12-2004

Structure-Function and Regulation of Yeast Ribonucleotide Reductase Inhibitor, Sml1

Tomoaki Uchiki
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Tomoaki Uchiki entitled "Structure-Function and Regulation of Yeast Ribonucleotide Reductase Inhibitor, Sml1." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

Chris Dealwis, Major Professor

We have read this dissertation and recommend its acceptance:

Robert Hettich, Jeffrey Becker, Engin Serpersu, Elizabeth Howell

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting here with a dissertation written by Tomoaki Uchiki entitled "Structure-function and regulation of yeast ribonucleotide reductase inhibitor, Sml1" I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

Chris Dealwis
Major Professor

We have read this dissertation
and recommend its acceptance:

Robert Hettich

Jeffrey Becker

Engin Serpersu

Elizabeth Howell

Accepted for the Council:

Anne Mayhew
Vice Chancellor and Dean of
Graduate Studies

(Original signatures are on file with official student record)

STRUCTURE-FUNCTION AND REGULATION OF YEAST RIBONUCLEOTIDE REDUCTASE INHIBITOR, SML1

A Dissertation Presented for the Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Tomoaki Uchiki
December 2004

ACKNOWLEGEMENTS

I wish to thank all those who helped me complete my Doctor of Philosophy degree in Life Sciences. I am very grateful to my research advisors Dr. Chris Dealwis and Dr. Robert Hettich for their tireless effort and guidance for my learning and accomplishment in science, technology and personal development. I would also like to thank Dr. Jeffrey Becker, Dr. Elizabeth Howell, and Dr. Engin Serpersu, for serving on my committee. I would like to thank Dr. Rodney Rothstein's group at Columbia University and Dr. JoAnn Stubbes' group at MIT for kindly providing biological materials and critical information indispensable for this research. I would like to thank Dr. Daniel Roberts for his expert advice in protein phosphorylation. I would like to thank the members of ORNL organic and biological mass spectrometry group for their technical advice and assistance in mass spectrometry. I am particularly grateful to a former member of Deralwis' lab, Dr. Vibha Gupta, for her work in development of this project as well as personal guidance in my graduate education. I would also acknowledge a former member of Dealwis' lab, Mrs. Leslee Dice, for her construction of most of the Sml1 mutant expression plasmids used in this research. I am also grateful to both former and current members of Chris Dealwis' lab for their help in solving technical and scientific problems as well as their friendship. I would like to acknowledge UT-ORNL Graduate School of Genome Science and Technology, the Center for Excellence in Structural Biology at the University of Tennessee and National Institute of Health (1R01CA100827-01 and GM50237) for financial support. Lastly, I would like to thank my parents in Japan, whose help and encouragement made my graduate study possible.

ABSTRACT

Sml1 is a small protein in *Saccharomyces cerevisiae* that inhibits the activity of ribonucleotide reductase (RNR) through its interactions with the large subunit of RNR. RNR catalyzes the reduction of nucleotide diphosphates (NDPs) to deoxynucleotide diphosphates (dNDPs) that is the rate-limiting step of *de novo* deoxynucleotide triphosphate (dNTP) synthesis. The cellular level of Sml1 is regulated by DNA damage and replication block response through its phosphorylation by the Dun1 kinase. The goal of this dissertation research is to elucidate structure-function and regulation of Sml1. First, biochemical characterization of recombinant Sml1 was conducted using mass spectrometry and gel filtration chromatography (Chapter 3 and 4). The data shows that a disulfide bond and non-covalent interactions mediate Sml1 oligomerization. Furthermore, alkali metal adducts (Na^+/K^+) that bind strongly with Sml1 were found. Second, the phosphorylation of Sml1 by the Dun1 kinase was studied by a combination of mass spectrometry, site directed mutagenesis, and P32 incorporation (Chapter 5). Three phosphorylation sites of Sml1 (Ser56, Ser58 and Ser60) were identified. The data also reveals that the Dun1 kinase requires an acidic residue at the +3 position and there is cooperativity between the phosphorylation sites. Third, the relationship between Sml1 phosphorylation and the Sml1-Rnr1 interactions were investigated based on P32 incorporation, fluorescence spectroscopy, and a RNR activity assay (Chapter 6). We demonstrated that the Sml1-Rnr1 interactions reduced the phosphorylation levels of Sml1 by making the phosphorylation sites less accessible. Our data also suggest that

phosphorylation of Sml1 weakens the ability of Sml1 to inhibit RNR. Taken together, this work has provided in-depth insights of Sml1's structure-function and regulation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

