirm that Hpc2 is indeed a specific target of DNA damage checkpoint kinases and identify the precise site of Hpc2 phosphorylation.

Hpc2-TAP was purified from wild-type control cells or cells treated with the DNA alkylating agent MMS. In addition, Hpc2-TAP was also isolated from MMS-treated cells in which the DNA damage checkpoint had been de-activated. For this purpose, we used

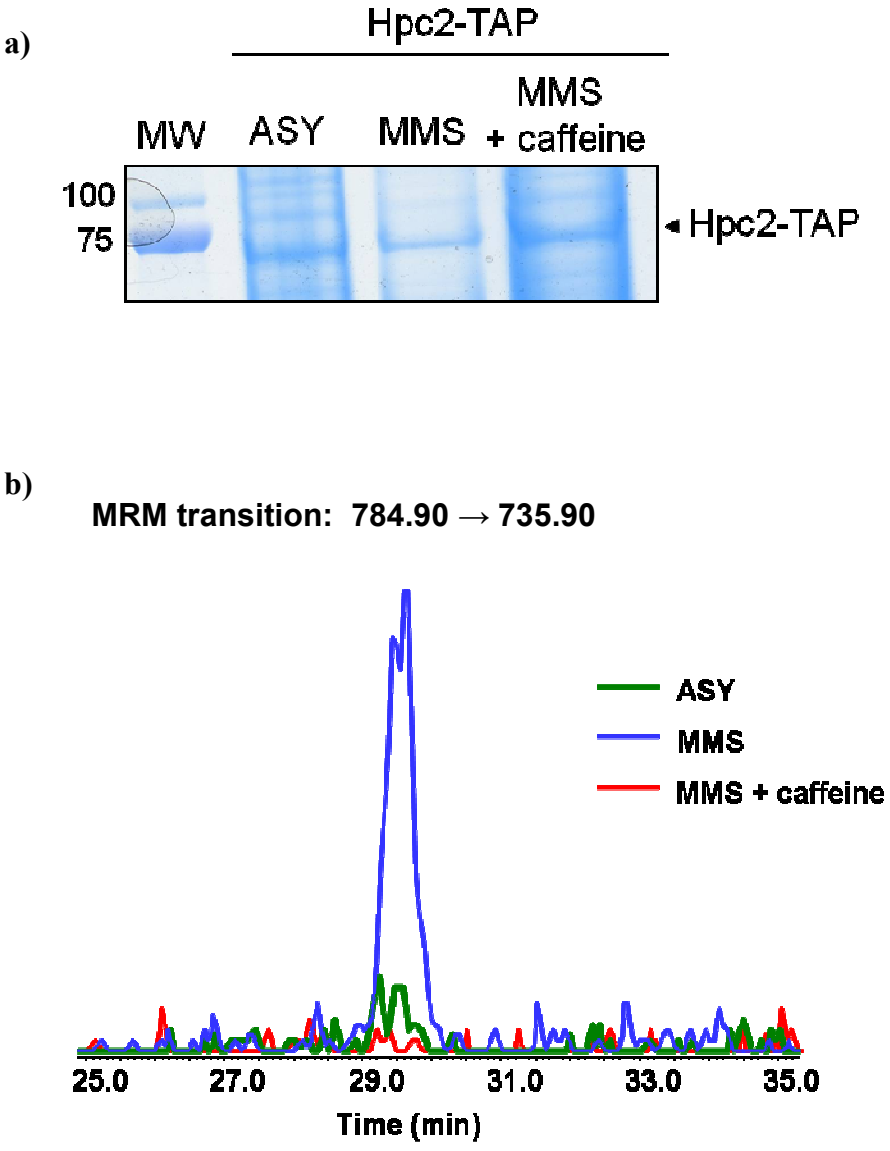
caffeine, a small molecule that has been extensively used to study the checkpoint kinases Mec1 and Tel1, the homologues of mammalian ATR and ATM, respectively (Hall-Jackson, 1999; Heffernan, 2002). In the presence of caffeine, both Mec1 and Tel1 kinase activities are effectively inhibited, thereby blocking checkpoint signaling, and eventually leading to dephosphorylation and de-activation of Rad53 and Dun1, two other kinases that require Mec1/Tel1 to be active. After the first affinity step, the eluted protein complex was analyzed by SDS-PAGE and Coomassie Blue staining (Figure 4.3a), which revealed a dominant band at the approximate molecular weight of Hpc2-TAP. The stained bands were also digested with trypsin and analyzed by LC-MS/MS which confirmed efficient isolation and identity of Hpc2 in all three samples. Next, we applied the MRM approach for sensitive and specific quantification of Hpc2 phosphorylation. More specifically, we evaluated the ^{328}SSSASAILPKPTTTK^{342} peptide by detecting a signature fragment ion that is unique to its phosphorylated form. For the Q1 mass, the theoretical phosphopeptide was calculated by adding 80 Da (one phosphate addition; HPO_3) to the doubly charged form of the peptide mass. The Q3 mass was calculated as the loss of 98 Da (loss of phosphoric acid; H_3PO_4) from the phosphopeptide. An MRM analysis based on the following Q1→ Q3 transition, 784.90 →735.90, was then performed. The dwell time and collision energy were adjusted for maximum sensitivity. Quantitative analysis was performed by calculating the peak area for the 784.90 →735.90 transition in wild type, MMS and MMS + caffeine treated samples. The results indicated that the signal of the phosphorylated ^{328}SSSASAILPKPTTTK^{342} peptide increased in response to DNA damage and decreased to basal levels after MMS-treated cells were exposed to caffeine for 30min (Figure 4.3b). These results demonstrated that, in response to DNA damage and replication arrest, Hpc2 is phosphorylated in a Mec1/Tel1 and Rad53-dependent manner. Ideally, results should have been normalized within each sample conditions for protein level by using non-phosphorylated Hpc2 peptides to correct the data. However, this was not done for this specific analysis. Nonetheless, as results clearly show that Hpc2 protein levels are clearly lower in the MMS-treated condition compared to ASY and MMS + caffeine treated conditions (Figure 4.3a), a 9-fold induction in Hpc2 phosphorylation is still observed in response to MMS-induced DNA

damage (Figure 4.3c). These observations confirm that the intensity increase observed in the MMS-treated condition is not due to variation in Hpc2 protein abundance and truly reflect an increase in Hpc2 phosphorylation.

To establish the site of Hpc2 phosphorylation induced in response to DNA damage and checkpoint activation, an information-dependent-acquisition (IDA) experiment was designed to generate an MRM-dependent MS/MS spectrum of the ^{328}SSSASAILPKPTTTK^{342} Hpc2 phosphopeptide. Hence, when the signal corresponding to the 784.90 →735.90 transition reached a specified threshold, an enhanced product ion (EPI) scan was triggered to confirm the sequence of the phosphopeptide and the precise location of the phosphorylation site.

Additionally, the EPI scan of the precursor *m/z* value 744.90, corresponding to the non-phosphorylated form of the ^{328}SSSASAILPKPTTTK^{342} peptide was also measured for comparison. Interpretation of the MS/MS spectra corresponding to the ^{328}SSSASAILPKPTTTK^{342} phosphopeptide (Figure 4.4a) revealed the presence of a fragment ion at *m/z* 736.0 corresponding to the loss of phosphoric acid from the precursor ion. Additionally, the signal observed at *m/z* 648.9963 is specifically due to the loss of H_3PO_4 from the doubly charged y_{13} fragment ion ($y_{13} - H_3PO_4$) and corresponds to Ser-330 in the ^{328}SSSASAILPKPTTTK^{342} phosphopeptide. Examination of the MS/MS spectra corresponding to the non-phosphorylated form of ^{28}SSSASAILPKPTTTK^{342} (Figure 4.4b) shows the presence of a fragment ion at *m/z* 1314.7467 corresponding to non-phosphorylated y_{13} fragment ion. Based on these findings, we can unambiguously demonstrate for the first time that Hpc2 is phosphorylated on Ser-330 in response to DNA damage. Additionally, phosphorylated Ser-330 is a specific target of DNA damage checkpoint kinases. Most direct targets of Mec1/Tel1 are preferentially phosphorylated in their [S/T][Q] motif, whereas non [S/T][Q] motifs are commonly found to be dependent on Rad53, the effector kinase that is activated by Mec1/Tel1 (Chen, 2007; Smolka, 2007). There is no strict consensus sequence for serine or threonine residues that are

phosphorylated by Rad53. However, given that Hpc2 Ser-330 is followed by an alanine residue at the n + 1 position, it seems unlikely that this residue is directly phosphorylated by Mec1/Tel1. Further experiments would be needed to test whether Rad53 or Dun1, the kinases that act downstream of Mec1/Tel1, directly phosphorylate Hpc2 Ser-330.



c)

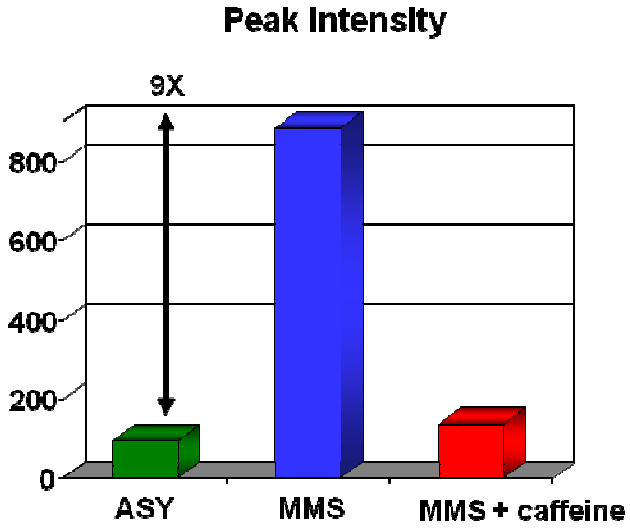
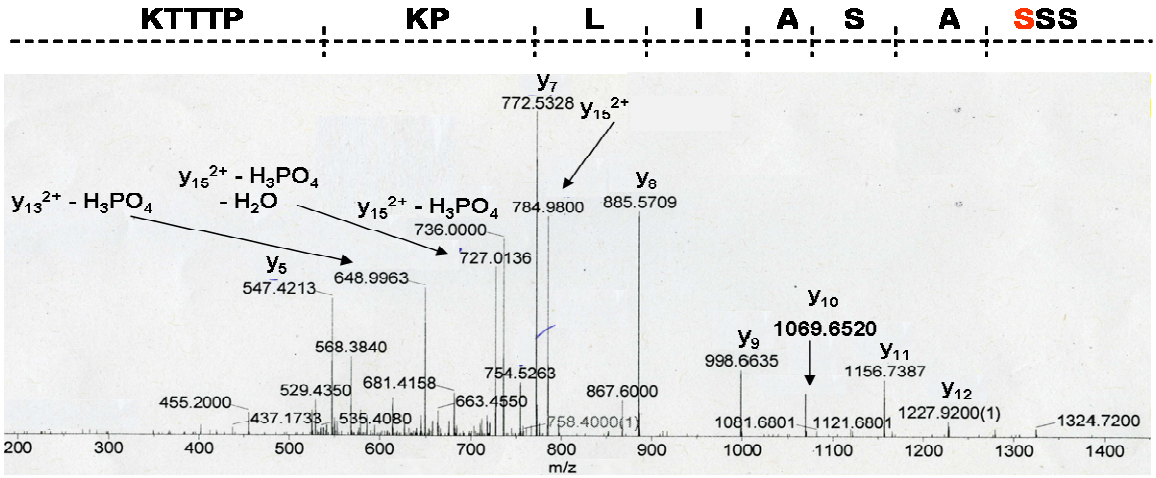


Figure 4.3 Hpc2 phosphorylation increases in response to the DNA damage checkpoint activation. a) Hpc2-TAP immunoprecipitates from asynchronous, MMS and MMS + caffeine treated cells were analyzed by SDS-PAGE and stained with Coomassie Blue. The protein band corresponding to Hpc2-TAP is indicated. **b)** Extracted ion chromatogram (XIC) displays an overlay for the 784.90 →735.90 MRM transition monitoring the loss of H_3PO_4 from the phosphorylated ^{328}SSSASAILPKPTTTK^{342} Hpc2 peptide from asynchronous (green), MMS (blue) and MMS + caffeine (red) treated cells. **c)** The relative change in the amount of the phosphorylated ^{328}SSSASAILPKPTTTK^{342} Hpc2 peptide in asynchronous (green), MMS (blue) and MMS + caffeine (red) treated cells was calculated by measuring the peak area for the 784.90 →735.90 MRM transition.

a) **EPI-MS^2 spectrum**
Precursor ion: 784.90^{2+}
SSpSASAILPKPTTTK



b) **EPI-MS^2 spectrum**
Precursor ion: 744.90^{2+}
SSSASAILPKPTTTK

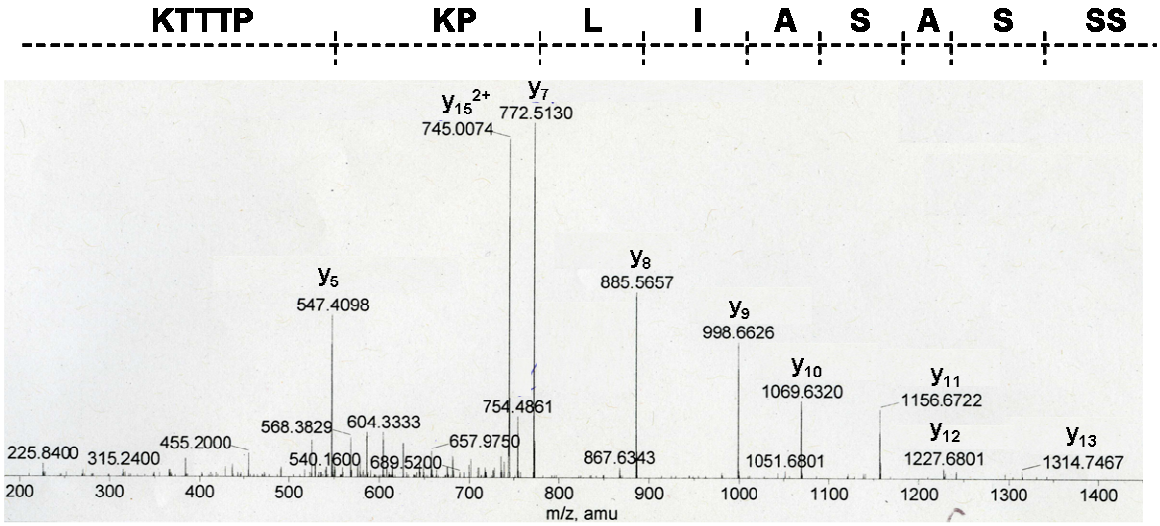


Figure 4.4 Hpc2 is phosphorylated on Ser-330 in response to DNA damage. a) MS^2 spectrum for the phosphorylated ^{328}SSSASAILPKPTTTK^{342} Hpc2 peptide showing specific phosphorylation on Ser-330. **b)** MS^2 spectrum showing the unmodified form of the ^{328}SSSASAILPKPTTTK^{342} Hpc2 peptide. The sequence of the identified peptide is indicated. Phosphoserine (pS) is shown in bold red. All detected *y*-fragment ions, including loss of H_3PO_4 are indicated.

4.4. Discussion

Hir proteins must efficiently and very precisely control the levels of histone mRNA being produced throughout the cell cycle and when replication slows down in response to DNA damage. This is important because cells need to maintain a very delicate balance between histone and DNA synthesis. Yeast cells lacking any of the four Hir proteins are incapable of repressing histone genes, either outside of S-phase or following replication arrest (Osley, 1987). In this chapter, I set out to determine whether Hir proteins are targets of protein phosphorylation *in vivo* in our quest to better understand the regulatory mechanisms that control Hir protein`s ability to repress histone genes and prevent cells from making too much or too little newly synthesized histones during different stages of the cell cycle. Here, we demonstrate that one subunit of the complex, Hpc2, is phosphorylated at multiple sites in yeast and, interestingly, many of these sites correspond to CDK and DDK consensus motifs. Although human HIRA, the homologue of yeast Hir1 and Hir2 was a known target of CDK phosphorylation (Hall, 2001), our preliminary analysis revealed the first evidence that Hpc2 is also a very prominent target of post-translational modifications.

In yeast, transcription of histone genes is activated during the transition from G1 into S-phase in a step requiring the S-phase specific CDK, and, surprisingly, repressed shortly following entry into S-phase in a Cdc7-dependent step (Hereford, 1982; Spector, 1997). Because cells need to synthesize large amounts of histones during DNA replication, one hypothesis could be that the phosphorylation of Hpc2 by CDKs may inactivate the ability of Hir proteins to repress histone genes during the G1/S-phase transition. Alternatively, phosphorylation of the CDK sites in Hpc2 may promote histone gene repression in G1, G2 and/or M-phase. Among other experiments, studies designed to determine when Hpc2 CDK sites are phosphorylated during the cell cycle would be an important first step to assess the function of this phosphorylation.

The potential function of the phosphorylation of Hpc2 on DDK sites is also unclear. DDK-mediated phosphorylation of Hpc2 cannot be absolutely essential to inactivate the Hir complex and, as a result, promote histone gene transcription. This is because cells completely lacking Cdc7 are viable provided that they also contain a mutation in MCMs that bypasses the need for MCM phosphorylation (Hardy, 1997). However, another explanation could be that Cdc7-Dbf4 may act to repress histone gene transcription in response to the decrease in the rate of DNA synthesis that occurs during late S-phase. Rates of DNA synthesis are not constant during the cell cycle. For instance, fork movement rates vary at different stages during S-phase and falls within the range of 0.3-6 Kbp/min in mammalian cells. In addition, in early S-phase, the number of replication foci is much greater than in late S-phase (Berezney, 2000; Leonhardt, 2000; Somanathan, 2001). Thus similarly to Rad53 in response to DNA damage, when histone levels have reached a critical threshold, Cdc7-Dbf4 may trigger histone gene repression as the rate of DNA synthesis is declining in late S-phase by phosphorylating the Hpc2 subunit of the Hir protein complex. A second explanation could be that Cdc7-Dbf4 may trigger phosphorylation of Hir proteins in response to replication arrest or DNA damage in S-phase and consequently cause repression of histone genes for the repair of damaged bases. Consistent with this notion, a direct role for Cdc7 in DNA repair as well as DNA damage tolerance has been proposed (Jares, 2000; Pessoa-Brandao, 2004; Sclafani, 2000).

It was previously found that Rad53 is required to repress histone gene transcription following DNA damage and replication arrest during S-phase (Paik, In preparation). In support of this, we demonstrated that a specific site on Hpc2, Ser-330, is a target of phosphorylation in response to MMS-induced DNA damage during S-phase. Given the fact that it is not located within a known consensus motif for Mec1/Tel1 kinases, we suggest that Ser-330 is a Rad53-dependent phosphorylation site. Interestingly, Hir3 has been shown to co-purify with Asf1 and Rad53 (Gavin, 2002), indicating that the Hir protein complex might interact with Rad53. Furthermore, Asf1/Hir protein function in heterochromatin-mediated silencing is also dependent on Rad53 (Sharp, 2005). These results suggest that,

following replication arrest, Rad53 may activate the ability of the Hir protein complex to repress histone gene transcription by phosphorylating Hpc2.

These preliminary findings also suggest that Hpc2 is a potentially important regulatory subunit of the Hir protein complex, whose ability to repress histone genes must be precisely controlled. The mechanism by which Hir proteins prevent inappropriate histone gene transcription during normal cell cycle progression and in response to DNA damage is not clear. One possibility implicates the RSC complex, an ATP-dependent chromatin remodeling complex that is involved in histone gene regulation. RSC associates with histone gene promoters in a manner that is strongly dependent upon both the Hir proteins and the *NEG* DNA sequence element present in histone gene promoters (Ng, 2002). Furthermore, the presence of RSC on histone gene promoters increases both outside of S-phase as well as following replication arrest. Therefore transcriptional repression of histone genes coincides with the presence of RSC at their promoters (Ng, 2002). Recently, our laboratory has demonstrated that in response to DNA damage and replication arrest, Rad53 triggers the repression of histone gene transcription by signaling the recruitment of the RSC chromatin remodeling complex to histone gene promoters (Paik, In preparation). Thus one possibility is that phosphorylation of Hpc2 by Rad53 may be needed to recruit RSC to histone gene promoters which, in turn, would alter nucleosome positioning and repress histone gene transcription. A second possibility lies in the ability of Hir proteins to bind histones and promote their assembly into nucleosomes. Mutations in Hir proteins have been found to act synergistically with mutations in CAF-1 to disrupt telomeric silencing, which depends upon formation of a repressive chromatin structure over the reporter gene (Kaufman, 1998; Qian, 1998). Furthermore, nucleosomes at the *HTA1-HTB1* locus are organized into an extended and precisely positioned array (Norris, 1988). In the absence of Hir proteins, this chromatin organization is disrupted, concomitant with the relief of transcriptional repression at the *HTA1-HTB1* locus (Dimova, 1999). Recently, Asf1 was shown to be required for disassembly of histones H3 and H4 to allow transcriptional activation of the *PHO5* promoter (Adkins, 2004). Although this study also showed that

Asf1, as well as the Hir proteins, were not needed to re-assemble nucleosomes during repression of the *PHO5* promoter, it seems possible that phosphorylation may trigger Hir proteins to be recruited to histone gene promoters to deposit histones and repress the transcription of histone genes. Irrespective of the molecular mechanism, mutation of the DNA damage-inducible Hpc2 phosphorylation site (Ser-330) to a non-phoshorylatable residue might be expected to prevent Hir proteins from repressing histone genes in response to DNA damage.

This preliminary study does not identify a direct function for the phosphorylation of Hpc2. Further studies will be needed to define the molecular function of Hpc2 phosphorylation in transcriptional repression of histone genes following DNA damage and replication arrest. In human cells, formation of senescence-associated heterochromatin foci (SAHF), which are formed by condensation of large regions of human chromosomes into isolated heterochromatic domains, requires an interaction between Asf1 and HIRA (Zhang, 2007, 2005). A novel HIRA binding protein and human homologue of Hpc2, called UBN1, has recently been described as essential for the HIRA/Asf1-mediated chromatin remodeling pathway and SAHF formation (Banumathy, 2009). Although the function of this Hpc2 homologue in humans needs to be further examined, these new findings provide evidence that yeast Hpc2 is evolutionarily conserved and that further examination of its regulation by phopshorylation in yeast might provide valuable insights into the function and regulation of the Hir complex in higher eukaryotes.

4.5. References

Adkins, M.W., Howar, S.R., Tyler, J.K. (2004). Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. Molecular cell *14*, 657-666.

Banumathy, G., Somaiah, N., Zhang, R., Tang, Y., Hoffmann, J., Andrake, M., Ceulemans, H., Schultz, D., Marmorstein, R., Adams, P.D. (2009). Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Molecular & Cellular Biology *29*, 758-770.

Berezney, R., Dubey, D.D., Huberman, J.A. (2000). Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma *108*, 471-484.

Berloco, M., Fanti, L., Breiling, A., Orlando, V., Pimpinelli, S. (2001). The maternal effect gene, abnormal oocyte (abo), of *Drosophila melanogaster* encodes a specific negative regulator of histones. Proc Proceedings of the National Academy of Sciences of the United States of America *98*, 12126-12131.

Blackwell, C., Martin, K.A., Greenall, A., Pidoux, A., Allshire, R.C., Whitehall, S.K. (2004). The *Schizosaccharomyces pombe* HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. Molecular & Cellular Biology *24*, 4309-4320.

Butler, J.S., Sadhale, P.P., Platt, T. (1990). RNA processing in vitro produces mature 3' ends of a variety of Saccharomyces cerevisiae mRNAs. Molecular & Cellular Biology *10*, 2599-2605.

Campbell, S.G., li del Olmo, M., Beglan, P., Bond, U. (2002). A sequence element downstream of the yeast HTB1 gene contributes to mRNA 3' processing and cell cycle regulation. Molecular & Cellular Biology *22*, 8415-8425.

Canavan, R., Bond, U. (2007). Deletion of the nuclear exosome component RRP6 leads to continued accumulation of the histone mRNA HTB1 in S-phase of the cell cycle in Saccharomyces cerevisiae. Nucleic Acids Research *35*, 6268-6279.

Chen, S.H., Smolka, M.B., Zhou, H. (2007). mechanism of Dun1 activation by Rad53 phosphorylation in Saccharomyces cerevisiae. Journal of Biological Chemistry *282*, 986-995.

De Lucia, F., Lorain, S., Scamps, C., Galissson, F., MacHold, J., Lipinski, M. (2001). Subnuclear localization and mitotic phosphorylation of HIRA, the human homologue of Saccharomyces cerevisiae transcriptional regulators Hir1p/Hir2p. Biochemical Journal *358 (Pt 2)*, 447-455.

Dimova, D., Nackerdien, Z., Furgeson, S., Eguchi, S., Osley, M.A. (1999). A role for transcriptional repressors in targeting the yeast Swi/Snf complex. Molecular Cell *4*, 75-83.

Dominski, Z., Erkmann, J.A., Yang, X.,Sanchez, R., Marzluff, W.F. (2002). A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. Genes & Development *16*, 58-71.

Dominski, Z., Yang, X.C., Kaygun, H., Dadlez, M., Marzluff, W.F. (2003). A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. Molecular Cell *12*, 295-305.

Dominski, Z., Zheng, L.X., Sanchez, R., Marzluff, W.F. (1999). Stem-loop binding protein facilitates 3'-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. Molecular & Cellular Biology *19*, 3561-3570.

Fahrner, K., Yarger, J., Hereford, L. (1980). Yeast histone mRNA is polyadenylated. Nucleic Acids Research *8*, 5725-5737.

Gao, G., Bracken, A.P., Burkard, K., Pasini, D., Classon, M., Attwooll, C., Sagara, M., Imai, T., Helin, K., Zhao, J. (2003). NPAT expression is regulated by E2F and is essential for cell cycle progression. Molecular & Cellular Biology *23*, 2821-2833.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Green, E., Antczak, A., Bailey, A., Franco, A., Wu, K., Yates, J., and Kaufman, P. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Current Biology *15*, 2044 - 2049.

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

Gunjan, A., Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in Saccharomyces cerevisiae. Cell *115*, 537-549.

Hall-Jackson, C.A., Cross, D.A., Morrice, N., Smythe, C. (1999). ATR is a caffeine-sensitive, DNA-activated protein kinase with a substrate specificity distinct from DNA-PK. Oncogene *18*, 6707-6713.

Hall, C., Nelson, D.M., Ye, X., Baker, K., DeCaprio, J.A., Seeholzer, S., Lipinski, M., Adams, P.D. (2001). HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel cyclin-cdk2 substrate whose expression blocks S-phase progression. Molecular & Cellular Biology *21*, 1854-1865.

Han, M., Chang, M., Kim, U.J., Grunstein, M. (1987). Histone H2B repression causes cell-cycle-specific arrest in yeast: Effects on chromosomal segregation, replication, and transcription. Cell *48*, 589-597.

Hardy, C.F.J., Dryga, O., Seematter, S., Pahl, P.M.B., Sclafani, R.A. (1997). mcm5/cdc46-bob1 bypasses the requirement for the S phase activator□Cdc7p. Proceedings of the National Academy of Sciences of the United States of America *94*, 3151-3155.

Harris, M.E., Bohni, R., Schneiderman, M.H., Ramamurthy, L., Schumperli, D., Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Molecular & Cellular Biology *11*, 2416-2424.

Heffernan, T.P., Simpson, D.A., Frank, A.R., Heinloth, A.N., Paules, R.S., Cordeiro-Stone, M., Kaufmann, W.K. (2002). An ATR- and Chk1-dependent S checkpoint inhibits replicon initiation following UVC-induced DNA damage. Molecular & Cellular Biology *22*, 8552-8561.

Hereford, L., Bromley, S., Osley, M.A. (1982). Periodic transcription of yeast histone genes. Cell *30*, 305-310.

Jares, P., Donaldson, A., Blow, J. (2000). The Cdc7/Dbf4 protein kinase: target of the S phase checkpoint? . EMBO Reports *1*, 319-322.

Kaufman, P.D., Cohen, J.L., Osley, M.A. (1998). Hir Proteins Are Required for Position-Dependent Gene Silencing in Saccharomyces cerevisiae in the Absence of Chromatin Assembly Factor I. Molecular & Cellular Biology *18*, 4793-4806.

Kim, U.J., Han, M., Kayne, P., Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. The EMBO Journal *7*, 2211-2219.

Kirov, N., Shtilbans, A., Rushlow, C. (1998). Isolation and characterization of a new gene encoding a member of the HIRA family of proteins from *Drosophila melanogaster*. Gene *212*, 323-332.

Lamour, V., Lécluse, Y., Desmaze, C., Spector, M., Bodescot, M., Aurias, A., Osley, M.A., Lipinski, M. (1995). A human homolog of the Saccharomyces cerevisiae HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. Human Molecular Genetics *4*, 791-799.