°C	degree Celsius
cm^{-1}	per centimeter
$cm^{-1}M^{-1}$	per centimeter per molar
cpm	counts per minute
Da	Dalton
dpm	dose per minute
ε_{259}	extinction coefficient at 259nm
g	gram
K_d	dissociation constant
kDa	kilodalton
kHz	kilohertz
kPa	kiloPascal
K_m	Michaelis-Menten constant
µg	microgram
µl	microliter
µM	micromolar
µCi	microcurie
mg	milligram
ml	milliliter
mm^2	millimeter square
msec	millisecond
m/z	mass/charge
mM	millimolar
M^{-1}	per molar
Mr	nominal mass
OD_{600}	optical density at 600nm
Nm	nanometer
ppm	part per million
pmol	picomole
psi	pound per square
S^{-1}	per second
x g	gravity
v/v	volume per volume
wt/vol	weight per volume

Abbreviations

ATP	adenosine tri-phosphate
Ala	alanine
Asp	aspartic acid
BME	beta-mercaptoethanol
BSA	bovine serum albumin

CAD	collisional activated dissociation
cAMP	cyclic adenosine monophosphate
Cdk	cyclin dependent kinase
CDP	cytidine diphosphate
CHK	Checkpoint kinase
CKII	casein kinase II
CNBr	cyanogen bormide
Cys	cysteine
CV	column volume
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dATP	deoxyadenosine triphosphate
dCDP	deoxycytidine diphosphate
dNDP	deoxynucleotide diphosphate
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
E.coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
FTICR	Fourier transform ion cyclotron
FHA domain	forkhead associated domain
Glu	glutamic acid
GST	glutathione S-transferase
HU	hydroxyurea
FWHM	full width half maximum
IANBD-amide	*N*,*N*'-dimethyl-*N*-(iodoacetyl)-*N*'-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamine
ICR	ion cyclotron resonance
I.D.	inner diameter
IMAC	immobilized metal affinity chromatography
IOD	image optical density
MALDI	matrix-assisted laser desorption Ionization
MMS	methyl methane sulfonate
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MWCO	molecular weight cut off
NDP	nucleotide diphosphate
ORF	open reading frame
PCNA	proliferating cell nuclear antigen
PBS	phosphate buffer saline
PIKK	PI3 kinase like kinase
PKA	cAMP-dependent kinase
PMSF	phenylmethylsulfonyl fluoride
rf	radio frequency

RNR	ribonucleotide reductase
S.cerevisiae	*Saccharomyces cerevisiae*
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacyrlamide gel electrophoresis
Ser	serine
SORI	sustained off resonance irradiation
Std	standard deviation
Thr	threonine
TCEP	tris(2-carboxyethyl) phosphine hydrochloride
TB	Terrific Broth media
TOF	time of flight
Tyr	tyrosine

Chapter 1. Background and significance

Introduction

Sml1 is a small protein in *Saccharomyces cerevisiae* that inhibits the activity of ribonucleotide reductase (RNR). RNR catalyzes reduction of nucleotide diphosphate (NDP) to deoxynucleotide diphosphate (dNDP), which is the rate-limiting step of *de novo* deoxynucleotide triphosphate (dNTP) synthesis. The cellular level of Sml1 is regulated by DNA damage and replication block response through its phosphorylation.

The goal of this chapter is to discuss the role of Sml1 during DNA damage and replication block responses. First, it is important to define DNA damage and replication block responses along with the historical background of the field. Second, a description of the major upstream components of DNA damage and replication pathways, namely phosphoinositide 3 kinase like kinase (PIKK) and checkpoint kinase (CHK), which control most of the biological processes involving the DNA damage response, will be given. Because Sml1 is found in yeast, *S. cerevisiae*, and yeast has been most commonly used as a model organism to study DNA damage and replication block, the discussion will be mainly based on discoveries made in yeast. Thirdly, the Dun1 kinase will be described in detail, as cellular levels of Sml1 is controlled by Dun1. Fourth, the regulation of RNR at multiple levels will be described. Finally, the biochemical properties of Sml1 and the possible mechanism of its degradation will be discussed.

DNA damage and replication block responses

Cells can be potentially exposed to genotoxic stresses that alter the chemical structure of DNA. DNA damage can cause serious problems in DNA replication, chromosome segregation and gene transcription (Elledge 1996; Weinert 1998; Wilson 2004). If DNA damage persists, it can lead to mutations that will be passed onto daughter cells. Such mutations are potentially deleterious or even lethal to the organism; thus, it is necessary for the cell to detect and repair DNA damage promptly. From earlier days, it has been observed that cell cycle progression is halted at certain time points after exposure of the cells to ionizing radiation that causes DNA damage (Wilson 2004). Initially the radiation induced cell cycle arrest was considered a passive cellular consequence of DNA damage (Wilson 2004). However, after discoveries of mutations leading to the absence of the cell cycle arrest after radiation damage, it became obvious that cell cycle arrest was an integral part of the progression of the cell cycle (Terasima and Tolmach 1963; Weinert and Hartwell 1988; Hartwell and Weinert 1989; Weinert and Hartwell 1990; Hartwell, Weinert et al. 1994).