Leonhardt, H., Rahn, H.P., Weinzierl, P., Sporbert, A., Cremer, T., Zink, D., Cardoso, M.C. (2000). Dynamics of DNA replication factories in living cells. The Journal of Cell Biology *149*, 271-280.

Lycan, D.E., Osley, M. A., Hereford, L. M. (1987). Role of transcriptional and posttranscriptional regulation in expression of histone genes in *Saccharomyces cerevisiae*. Mol Cell Biol *7*, 614-621.

Ma, T., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T., Harper, J.W (2000). Cell cycle-regulated phosphorylation of p220NPAT by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. Genes & Development *14*, 2298-2313.

Meeks-Wagner, D., Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell *44*, 43-52.

Morgan, D.O. (1997). CYCLIN-DEPENDENT KINASES: Engines, Clocks, and Microprocessors. Annual Review of Cell and Developmental Biology *13*, 261-291.

Morgan, D.O. (2007). The cell cycle: principles of control (London, England, New Science Press Ltd).

Nelson, D.M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G.D., Yen, T.J., Harper, J.W., Adams, P.D. (2002). Coupling of DNA Synthesis and Histone Synthesis in S Phase Independent of Cyclin/cdk2 Activity. Molecular & Cellular Biology *22*, 7459-7472.

Ng, H.H., Robert, F., Young, R.A. Struhl, K. (2002). Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. Genes & Development *16*, 806-819.

Norris, D., Dunn, B., Osley, M. A. (1988). The effect of histone gene deletions on chromatin structure in Saccharomyces cerevisiae. Science *242*, 759-761.

Osley, M.A. (1991). The regulation of histone synthesis in the cell cycle. Annual Review of Biochemistry *60*, 827-861.

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

Osley, M.A., Lycan, D. (1987). Trans-acting regulatory mutations that alter transcription of Saccharomyces cerevisiae histone genes. Molecular & Cellular Biology *7*, 4204-4210.

Paik, J., Gunjan, A., Verreault, A. (In preparation). The Rad53 checkpoint kinase represses histone gene transcription in response to DNA damage and replication arrest.

Paulovich, A.G., Hartwell, L.H. (1995). A checkpoint regulates the rate of progression through S phase in *Saccharomyces cerevisiae* in response to DNA damage. Cell *8*, 841-847.

Pessoa-Brandao, L., Sclafani, R.A. (2004). CDC7/DBF4 functions in the translesion synthesis branch of the RAD6 epistasis group in saccharomyces cerevisiae. Genetics *167*, 1597-1610.

Prochasson, P., Florens, L., Swanson, S.K., Washburn, M.P., Workman, J.L. (2005). The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. Genes & Development *19*, 2534-2539.

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

Qian, Z., Huang, H., Hong, J.Y., Burck, C.L., Johnston, S.D., Berman, J., Carol, A., Liebman, S.W. (1998). Yeast Ty1 retrotransposition is stimulated by a synergistic

interaction between mutations in chromatin assembly factor I and histone regulatory proteins. Molecular & Cellular Biology *18*, 4783-4792.

Reis, C.C., Campbell, J.L. (2007). Contribution of Trf4/5 and the nuclear exosome to genome stability through regulation of histone mRNA levels in saccharomyces cerevisiae. Genetics *175*, 993-1010.

Sanchez, R., Marzluff, W.F. (2002). The Stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro. Molecular & Cellular Biology *22*, 7093-7104.

Sclafani, R.A. (2000). Cdc7p-Dbf4p becomes famous in the cell cycle. Journal of Cell Science *113*, 2111-2117.

Sharp, J.A., Rizki, G., Kaufman, P.D. (2005). Regulation of histone deposition proteins Asf1/Hir1 by multiple DNA damage checkpoint kinases in saccharomyces cerevisiae. Genetics *171*, 885-899.

Sherwood, P.W., Tsang, S.V., Osley, M.A. (1993). Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in Saccharomyces cerevisiae. Molecular & Cellular Biology *13*, 28-38.

Singh, R.K., Kabbaj, M.H.M., Paik, J., Gunjan, A. (2009). Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. Nature Cell Biology *11*, 925-933.

Smolka, M.B., Albuquerque, C.P., Chen, S.H., Zhou, H. (2007). Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. Proceedings of the National Academy of Sciences of the United States of America *104*, 10364-10369.

Somanathan, S., Suchyna, T.M., Siegel, A.J., Berezney, R. (2001). Targeting of PCNA to sites of DNA replication in the mammalian cell nucleus. Journal of Cellular Biochemistry *81*, 56-67.

Spector, M.S., Raff, A., DeSilva, H., Lee, K., Osley, M. A. (1997). Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle. Molecular & Cellular Biology *17*, 545-552.

Su, C., Gao, G., Schneider, S., Helt, C., Weiss, C., O'Reilly, M.A., Bohmann, D., Zhao, J. (2004). DNA damage induces downregulation of histone gene expression through the G1 checkpoint pathway. The EMBO Journal *23*, 1133-1143.

Sullivan, E., Santiago, C., Parker, E.D., Dominski, Z., Yang, X., Lanzotti, D.J., Ingledue, T.C., Marzluff, W.F., Duronio, R.J. (2001). *Drosophila* stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. Genes & Development *15*, 173-187.

Xu, H., Kim, U.J., Schuster, T., and Grunstein, M. (1992). Identification of a new set of cell cycle-regulatory genes that regulate S-phase transcription of histone genes in Saccharomyces cerevisiae. Molecular & Cellular Biology *12*, 5249-5259.

Xu, H.X., Johnson, L., Grunstein, M. (1990). Coding and noncoding sequences at the 3' end of yeast histone H2B mRNA confer cell cycle regulation. Molecular & Cellular Biology *10*, 2687-2694.

Ye, X., Wei, Y., Nalepa, G., Harper, J.W. (2003). The cyclin E/Cdk2 substrate p220NPAT is required for S-phase entry, histone gene expression and Cajal body maintenance in human somatic cells. Molecular & Cellular Biology *23*, 8586-8600.

Zhang, R., Chen, W., Adams, P.D. (2007). Molecular dissection of formation of senescence-associated heterochromatin foci. Molecular & Cellular Biology *27*, 2343-2358.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., Pehrson, J.R., Berger, J.M., Kaufman, P.D., Adams, P.D. (2005). Formation of macroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Developmental Cell *8*, 19-30.

Zhao, J., Dynlacht, B., Imai, T., Hori, T.A., Harlow, E. (1998). Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S-phase entry. Genes & Development *12*, 456-461.

Zhao, J., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A., Harlow, E. (2000). NPAT links cyclinE-Cdk2 to the regulation of replication-dependent histone gene transcription. Genes & Development *14*, 2283-2297.

5. General conclusion and future perspectives

The structural organization of DNA into chromatin is of key importance as it regulates a wide variety of cellular functions, such as gene transcription, DNA replication and DNA repair. Maintenance of chromatin organization is thus crucial to preserve genomic stability. During S-phase of each cell cycle, the duplication of chromatin structure occurs in tight coordination with DNA replication. To be efficiently coupled to DNA synthesis, this requires histone synthesis and their deposition onto DNA by chromatin assembly factors (Gunjan, 2005). Importantly, even slight defects in nucleosome assembly can lead to genomic instability and sensitivity to genotoxic agents (Emili, 2001; Game, 1999; Myung et al., 2003). Therefore, a thorough understanding of the molecular mechanisms that control the activity of chromatin assembly factors is crucial to establish molecular links between nucleosome assembly and the maintenance of genomic integrity.

The aim of this thesis was to investigate the regulatory post-translational mechanisms that control the process of chromatin assembly during DNA replication and repair. Our study was mainly focused on using mass spectrometry as a tool to indentify and characterize the phosphorylation of proteins involved in chromatin assembly, namely the proliferating cell nuclear antigen (PCNA), chromatin assembly factor 1 (CAF-1) and the Hir protein complex. To understand the regulation of these chromatin assembly factors and their function in nucleosome assembly, we used the budding yeast *Saccharomyces cerevisiae* as a model organism. Budding yeast is an excellent model because many of the processes that occur in yeast, such as cell cycle control, DNA replication and chromatin dynamics, are similar in humans and are regulated by homologous proteins. In addition, budding yeast is genetically tractable, easy to manipulate, and provides a model system in which hypotheses can be rapidly and rigorously tested (Botstein, 1988; Smutzer, 2001). Finally, budding yeast is also compatible with proteomic techniques, since the affinity purification of epitope-tagged protein complexes and the subsequent identification of their post-translational modifications by mass spectrometry can be performed with relative ease using yeast.

Mass spectrometry has become the standard tool to map protein phosphorylation sites. Nonetheless, their analysis by mass spectrometry is still challenging for several reasons. These include the low stoichiometry of protein phosphorylation, which often requires an enrichment step as a prerequisite to identify sites of phosphorylation. In addition, phosphopeptide ion suppression is a problem frequently encountered during mass spectrometry analyses (Aebersold, 2001; Mann, 2002; Steen, 2006). Furthermore, during an in-depth analysis of phosphorylation sites on PCNA, the eukaryotic sliding clamp, we came across a surprisingly large number of what we originally thought as phosphopeptides. Suspiciously, these peptides were uniquely present in samples derived from silver-stained polyacrylamide gels and completely absent from Coomassie Blue-stained gels. Interestingly, the phosphorylation and sulfonation of proteins both impart a mass shift of 80 Da on serine, threonine and tyrosine residues. However, since the mass difference between these two types of modification is only 0.0095 Da, the distinction between sulfonation and phosphorylation cannot be achieved using low resolution mass spectrometers, such as quadrupole time-of-flight and ion trap instruments. We were only able to differentiate between these two modifications when analyzing yeast PCNA peptides on an LTQ-Orbitrap, an MS instrument which provides higher resolution and better mass accuracy. We eventually found that all the PCNA modifications that were originally ascribed to phosphorylation after silver staining, were instead artifactually sulfated residues. In this context, the origin of the sulfation artifact, which was important to understand prior to conducting our study of chromatin assembly factors, was evaluated in chapter 2. Sodium thiosulfate ($Na_2S_2O_3$) is a commonly used reducing reagent that catalyses the conversion of silver cations into metallic silver deposits that coat proteins during silver staining (Herrmann, 1986). Since this chemical is the only potential source of sulfate used in the procedure, we investigated its effect on the sulfation of serine, threonine and tyrosine residues.

We performed a detailed MS analysis of gel-separated enolase and confirmed that several serine, threonine and tyrosine residues were sulfated using silver staining, but not following Coomassie Blue staining. Although the +80 Da modification can be attributed to either phosphorylation or sulfation, accurate mass measurements of these modified peptides were in agreement with the theoretical values expected for the presence of sulfate (79.9568 Da), rather than phosphate (79.9663 Da). Moreover, a number of peptide ions displaying identical mass values and fragmentation patterns were shown to elute at different chromatographic retention times. An analysis performed by β-elimination revealed that this sulfation can occur at different locations when serine, threonine or tyrosine residues are present within a same peptide. These observations suggested that the silver staining-induced sulfation artifact affects serine, threonine, and tyrosine residues randomly. The abundance of enolase sulfopeptides was measured when silver staining was performed with increasing concentrations of $Na_2S_2O_3$. Interestingly, the degree of sulfation correlated with increasing concentrations of sodium thiosulfate up to 0.02%, which is typically used for silver staining. This unambiguously established the source of reactive sulfate. We hypothesize that silver ions might interact with and activate sodium thiosulfate thus promoting the sulfonation of residues and suggested a reaction mechanism. Furthermore, after examining the distribution of amino acids around the sulfated residues, we found that serine, threonine and tyrosine residues in close proximity to acidic residues have a higher propensity of being artifactually sulfated. Results obtained using the phosphoprotein ERK1 demonstrated that differentiation between the sulfation artifact and real phosphorylation can be relatively straightforward using instruments that provide high mass accuracy. Otherwise, manual validation of the MS/MS spectra, along with the detection of the signature $[M + 2H - 98]^{2+}$ or $[M + 2H - 80]^{2+}$ fragments, is absolutely required to distinguish phosphopeptides from sulfopeptides. Lastly, the sulfation artifact caused by silver staining was shown to globally affect a wide range of bacterial proteins that are not known to naturally bear this modification in the cell.

The data collected in this work highlighted important pitfalls of silver staining in the assignment of protein phosphorylation by mass spectrometry. Most importantly, the artifact appears to be a low yield side reaction (2%) that occurs at levels comparable to some *in vivo* phosphorylation events (1-2%) (Reinders, 2005) and can therefore be easily mis-assigned when using silver staining. Although silver staining is often used for its high sensitivity (Graham, 2005; Shevchenko, 1996), a feature that is often necessary for many proteomics applications, whenever possible, it should be avoided for accurate identification of protein phosphorylation by MS. Previously reported phosphorylation and even sulfation sites derived from silver stained gel bands with no functional validation should be verified before being accepted as *bona fide in vivo* post-translational modifications. Given the profound consequences that silver staining artifacts impose on biological studies aimed at identifying *in vivo* phosphorylation, silver staining is no longer used for protein phosphorylation studies in our laboratory. Instead, Coomassie Blue staining is the preferred method currently employed.

In chapter 3, we initiated a study on CAF-1 phosphorylation. The main objective of the study presented in this chapter was to obtain a more comprehensive understanding of the role of phosphorylation in regulating the function of CAF-1 *in vivo*. Although *in vitro* studies had revealed that CAF-1 phosphorylation and CAF-1-mediated nucleosome assembly were controlled by cyclinA/E-CDK2 and Cdc7-Dbf4 (Gerard, 2006; Hall, 2001; Keller, 2000), two protein kinases required for the initiation of DNA replication, the sites of CAF-1 modification and their physiological significance had never been established. Our present study brings a new perspective in this area. We addressed this important issue by first isolating the CAF-1 complex from budding yeast using the tandem affinity purification strategy. Using mass spectrometry, we unambiguously identified several phosphorylation sites on the Cac1 subunit (orthologue of human p150). More importantly, we found that Ser-94 and Ser-515 as well as Ser-238 and Ser-503 correspond to consensus sequences that are recognized and phosphorylated by cyclin-dependent kinases and Cdc7-Dbf4, respectively. In addition to identifying for the first time CAF-1 *in vivo* phosphorylation

sites, we also provided the first evidence that CAF-1 is potentially regulated by CDKs and DDKs *in vivo*, thereby complementing previous *in vitro* evidences of this regulation (Gerard, 2006; Hall, 2001; Keller, 2000). Secondly, the function of CAF-1 in heterochromatin-mediated telomeric silencing was investigated. We found that, among all of the phosphorylation sites mapped on Cac1, only Ser-503 was crucial for CAF-1 function in telomeric silencing, as manifested by the failure of the Cac1-S503A mutant strain to repress the telomere-integrated *URA3* reporter gene and grow in the presence of 5-FOA. Furthermore, the ability of the Cac1-S503D phosphomimetic mutant to promote telomeric silencing was consistent with the notion that loss of silencing is specifically due to the loss of phosphorylation on Ser-503. However phosphorylation of Ser-503 did not have any effect on cell viability and did not appear to be required for other known CAF-1 dependent functions, including DNA damage resistance and mitotic chromosome segregation. Taken together, these results provide support for a specific role of Cac1 Ser-503 phosphorylation in promoting CAF-1 function in heterochromatin-mediated gene silencing independently of its other functions. Third, we discovered a novel *in vivo* association between the Cdc7-Dbf4 kinase and CAF-1. In affinity purifications, we in fact determined that Cdc7 co-immunoprecipitated with Cac1. Supporting this physical interaction, *in vitro* kinase assays also revealed that Cdc7-Dbf4 phosphorylates Cac1 with high specificity. This is consistent with Cac1 being a physiologically relevant substrate of Cdc7-Dbf4 *in vivo*. Finally, mass spectrometry analysis confirmed that Cdc7-Dbf4 directly phosphorylates Cac1 on serine 503. Thus, one essential and novel conclusion of this study is the implication of the Cdc7/Dbf4 kinase in heterochromatin function, which appears to be mediated through the phosphorylation of CAF-1 on Cac1 serine 503.

A crucial question that needs to be addressed is the underlying molecular mechanism that links CAF-1 phosphorylation and the propagation of epigenetic information in proliferating cells. One attractive possibility is that phosphorylation mediated by Cdc7-Dbf4 functions as a specific mark to target CAF-1 to PCNA during DNA replication of specific chromosomal regions. Although we found that blocking

phosphorylation of Ser-503 does not significantly reduce the association of CAF-1 with chromatin, CAF-1 Ser-503 phosphorylation may be required to enhance the affinity of CAF-1 PIPs to PCNA specifically in sub-telomeric regions. Consistent with this hypothesis, phosphorylation of CAF-1 p150 by Cdc7/Dbf4 has been shown to promote conversion of p150 from its dimeric to its monomeric form (Gerard, 2006). Consistent with this, we also found that yeast Cac1 can also exist as a dimer *in vivo*. It has been proposed that the monomerization of PCNA favours CAF-1 interaction with PCNA and thus promotes replication-dependent nucleosome assembly *in vitro* (Gerard, 2006). This additional level of regulation would ensure that CAF-1 interaction with PCNA would not interfere with other PCNA binding enzymes and PCNA-dependent processes that occur immediately behind replication forks. An alternative possibility is that phosphorylation of Cac1 Ser-503 by Cdc7-Dbf4 may regulate the association of CAF-1 with proteins that promote heterochromatin formation and silencing, such as Sir proteins and Rtt106, at heterochromatic regions. In support of this notion, CAF-1 physically interacts with Rtt106 and Sir proteins *in vivo* and a reduced association of Sir proteins with telomeric heterochromatin has been observed in cells lacking the histone chaperones CAF-1 and Rtt106 (Huang, 2005, 2007; Sharp, 2003). To date, the true function of CAF-1 phosphorylation in nucleosome assembly and heterochromatin function remains a mystery and should therefore be a primary focus in future research.

Finally, in chapter 4, we were principally interested in determining if the Hir protein complex which, in addition to being a chromatin assembly factor also contributes to histone gene regulation during normal cell cycle progression and the DNA damage response (Kaufman, 1998; Osley, 1986; Osley, 1987; Sharp, 2001), is phosphorylated in yeast. Hir proteins transcriptionally repress histone genes in G1, G2 and in response to DNA damage that slows down replication during S-phase. On the other hand, the ability of Hir proteins to repress the transcription of histone genes is suppressed during DNA replication in S-phase because cells need an abundant supply of histones for the assembly of nucleosomes on newly replicated DNA (Gunjan, 2005). In order to obtain a better

understanding of how Hir-dependent histone gene repression is regulated, we characterized the phosphorylation of Hir proteins by mass spectrometry. We showed for the first time that the Hpc2 subunit is highly phosphorylated on several sites, including DDK and CDK-like sites. These two kinases are known to regulate a number of cell cycle events, such as DNA synthesis, chromatin assembly, mitosis and other biosynthetic processes (Morgan, 1997; Sclafani, 2000). Thus, our initial results also suggest that the S-phase specific expression of histone genes, another important event of the cell cycle, is likely modulated through cell-cycle regulated phosphorylation of the Hir proteins by DDK and/or CDKs. Whether Hpc2 phosphorylation occur during S-phase to inactivate the ability of Hir proteins to repress histone genes or outside of S-phase to promote histone gene repression remains to be determined. Nonetheless, in response to MMS-induced DNA damage, we showed that Hpc2 is specifically phosphorylated on Ser-330 and that this phosphorylation was dependent upon DNA damage response kinases. Given that DNA damage and activation of checkpoint kinases trigger histone gene repression (Paik, In preparation), these results suggest that MMS-induced phosphorylation of Hpc2 is required to activate the ability of Hir proteins to promote transcriptional repression of histone genes. This response is important to prevent the accumulation of excess histones when DNA replication is slowed down by genotoxic agents.

Our results are consistent with the hypothesis that Hpc2 is likely an important regulatory subunit whose phosphorylation mediates the activation of the Hir protein complex when histone gene repression is required. Indeed, during the course of writing this thesis, studies recently performed in our laboratory have suggested that, although the stability of Hir proteins (Hir1 and Hir2) does not seem to change throughout the cell cycle (Blackwell, 2004; De Lucia, 2001; Sherwood, 1993), Hpc2 is highly cell cycle regulated. Protein stability assays have in fact demonstrated that Hpc2 is stable in G1, when histone genes are repressed, but highly unstable at the Cdc7 arrest point, when histone gene transcription is activated prior to initiation of DNA replication. Moreover, Hpc2 is re-stabilized when histone genes are repressed in response to DNA damage. We therefore

suggest that MMS-induced phosphorylation might contribute to increase the stability of Hpc2 and regulate the activity of the Hir protein complex. The active complex may subsequently recruit RSC to histone gene promoters to alter nucleosome positioning or switch on the nucleosome assembly activity of the Hir complex to repress the transcription of histone genes. It will therefore be interesting to conduct further mutational analysis of Hpc2 phosphorylation sites and investigate the phenotypic consequences of blocking Hpc2 phosphorylation on histone gene expression, Hpc2 stability, cell cycle progression and RSC recruitment to histone gene loci.

In summary, the mapping of post-translational modifications by mass spectrometry has yielded novel insights into the control of histone synthesis and the regulation of chromatin assembly by phosphorylation. However, there is still much to understand and thus, long-term follow-up studies would be of major interest to determine the molecular relationship between CAF-1 and Hpc2 phosphorylation and chromatin assembly events that require a tight coordination with histone synthesis. CAF-1 and Hir proteins function at the crossroads of many biological processes that are essential for genomic stability and are therefore linked to major diseases like cancer. For this reason, understanding the molecular switches that dictate the functions of these proteins is of great importance.

5.1. References

Aebersold, R., Goodlett, D.R. (2001). Mass Spectrometry in Proteomics. Chemical Reviews *101*, 269-296.

Blackwell, C., Martin, K.A., Greenall, A., Pidoux, A., Allshire, R.C., Whitehall, S.K. (2004). The *Schizosaccharomyces pombe* HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. Mol Cell Biol *24*, 4309-4320.

Botstein, D., Fink, G.R. (1988). Yeast: an experimental organism for modern biology. Science *240*, 1439-1443.

De Lucia, F., Lorain, S., Scamps, C., Galissson, F., MacHold, J., Lipinski, M. (2001). Subnuclear localization and mitotic phosphorylation of HIRA, the human homologue of Saccharomyces cerevisiae transcriptional regulators Hir1p/Hir2p. Biochem J *358 (Pt 2)*, 447-455.

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

Game, J.C., Kaufman, P.D. (1999). Role of saccharomyces cerevisiae chromatin assembly factor-I in repair of ultraviolet radiation damage in vivo. Genetics *151*, 485-497.

Gerard, A., Koundrioukoff, S., Ramillon, V., Sergere, J-C., Mailand, N., Quivy, J-P., Almouzni, G. (2006). The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. EMBO Rep *7*, 817-823.

Graham, D.R.M., Elliott, S.T., van Eyk, J.E. (2005). Broad-based proteomic strategies: a practical guide to proteomics and functional screening. The Journal of Physiology *563*, 1-9.

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

Hall, C., Nelson, D.M., Ye, X., Baker, K., DeCaprio, J.A., Seeholzer, S., Lipinski, M., Adams, P.D. (2001). HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel

cyclin-cdk2 substrate whose expression blocks S-phase progression. Mol Cell Biol *21*, 1854-1865.

Herrmann, W.A. (1986). Multiple Bonds between Transition Metals and ldquoBarerdquo Main Group Elements: Links between Inorganic Solid State Chemistry and Organometallic Chemistry. Angewandte Chemie International Edition in English *25*, 56-76.

Huang, S., Zhou, H., Katzmann, D., Hochstrasser, M., Atanasova, E., Zhang, Z. (2005). Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. Proceedings of the National Academy of Sciences of the United States of America *102*, 13410-13415.

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

Kaufman, P.D., Cohen, J.L., Osley, M.A. (1998). Hir Proteins Are Required for Position-Dependent Gene Silencing in Saccharomyces cerevisiae in the Absence of Chromatin Assembly Factor I. Mol Cell Biol *18*, 4793-4806.

Keller, C., Krude, T. (2000). Requirement of Cyclin/Cdk2 and Protein Phosphatase 1 Activity for Chromatin Assembly Factor 1-dependent Chromatin Assembly during DNA Synthesis. Journal of Biological Chemistry *275*, 35512-35521.

Mann, M., Ong, S-E., Grønborg, M., Steen, H., Jensen, O. N., Pandey, A. (2002). Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends in Biotechnology *20*, 261-268.

Morgan, D.O. (1997). CYCLIN-DEPENDENT KINASES: Engines, Clocks, and Microprocessors. Annual Review of Cell and Developmental Biology *13*, 261-291.

Myung, K., Pennaneach, V., Kats, E.S., and Kolodner, R.D. (2003). Saccharomyces cerevisiae chromatin-assembly factors that act during DNA replication function in the maintenance of genome stability. Proceedings of the National Academy of Sciences of the United States of America *100*, 6640-6645.

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

Osley, M.A., Lycan, D. (1987). Trans-acting regulatory mutations that alter transcription of Saccharomyces cerevisiae histone genes. Mol Cell Biol *7*, 4204-4210.

Paik, J., Gunjan, A., Verreault, A. (In preparation). The Rad53 checkpoint kinase represses histone gene transcription in response to DNA damage and replication arrest.

Reinders, J., Sickmann, A. (2005). State-of-the-art in phosphoproteomics. Proteomics *5*, 4052-4061.

Sclafani, R.A. (2000). Cdc7p-Dbf4p becomes famous in the cell cycle. J Cell Sci *113*, 2111-2117.

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

Sharp, J.A., Krawitz, D.C., Gardner, K.A., Fox, C.A., Kaufman, P.D. (2003). The budding yeast silencing protein Sir1 is a functional component of centromeric chromatin. Genes & Development *17*, 2356-2361.

Sherwood, P.W., Tsang, S.V., Osley, M.A. (1993). Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in Saccharomyces cerevisiae. Mol Cell Biol *13*, 28-38.

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

Smutzer, G. (2001). Yeast: an attractive, yet simple model. The Scientist *15*, 24-25.

Steen, H., Jebanathirajah, J.A., Rush, J., Morrice, N., Kirschner, M.W. (2006). Phosphorylation Analysis by Mass Spectrometry. Molecular & Cellular Proteomics *5*, 172-181.

APPENDIX A

Supplementary information for chapter 2

Table A1. Proteins identified from silver and Commassie stained polyacrylamide gel

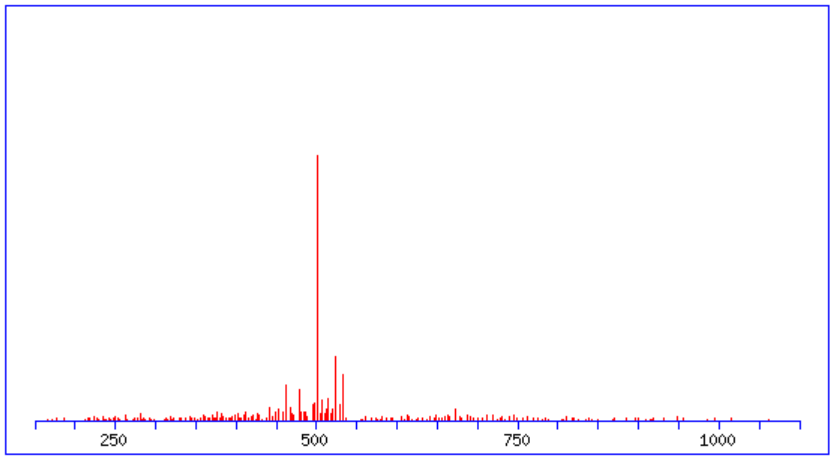
Protein ID	IPI entry #	Database source	Protein	Condition	Protein score	Total # of peptides	% sequence coverage
1	gi/ 119336	NCBI	Enolase	Silver	4295	37	83
				Coommassie	3591	35	81
2	Q63844*	Swiss-Prot	ERK-1	Silver	6963	23	69
				Coommassie	3421	26	59
3	gi/ 42958331	NCBI	Elongation Factor Tu	Silver	6468	31	91
				Coommassie	4859	32	92
4	gi/ 24050373	NCBI	Chain elongation factor EF-Ts	Silver	2101	26	82
				Coommassie	2201	26	82
5	gi/ 24053793	NCBI	50S ribosome subunit	Silver	1498	17	69
				Coommassie	1372	17	77

* This sequence belongs to the rat p44^{mapk} (ERK1) and shares 98.6% homology with the hamster ERK1. The exact sequence of hamster ERK1 used in this study can be found in the following publication: Meloche S., Pagès, G. and Pouységgur J. (1991). Functional expression and growth factor activation of an epitope-tagged p44 mitogen-activated protein kinase, p44^{mapk}. *Mol. Biol. of the Cell*, 3, 63-71.