Studies have shown that such active processes either arrest or slow down cell cycle progression at specific periods in G1, S, and G2 phases of the cell cycle (Hartwell, Culotti et al. 1974; Elledge 1996; Weinert 1998; Zhou and Elledge 2000; Wilson 2004); today such periods of cell cycle arrest or slowdown are referred to as DNA damage checkpoints (Weinert and Hartwell 1988). The mutations that cause defects in the cell cycle arrest result in higher sensitivity (less resistance) to agents that create DNA

damages (Weinert and Hartwell 1988; Zhou and Elledge 1993; Elledge 1996; Weinert 1998; Zhao, Muller et al. 1998; Wilson 2004). The exact mechanism and role of the cell cycle arrest or slowdown is not yet fully understood. In the conventional paradigm, cell cycle arrest or slowdown is considered to be a process that allows enough time for the cells to repair DNA (Weinert and Hartwell 1988; Hartwell, Weinert et al. 1994; Elledge 1996; Wilson 2004). However, at least in the case of S phase cell cycle slowdown, the DNA repair process is the major DNA damage response, which consequently causes delay in DNA replication, and the defect of the cell cycle slowdown in some mutant cells is due to a defect of the DNA repair process (Rhind and Russell 2000).

During the study of the pathways responsible for cell cycle arrest, it was discovered that the same pathways not only induce cell cycle arrest, but also up-regulate the DNA repair process and transcription of genes responsible for DNA repair (Elledge 1996; Rhind and Russell 2000). Furthermore, these pathways up-regulate ribonucleotide reductase (RNR) activity in multiple fashions that lead to increased levels of cellular deoxynucleotide triphosphate pools (Zhou and Elledge 1993; Zhao, Muller et al. 1998; Ouspenski, Elledge et al. 1999). Today, cell cycle arrest, up-regulation of DNA repair, transcriptional activation of repair genes, and up-regulation of RNR are considered to be components of a cellular response to DNA damage that increases the cell's ability to repair DNA. The whole process is termed the DNA damage response.

DNA damage in the S-phase causes an additional problem that is not encountered in other phases of the cell cycle termed the replication block (Loeb and Kunkel 1982; Rhind and Russell 2000; Osborn, Elledge et al. 2002). If DNA damage is present during

the S-phase, progression of the replication fork is stalled at the point of the DNA damage. The stalled replication forks activate the DNA damage response pathways to repair the damage, and the repair process prohibits replication origin firing[1] (Rhind and Russell 2000). In addition to DNA damage, the replication block response can be elicited by other stimuli that stop progression of replication forks, such as reduced levels of dNTP pools induced by inhibition of ribonucleotide reductase with hydroxyurea (HU) (Elledge and Davis 1989; Lopes, Cotta-Ramusino et al. 2001) and mutation of proteins responsible for DNA replication such as DNA polymerases (Bhaumik and Wang 1998; Liu, Bhaumik et al. 1999).

The perturbation of replication can lead to a number of deleterious events including increased mutagenesis and chromosomal aberration (Loeb and Kunkel 1982). If the damage is not repaired or replication forks cannot progress after a certain period of time, the cells resume replication, skipping the point of the damage (Sandell and Zakian 1993; Toczyski, Galgoczy et al. 1997). This resumption of DNA replication is called adaptation. In such cases, the DNA damage persists as mutations or double strand breaks that can lead to chromosome loss (Sandell and Zakian 1993) or genomic instability (Toczyski, Galgoczy et al. 1997). For multi-cellular organisms, DNA damage can potentially give rise to cancer. The advantage of adaptation to a organism's survival has

[1] Replication origin firing is the initiation of the replication process at an origin of replication. At the origin of replication, several proteins including DNA polymerase are assembled, and DNA double strand is unwound upon the origin firing. Note that Eukaryotes have multiple origins of replication, and initiation of replication (firing) at some of them occurs in early stage of S-phase while firing at some of them occurs in a later stage.

not been demonstrated (Toczyski, Galgoczy et al. 1997; Weinert 1998). However, adaptation is likely to be an advantage for unicellular organisms.