Figure A1: MS^2 and MS^3 spectra from modified peptides of yeast enolase

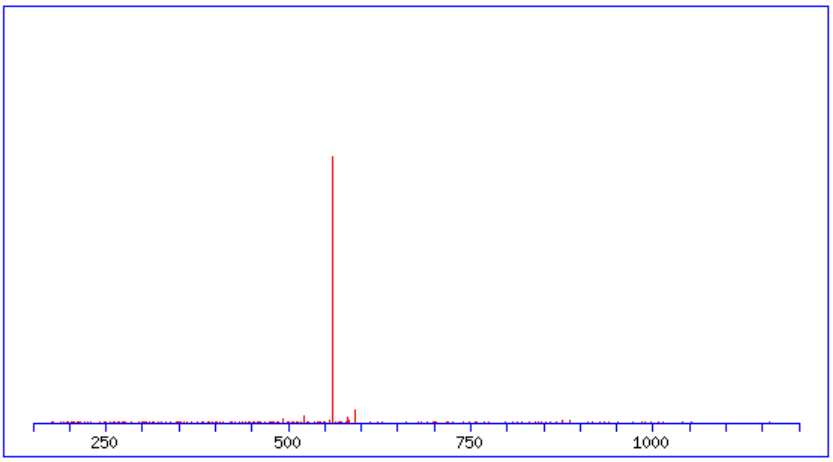
ERK-1 sulfopeptides
Peptide #1:
<u>APEIMLNSK + sulfonation</u>
***m/z*: 541.75**
Charge: 2+
RT: 28.2 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 514.82



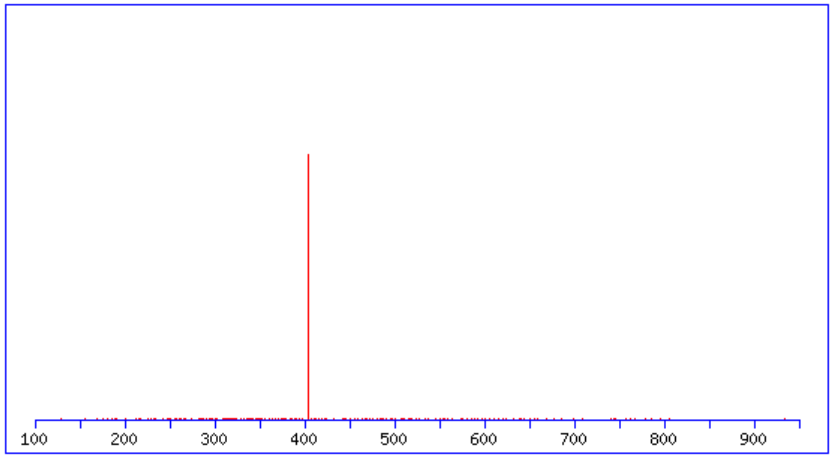
Peptide #2:
<u>MLTFNPNKR+ sulfonation</u>
***m/z*: 600.78**
Charge: 2+
RT: 31.6 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 560.83

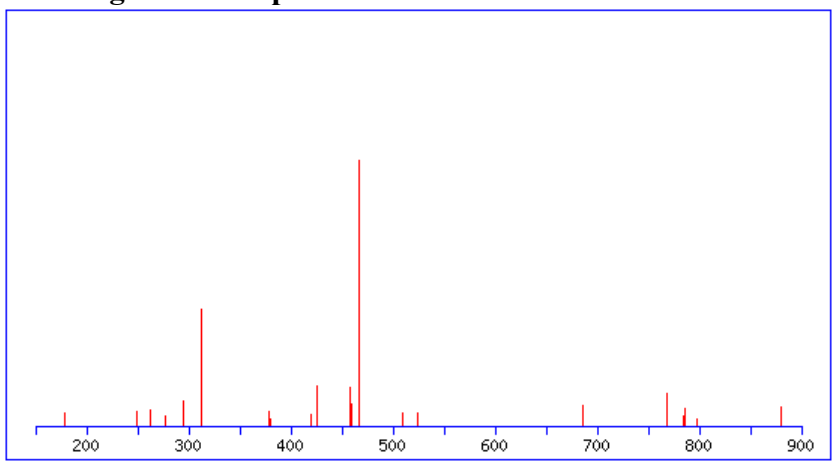


Peptide #3:
<u>YIHSANVLHR + sulfonation</u>
***m/z*: 430.55**
Charge: 3+
RT: 23.8 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 403.13



MS^3 fragmentation pattern



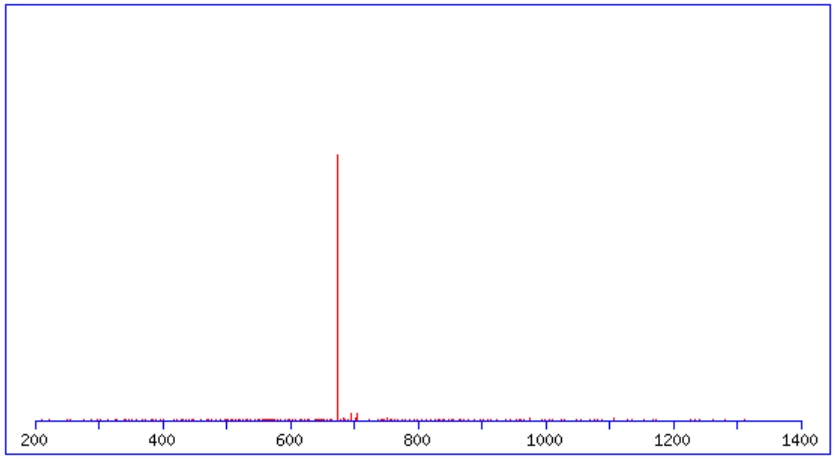
Variable modifications:
Y1 or S4: Sulfonation
Matched ions:

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	146.06	73.53					Y							10
2	259.14	130.08					I	1046.59	523.80	1029.56	515.28	1028.57	514.79	9
3	396.20	198.61					H	933.50	467.25	916.47	458.74	915.49	458.25	8
4	483.24	242.12			465.22	233.12	S	**796.44**	398.72	779.42	390.21	778.43	389.72	7
5	554.27	277.64			536.26	268.63	A	709.41	355.21	692.38	346.70			6
6	668.32	334.66	651.29	326.15	650.30	325.66	N	638.37	319.69	621.35	311.18			5
7	**767.38**	384.20	750.36	375.68	749.37	375.19	V	**524.33**	262.67	507.30	254.16			4
8	**880.47**	440.74	863.44	432.22	862.46	431.73	L	**425.26**	213.13	408.24	204.62			3
9	1017.53	509.27	1000.50	500.75	999.52	500.26	H	**312.18**	156.59	295.15	148.08			2
10							R	175.12	88.06	158.09	79.55			1

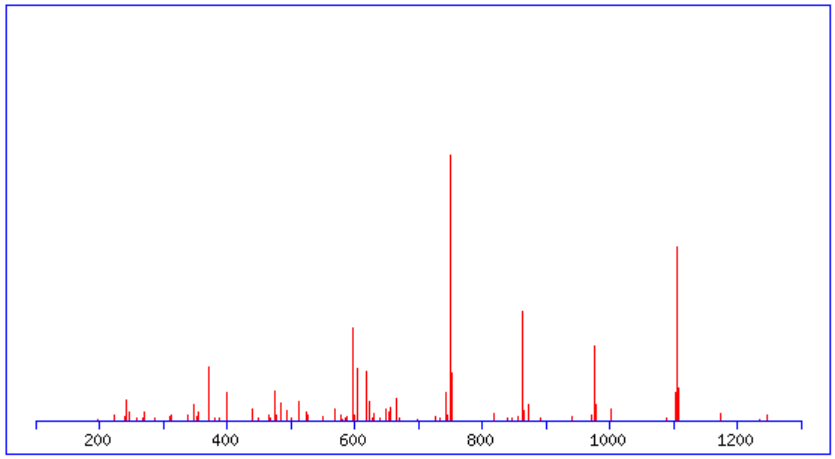
Peptide #4:

<u>LKELIFQETAR+ sulfonation</u>
***m/z*: 714.37**
Charge: 2+
RT: 35.2 min

MS^2 fragmentation spectrum
Neutral loss : 80 Da at *m/z* = 674.47



MS^3 fragmentation pattern

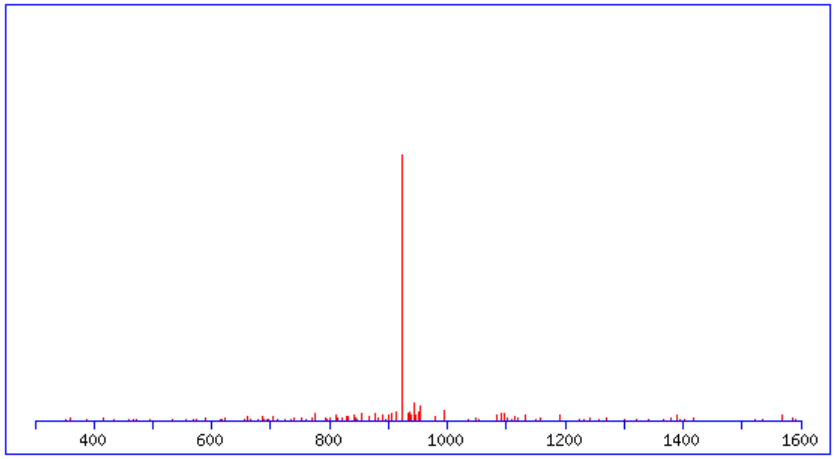


Variable modifications:
T9: Sulfonation
Matched ions:

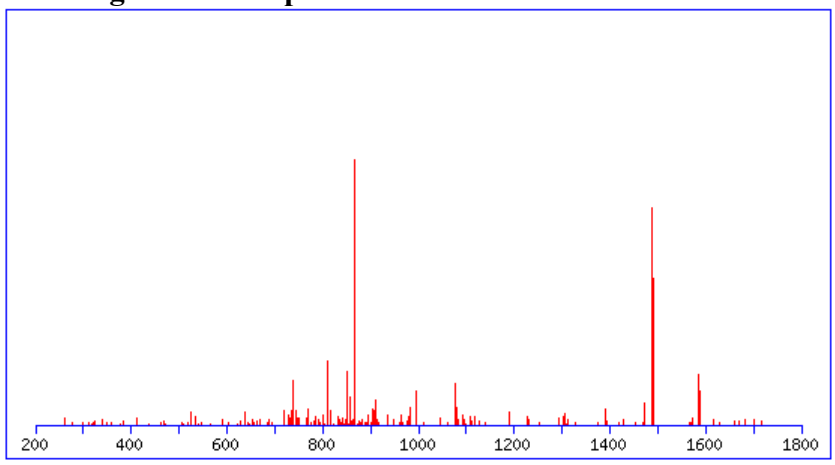
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	114.09	57.55					L							11
2	242.19	121.60	225.16	113.08			K	1234.68	617.84	1217.65	609.33	1216.67	608.84	10
3	371.23	186.12	354.20	177.60	353.22	177.11	E	1106.58	553.80	1089.56	545.28	1088.57	544.79	9
4	484.31	242.66	467.29	234.15	466.30	233.65	L	977.54	489.27	960.51	480.76	959.53	480.27	8
5	597.40	299.20	580.37	290.69	579.39	290.20	I	864.46	432.73	847.43	424.22	846.45	423.73	7
6	744.47	372.74	727.44	364.22	726.45	363.73	F	751.37	376.19	734.35	367.68	733.36	367.19	6
7	872.52	436.77	855.50	428.25	854.51	427.76	Q	604.30	302.66	587.28	294.14	586.29	293.65	5
8	1001.57	501.29	984.54	492.77	983.56	492.28	E	476.25	238.63	459.22	230.11	458.24	229.62	4
9	1102.61	551.81	1085.59	543.30	1084.60	542.81	T	347.20	174.11	330.18	165.59	329.19	165.10	3
10	1173.65	587.33	1156.62	578.82	1155.64	578.32	A	246.16	123.58	229.13	115.07			2
11							R	175.12	88.06	158.09	79.55			1

Peptide #5:
DLKPSNLLINTTC(carb.)DLK + sulfonation
***m/z*: 962.99**
Charge: 2+
RT: 42.6 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 923.32



MS^3 fragmentation spectrum



Variable modifications:
S5 or T11 or T12: Sulfonation
C13: Carbamidomethyl
Matched ions:

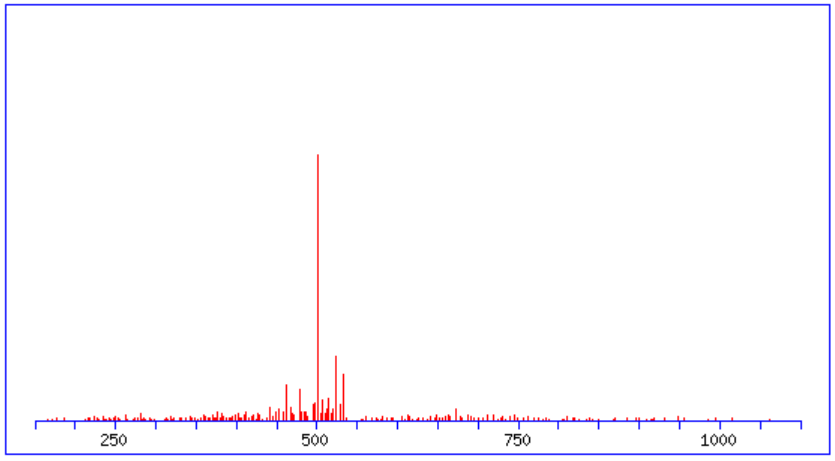
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	116.03	58.52			98.02	49.52	D							16
2	229.12	115.06			211.11	106.06	L	1729.95	865.48	1712.93	856.97	1711.94	856.47	15
3	357.21	179.11	340.19	170.60	339.20	170.10	K	1616.87	808.94	1599.84	800.42	1598.86	799.93	14
4	454.27	227.64	437.24	219.12	436.26	218.63	P	1488.77	744.89	1471.75	736.38	1470.76	735.88	13
5	541.30	271.15	524.27	262.64	523.29	262.15	S	1391.72	696.36	1374.69	687.85	1373.71	687.36	12
6	655.34	328.17	638.31	319.66	637.33	319.17	N	1304.69	652.85	1287.66	644.33	1286.68	643.84	11
7	768.43	384.72	751.40	376.20	750.41	375.71	L	1190.64	595.83	1173.62	587.31	1172.63	586.82	10
8	881.51	441.26	864.48	432.74	863.50	432.25	L	1077.56	539.28	1060.53	530.77	1059.55	530.28	9
9	994.59	497.80	977.57	489.29	976.58	488.79	I	964.48	482.74	947.45	474.23	946.47	473.74	8
10	1108.64	554.82	1091.61	546.31	1090.63	545.82	N	851.39	426.20	834.37	417.69	833.38	417.19	7
11	1209.68	605.35	1192.66	596.83	1191.67	596.34	T	737.35	369.18	720.32	360.67	719.34	360.17	6
12	1310.73	655.87	1293.70	647.36	1292.72	646.86	T	636.30	318.65	619.28	310.14	618.29	309.65	5
13	1470.76	735.88	1453.74	727.37	1452.75	726.88	C	535.25	268.13	518.23	259.62	517.24	259.13	4
14	1585.79	793.40	1568.76	784.88	1567.78	784.39	D	375.22	188.12	358.20	179.60	357.21	179.11	3
15	1698.87	849.94	1681.85	841.43	1680.86	840.93	L	260.20	130.60	243.17	122.09			2
16							K	147.11	74.06	130.09	65.55			1

Figure A2 : MS^2 and MS^3 spectra from modified peptides of ERK1

ERK-1 sulfopeptides

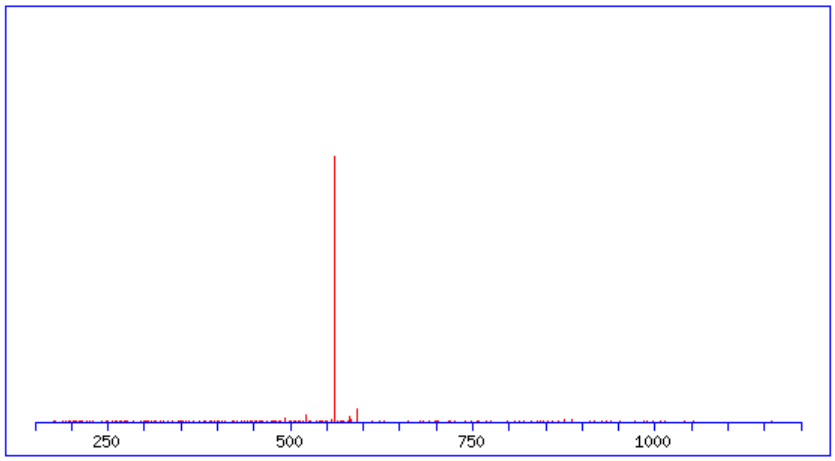
Peptide #1:
<u>APEIMLNSK + sulfonation</u>
***m/z*: 541.75**
Charge: 2+
RT: 28.2 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 514.82



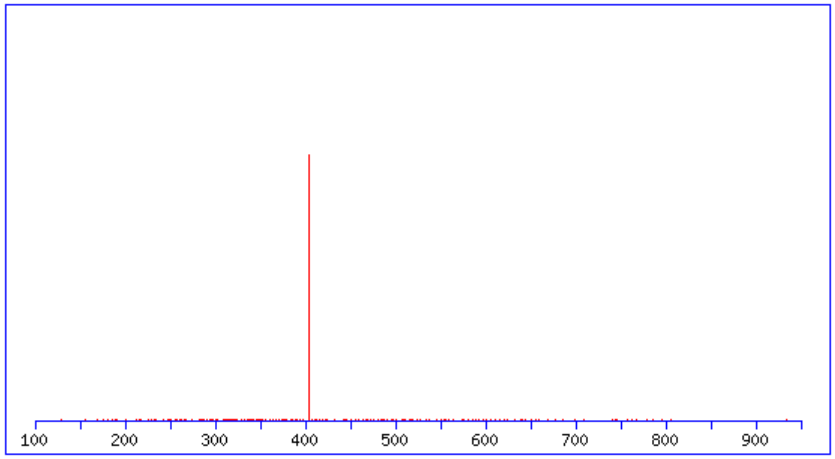
Peptide #2:
<u>MLTFNPNKR+ sulfonation</u>
***m/z*: 600.78**
Charge: 2+
RT: 31.6 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 560.83

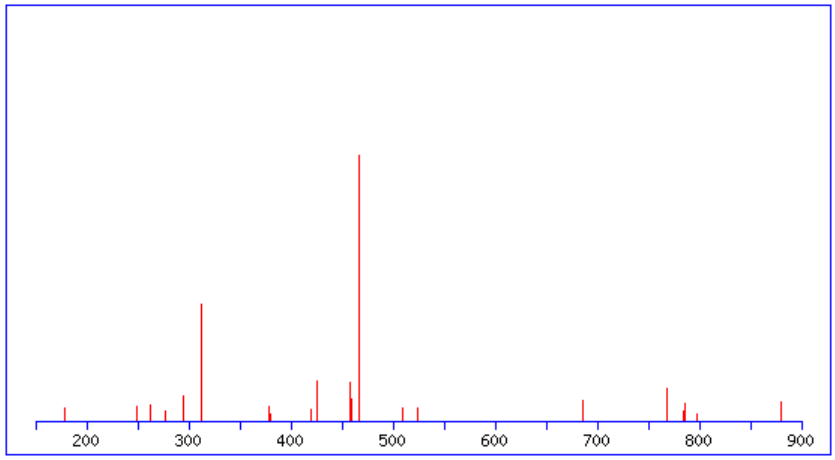


Peptide #3:
<u>YIHSANVLHR + sulfonation</u>
***m/z*: 430.55**
Charge: 3+
RT: 23.8 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 403.13



MS^3 fragmentation pattern



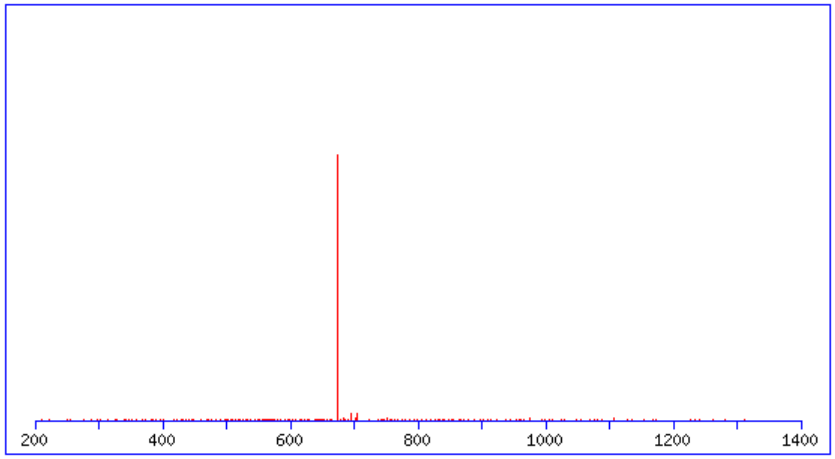
Variable modifications:
Y1 or S4: Sulfonation
Matched ions:

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	146.06	73.53					Y							10
2	259.14	130.08					I	1046.59	523.80	1029.56	515.28	1028.57	514.79	9
3	396.20	198.61					H	933.50	467.25	916.47	458.74	915.49	458.25	8
4	483.24	242.12			465.22	233.12	S	**796.44**	398.72	779.42	390.21	778.43	389.72	7
5	554.27	277.64			536.26	268.63	A	709.41	355.21	692.38	346.70			6
6	668.32	334.66	651.29	326.15	650.30	325.66	N	638.37	319.69	621.35	311.18			5
7	**767.38**	384.20	750.36	375.68	749.37	375.19	V	**524.33**	262.67	507.30	254.16			4
8	**880.47**	440.74	863.44	432.22	862.46	431.73	L	**425.26**	213.13	408.24	204.62			3
9	1017.53	509.27	1000.50	500.75	999.52	500.26	H	**312.18**	156.59	295.15	148.08			2
10							R	175.12	88.06	158.09	79.55			1

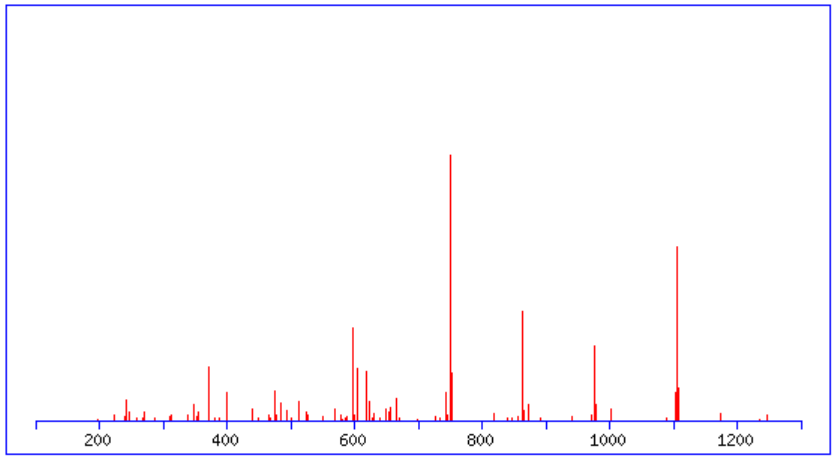
Peptide #4:

<u>LKELIFQETAR+ sulfonation</u>
***m/z*: 714.37**
Charge: 2+
RT: 35.2 min

MS^2 fragmentation spectrum
Neutral loss : 80 Da at *m/z* = 674.47



MS^3 fragmentation pattern

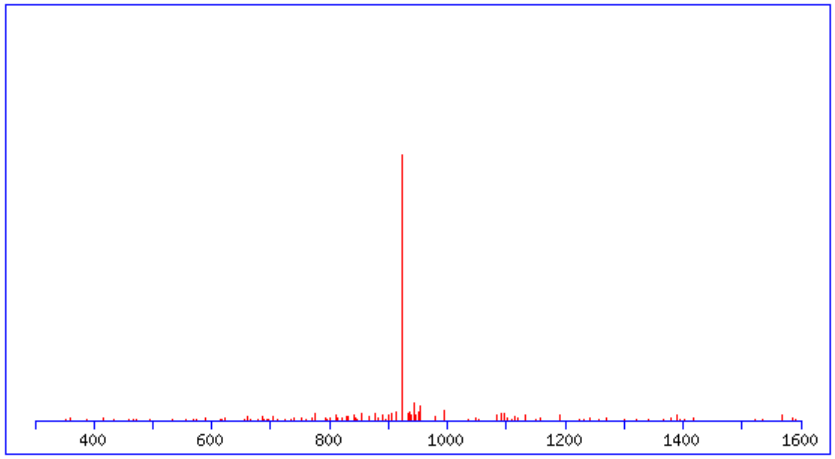


Variable modifications:
T9: Sulfonation
Matched ions:

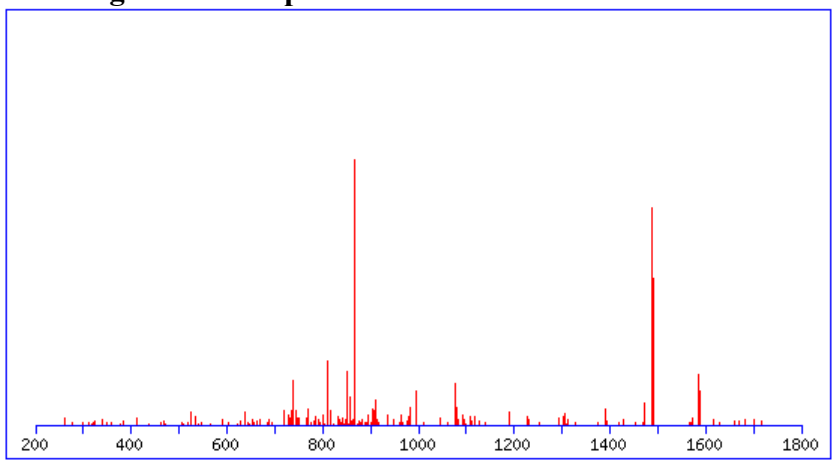
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	114.09	57.55					L							11
2	242.19	121.60	225.16	113.08			K	1234.68	617.84	1217.65	609.33	1216.67	608.84	10
3	371.23	186.12	354.20	177.60	353.22	177.11	E	1106.58	553.80	1089.56	545.28	1088.57	544.79	9
4	484.31	242.66	467.29	234.15	466.30	233.65	L	977.54	489.27	960.51	480.76	959.53	480.27	8
5	597.40	299.20	580.37	290.69	579.39	290.20	I	864.46	432.73	847.43	424.22	846.45	423.73	7
6	744.47	372.74	727.44	364.22	726.45	363.73	F	751.37	376.19	734.35	367.68	733.36	367.19	6
7	872.52	436.77	855.50	428.25	854.51	427.76	Q	604.30	302.66	587.28	294.14	586.29	293.65	5
8	1001.57	501.29	984.54	492.77	983.56	492.28	E	476.25	238.63	459.22	230.11	458.24	229.62	4
9	1102.61	551.81	1085.59	543.30	1084.60	542.81	T	347.20	174.11	330.18	165.59	329.19	165.10	3
10	1173.65	587.33	1156.62	578.82	1155.64	578.32	A	246.16	123.58	229.13	115.07			2
11							R	175.12	88.06	158.09	79.55			1

Peptide #5:
DLKPSNLLINTTC(carb.)DLK + sulfonation
***m/z*: 962.99**
Charge: 2+
RT: 42.6 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 923.32



MS^3 fragmentation spectrum



Variable modifications:
S5 or T11 or T12: Sulfonation
C13: Carbamidomethyl
Matched ions:

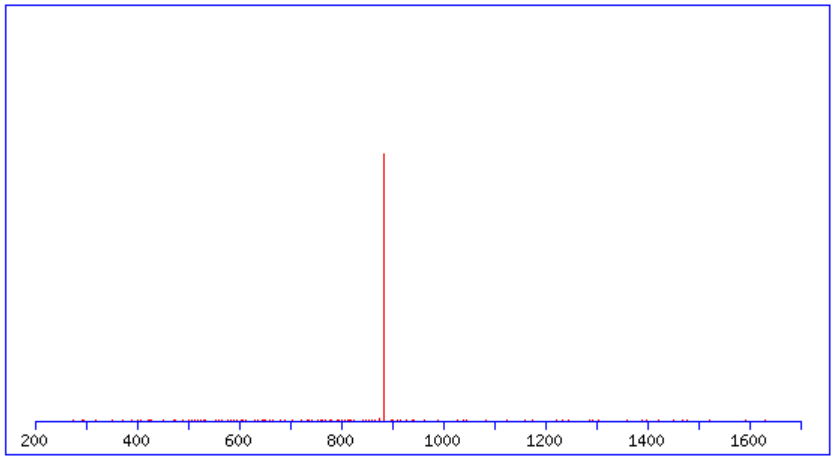
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	116.03	58.52			98.02	49.52	D							16
2	229.12	115.06			211.11	106.06	L	1729.95	865.48	1712.93	856.97	1711.94	856.47	15
3	357.21	179.11	340.19	170.60	339.20	170.10	K	1616.87	808.94	1599.84	800.42	1598.86	799.93	14
4	454.27	227.64	437.24	219.12	436.26	218.63	P	1488.77	744.89	1471.75	736.38	1470.76	735.88	13
5	541.30	271.15	524.27	262.64	523.29	262.15	S	1391.72	696.36	1374.69	687.85	1373.71	687.36	12
6	655.34	328.17	638.31	319.66	637.33	319.17	N	1304.69	652.85	1287.66	644.33	1286.68	643.84	11
7	768.43	384.72	751.40	376.20	750.41	375.71	L	1190.64	595.83	1173.62	587.31	1172.63	586.82	10
8	881.51	441.26	864.48	432.74	863.50	432.25	L	1077.56	539.28	1060.53	530.77	1059.55	530.28	9
9	994.59	497.80	977.57	489.29	976.58	488.79	I	964.48	482.74	947.45	474.23	946.47	473.74	8
10	1108.64	554.82	1091.61	546.31	1090.63	545.82	N	851.39	426.20	834.37	417.69	833.38	417.19	7
11	1209.68	605.35	1192.66	596.83	1191.67	596.34	T	737.35	369.18	720.32	360.67	719.34	360.17	6
12	1310.73	655.87	1293.70	647.36	1292.72	646.86	T	636.30	318.65	619.28	310.14	618.29	309.65	5
13	1470.76	735.88	1453.74	727.37	1452.75	726.88	C	535.25	268.13	518.23	259.62	517.24	259.13	4
14	1585.79	793.40	1568.76	784.88	1567.78	784.39	D	375.22	188.12	358.20	179.60	357.21	179.11	3
15	1698.87	849.94	1681.85	841.43	1680.86	840.93	L	260.20	130.60	243.17	122.09			2
16							K	147.11	74.06	130.09	65.55			1

Figure A3 : MS^2 and MS^3 spectra from modified peptides from *E. coli* protein extracts

EF-Tu sulfopeptides (from bacterial proteins)

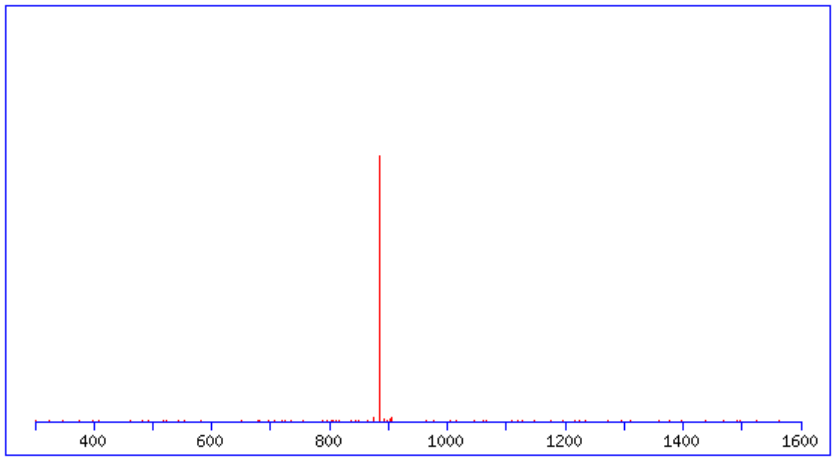
Peptide #1:
HYAHVDC(carb)PGHADYVK + sulfonation
***m/z* : 924.88**
Charge: 2+
RT: 21.1 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 885.06



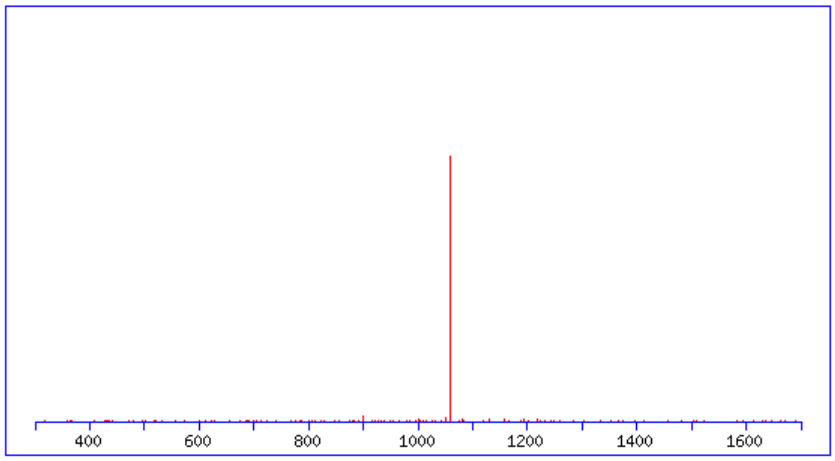
Peptide #2:
HYAHVDCPGHADYVK + sulfonation
***m/z*: 924.88**
Charge: 2+
RT: 21.7 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 884.95



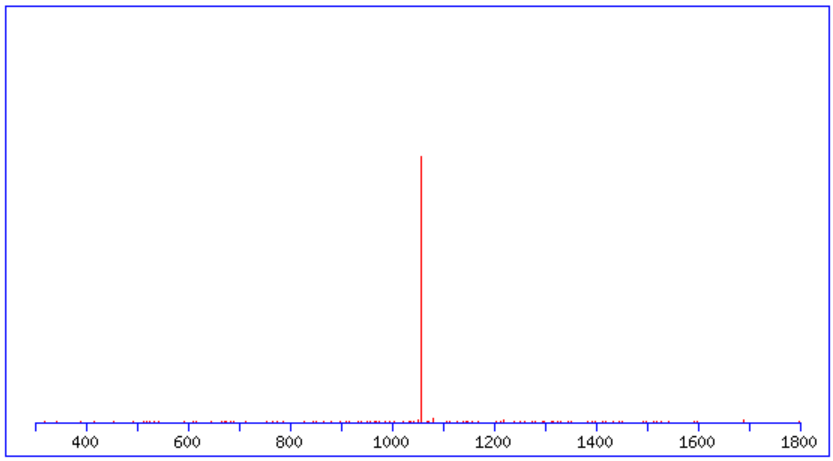
Peptide #3:
AIDKPFLLPIEDVFSISGR + sulfonation
***m/z*: 1099.07**
Charge: 2+
RT: 56.4 min

MS2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 1059.38



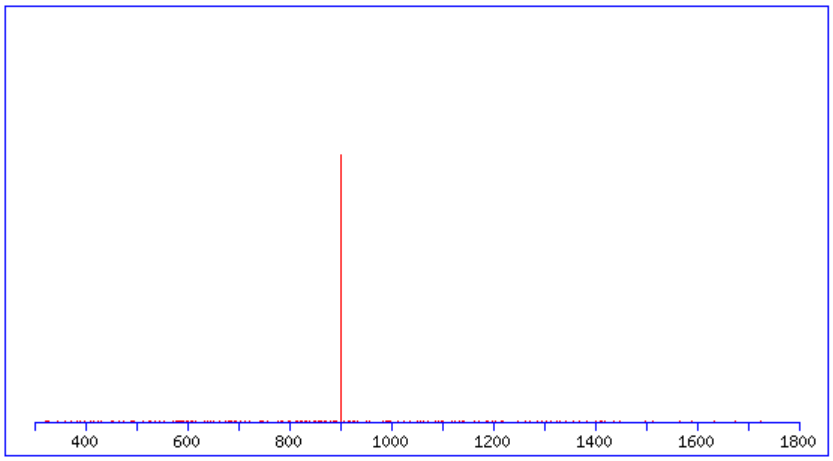
Peptide #4:
AIDKPFLLPIEDVFSISGR + sulfonation
***m/z*: 1099.07**
Charge: 2+
RT: 57.4 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 1059.03

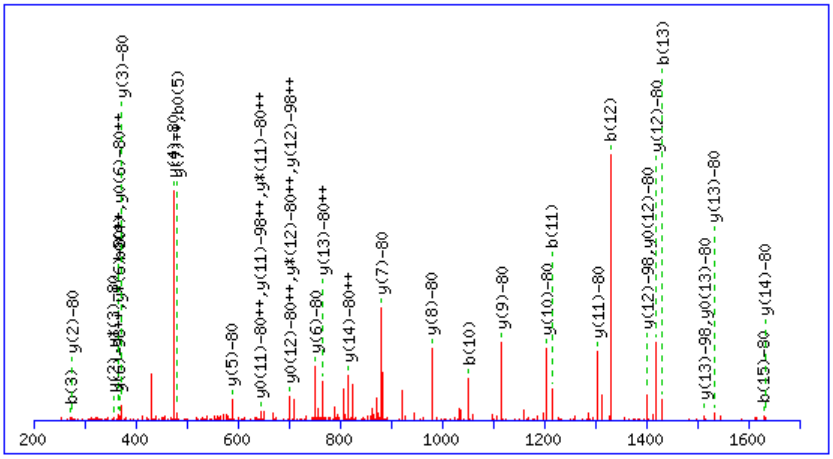


Peptide #5:
GITINTSHVEYDTPTR + sulfonation
***m/z* : 942.43**
Charge: 2+
RT: 31.8 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 902.40



MS^3 fragmentation spectrum

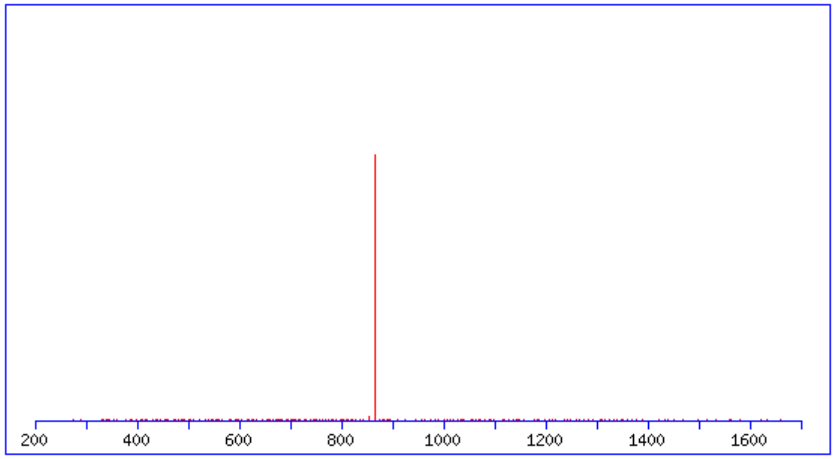


Variable modifications:
T2 or T4 or Y6 or S10 or T14: Sulfonation
Matched ions:

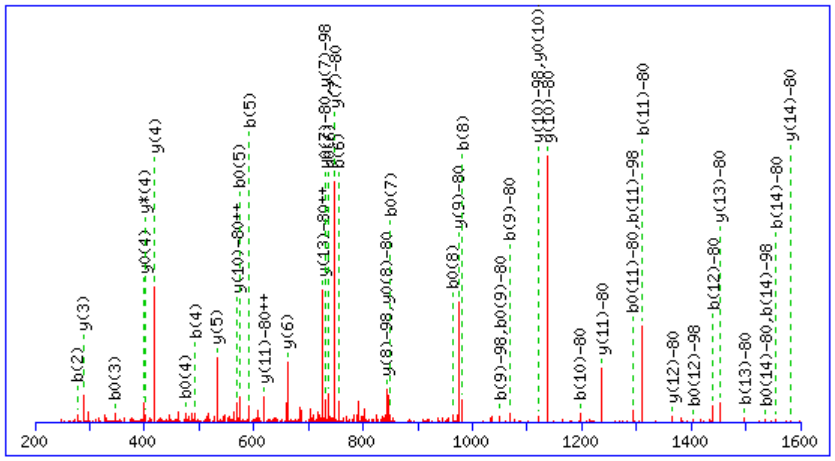
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	58.03	29.52					G							16
2	171.11	86.06					I	1746.87	873.94	1729.84	865.42	1728.86	864.93	15
3	**272.16**	136.58			254.15	127.58	T	**1633.78**	**817.39**	1616.76	808.88	1615.77	808.39	14
4	385.24	193.13			**367.23**	184.12	I	**1532.73**	**766.87**	1515.71	758.36	**1514.72**	757.87	13
5	499.29	250.15	482.26	241.63	**481.28**	241.14	N	**1419.65**	710.33	1402.62	**701.82**	**1401.64**	**701.32**	12
6	600.34	300.67	583.31	292.16	582.32	291.67	T	**1305.61**	653.31	1288.58	**644.79**	1287.60	**644.30**	11
7	687.37	344.19	670.34	335.67	669.36	335.18	S	**1204.56**	602.78	1187.53	594.27	1186.55	593.78	10
8	824.43	412.72	807.40	404.20	806.42	403.71	H	**1117.53**	559.27	1100.50	550.75	1099.52	550.26	9
9	923.49	462.25	906.47	453.74	905.48	453.25	V	**980.47**	490.74	963.44	482.22	962.46	481.73	8
10	**1052.54**	526.77	1035.51	518.26	1034.53	517.77	E	**881.40**	441.20	864.37	432.69	863.39	432.20	7
11	**1215.60**	608.30	1198.57	599.79	1197.59	599.30	Y	**752.36**	376.68	735.33	**368.17**	734.35	**367.68**	6
12	**1330.63**	665.82	1313.60	657.30	1312.62	656.81	D	**589.29**	295.15	572.27	286.64	571.28	286.15	5
13	**1431.68**	716.34	1414.65	707.83	1413.66	707.34	T	**474.27**	237.64	457.24	229.12	456.26	228.63	4
14	1528.73	764.87	1511.70	756.35	1510.72	755.86	P	**373.22**	187.11	**356.19**	178.60	355.21	178.11	3
15	**1629.78**	815.39	1612.75	806.88	1611.76	806.39	T	**276.17**	138.59	259.14	130.07	258.16	129.58	2
16							R	175.12	88.06	158.09	79.55			1

Peptide #6:
FESEVYILSKDEGGR + sulfonation
***m/z* : 904.91**
Charge: 2+
RT: 35.8 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 864.89



MS^3 fragmentation spectrum

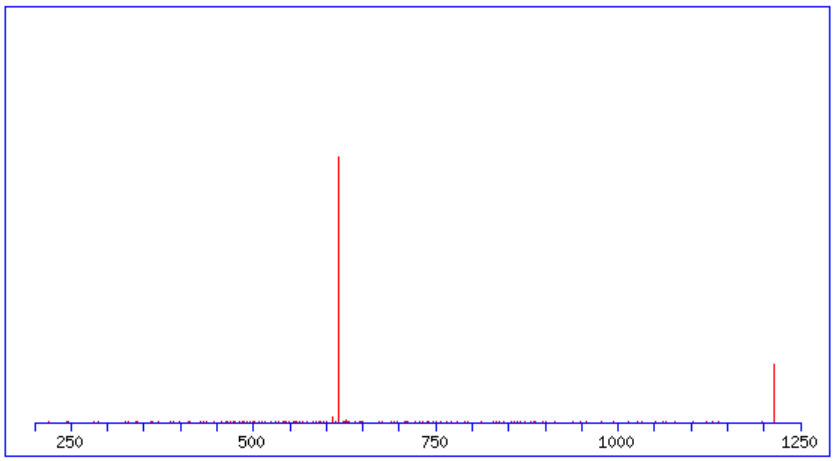


Variable modifications:
S3 or Y6 or S9: Sulfonation
Matched ions:

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	148.08	74.54					F							15
2	277.12	139.06			259.11	130.06	E	1581.78	791.39	1564.75	782.88	1563.76	782.39	14
3	364.15	182.58			346.14	173.57	S	1452.73	726.87	1435.71	718.36	1434.72	717.86	13
4	493.19	247.10			475.18	238.09	E	1365.70	683.35	1348.67	674.84	1347.69	674.35	12
5	592.26	296.63			574.25	287.63	V	1236.66	618.83	1219.63	610.32	1218.65	609.83	11
6	755.32	378.17			737.31	369.16	Y	1137.59	569.30	1120.56	560.79	1119.58	560.29	10
7	868.41	434.71			850.40	425.70	I	974.53	487.77	957.50	479.25	956.52	478.76	9
8	981.49	491.25			963.48	482.24	L	861.44	431.22	844.42	422.71	843.43	422.22	8
9	1068.52	534.77			1050.51	525.76	S	748.36	374.68	731.33	366.17	730.35	365.68	7
10	1196.62	598.81	1179.59	590.30	1178.61	589.81	K	661.33	331.17	644.30	322.65	643.32	322.16	6
11	1311.65	656.33	1294.62	647.81	1293.64	647.32	D	533.23	267.12	516.20	258.61	515.22	258.11	5
12	1440.69	720.85	1423.66	712.34	1422.68	711.84	E	418.20	209.61	401.18	201.09	400.19	200.60	4
13	1497.71	749.36	1480.68	740.85	1479.70	740.35	G	289.16	145.08	272.14	136.57			3
14	1554.73	777.87	1537.71	769.36	1536.72	768.86	G	232.14	116.57	215.11	108.06			2
15							R	175.12	88.06	158.09	79.55			1

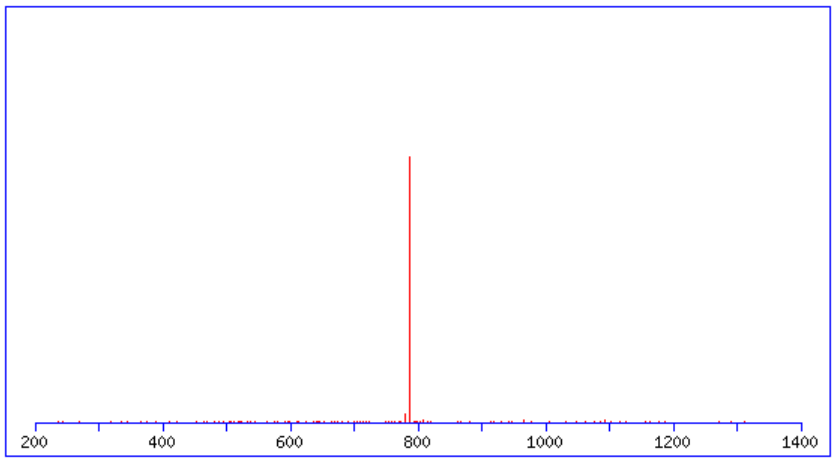
Peptide #7:
GYRPQFYFR + sulfonation
***m/z*: 657.29**
Charge: 2+
RT: 36.8 min

MS2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 617.64



Peptide #8:
GQVLAKPGTIKPHTK + sulfonation
***m/z*: 827.95**
Charge: 2+
RT: 22.2 min

MS^2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 788.08



Peptide #9:
TTLTAAITTVLAK + sulfonation
***m/z*: 692.38**
Charge: 2+
RT: 54.8

MS2 fragmentation spectrum
Neutral loss: 80 Da at *m/z* = 652.30

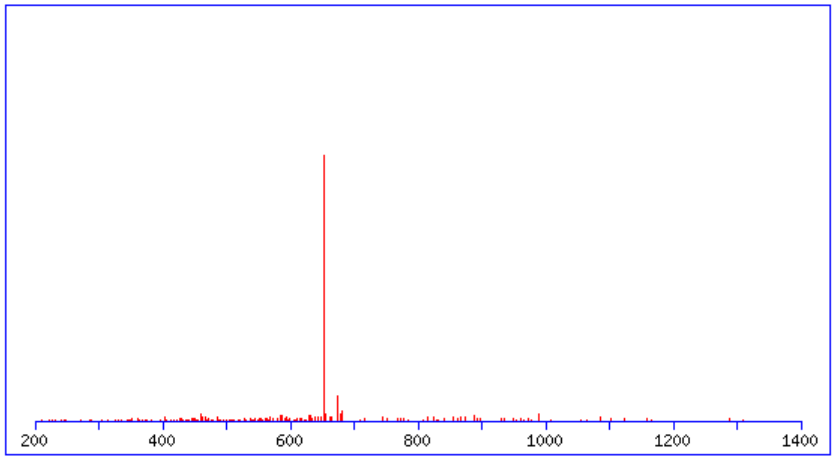
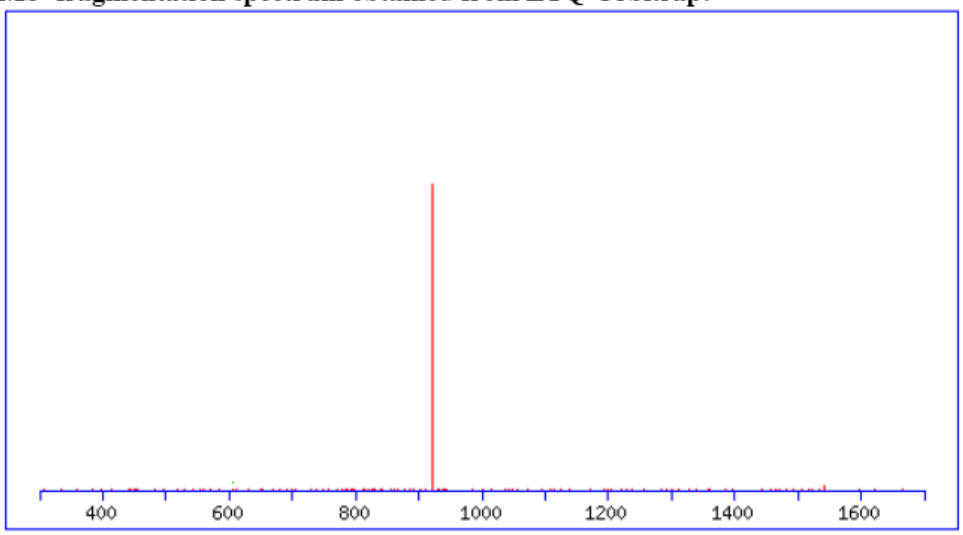


Figure A4 : MS^2 spectra from sulfated yeast enolse peptide SIVPSGASTGVHEALEMR obtained from LTQ-Orbitrap (top) and Q-TOF (bottom.

SIVPSGASTGVHEALEMR + sulfation (*m/z*: 960.95^{2+})
MS^2 fragmentation spectrum obtained from LTQ-Orbitrap:

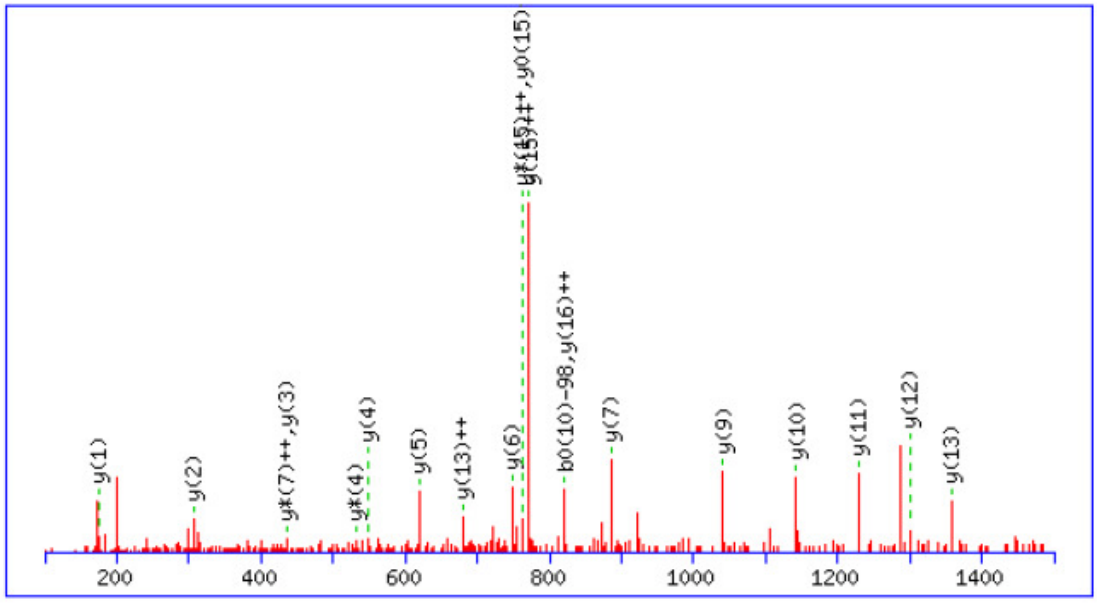


Variable modifications:
S7 or S8: Sulfation
Matched ions

#	b	b^{++}	b^0	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^0	y^{0++}	#
1	88.04	44.52	70.03	35.52	S							18
2	201.12	101.07	183.11	92.06	I	1735.88	868.44	1718.85	859.93	1717.87	859.44	17
3	300.19	150.60	282.18	141.59	V	1622.80	811.90	1605.77	803.39	1604.78	802.90	16
4	397.24	199.13	379.23	190.12	P	1523.73	762.37	1506.70	753.85	1505.72	753.36	15
5	466.27	233.64	448.26	224.63	S	1426.67	713.84	1409.65	705.33	1408.66	704.84	14
6	523.29	262.15	505.28	253.14	G	1357.65	679.33	1340.63	670.82	1339.64	670.32	13
7	594.32	297.67	576.31	288.66	A	1300.63	650.82	1283.60	642.31	1282.62	641.81	12
8	681.36	341.18	663.35	332.18	S	1229.59	615.30	1212.57	606.79	1211.58	606.30	11
9	782.40	391.71	764.39	382.70	T	1142.56	571.78	1125.54	563.27	1124.55	562.78	10
10	839.43	420.22	821.42	411.21	G	1041.51	521.26	1024.49	512.75	1023.50	512.26	9
11	938.49	469.75	920.48	460.75	V	984.49	492.75	967.47	484.24	966.48	483.74	8
12	1075.55	538.28	1057.54	529.27	H	885.42	443.22	868.40	434.70	867.41	434.21	7
13	1204.60	602.80	1186.59	593.80	E	748.37	374.69	731.34	366.17	730.36	365.68	6
14	1275.63	638.32	1257.62	629.31	A	619.32	310.17	602.30	301.65	601.31	301.16	5

15	1388.72	694.86	1370.71	685.86	**L**	548.29	274.65	531.26	266.13	530.28	265.64	**4**
16	1517.76	759.38	1499.75	750.38	**E**	435.20	218.10	418.18	209.59	417.19	209.10	**3**
17	1648.80	824.90	1630.79	815.90	**M**	306.16	153.58	289.13	145.07			**2**
18					**R**	175.12	88.06	158.09	79.55			**1**

SIVPSGASTGVHEALEMR + sulfation (m/z: 960.95^{2+})
MS^2 fragmentation spectrum obtained from Q-TOF:



Variable modifications:
S7 or S8: Sulfation
Matched ions

#	b	b^{++}	b^0	b^{0++}	Seq.	y	y^{++}	y*	y*$^{++}$	y^0	y^{0++}	#
1	168.01	84.51	150.00	75.50	S							18
2	281.09	141.05	263.08	132.04	I	1753.89	877.45	1736.86	868.94	1735.88	868.44	17
3	380.16	190.58	362.15	181.58	V	1640.81	**820.91**	1623.78	812.39	1622.80	811.90	16
4	477.21	239.11	459.20	230.10	P	1541.74	**771.37**	1524.71	**762.86**	1523.73	**762.37**	15
5	564.24	282.63	546.23	273.62	S	1444.68	722.85	1427.66	714.33	1426.67	713.84	14
6	621.26	311.14	603.25	302.13	G	**1357.65**	**679.33**	1340.63	670.82	1339.64	670.32	13
7	692.30	346.65	674.29	337.65	A	**1300.63**	650.82	1283.60	642.31	1282.62	641.81	12
8	779.33	390.17	761.32	381.17	S	**1229.59**	615.30	1212.57	606.79	1211.58	606.30	11
9	880.38	440.69	862.37	431.69	T	**1142.56**	571.78	1125.54	563.27	1124.55	562.78	10
10	937.40	469.20	919.39	460.20	G	**1041.51**	521.26	1024.49	512.75	1023.50	512.26	9
11	1036.47	518.74	1018.46	509.73	V	984.49	492.75	967.47	484.24	966.48	483.74	8
12	1173.53	587.27	1155.52	578.26	H	**885.42**	443.22	868.40	**434.70**	867.41	434.21	7
13	1302.57	651.79	1284.56	642.78	E	**748.37**	374.69	731.34	366.17	730.36	365.68	6
14	1373.61	687.31	1355.60	678.30	A	**619.32**	310.17	602.30	301.65	601.31	301.16	5

15	1486.69	743.85	1468.68	734.85	**L**	**548.29**	274.65	**531.26**	266.13	530.28	265.64	**4**
16	1615.74	808.37	1597.73	799.37	**E**	**435.20**	218.10	418.18	209.59	417.19	209.10	**3**
17	1746.78	873.89	1728.77	864.89	**M**	**306.16**	153.58	289.13	145.07			**2**
18					**R**	**175.12**	88.06	158.09	79.55			