Much of the knowledge about the organization of the DNA damage response came from studies involving yeasts *S. cerevisiae* and *S. pombe*, while studies vertebrate systems such as mammalian cells have made significant progress in recent years (Weinert 1998; Osborn, Elledge et al. 2002). The overall molecular anatomy of the DNA damage response is shared among all eukaryotes (Zhou and Elledge 2000). Essentially, the DNA damage response consists of three types of components; DNA damage sensors, signal transducers, and effectors (Elledge 1996; Zhou and Elledge 2000) (Figure1). DNA damage sensors are responsible for detecting DNA damage and trigger the signals that activate the transducers. The transducers in turn activate multiple pathways leading to the effectors. Effectors are the components that are responsible for cell cycle arrest and DNA damage repair.

PIKK and CHK: Major upstream components of DNA damage and replication block responses

Recently the focus of DNA damage response research has been shifting towards understanding the mechanisms by which individual components play their roles in the pathway. However, numerous components in the pathway are yet to be identified. The molecular nature of the sensors in the DNA damage response pathway is least well understood, but some proteins have been identified as candidates for damage sensors based on their ability to bind naked DNA and activate the DNA damage response

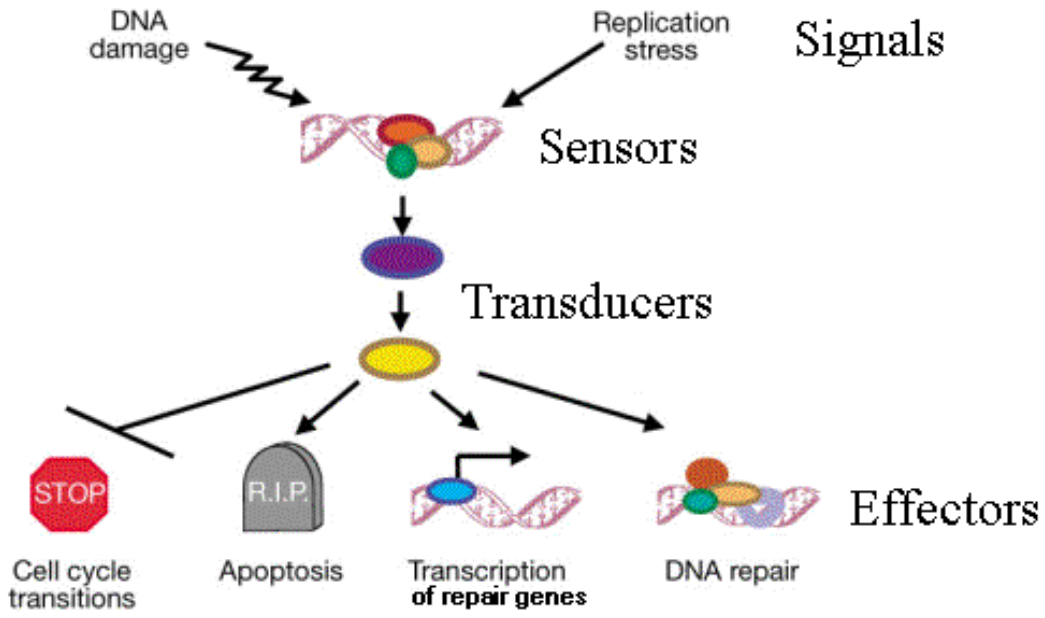

Figure 1. Current view of DNA damage response.

(The image was taken from Zhou and Elledge 2000) Arrowheads represent activating events and perpendicular ends represent inhibitory events. Replication stress is caused by stalled replication forks, DNA damage, low levels of dNTPs, and mutation/inhibition of components in replication machinery. The stop sign indicates cell cycle arrest and tombstone indicates activation of apoptosis leading to cell death.

(Elledge 1996; Zhou and Elledge 2000; Osborn, Elledge et al. 2002). One class of candidate sensor is distinct subfamily of phosphoinositide 3 kinase (PI3 kinase), "PI3 kinase like kinase" (PIKK) (Table 1). PIKK acts as both a DNA damage sensor and a signal transducer upstream of the DNA damage and replication block responses (Yang, Xu et al. 2004). Although they have sequence homology to other PI3 kinases, these proteins have protein kinase activities specific for their biological functions (Zakian 1995; Yang, Xu et al. 2004). In *S. cerevisiae*, Mec1[2] is the PIKK that induces all DNA damage (i.e. damage during G1, S and G2) and replication block responses (Elledge 1996; Weinert 1998; Abraham 2004). Furthermore, deletion or mutation in the *MEC1* gene[3] is lethal even in the absence of DNA damage, suggesting that *MEC1* is responsible for not only the DNA damage response but also regulation of the cell cycle in unperturbed cells (Zheng, Fay et al. 1993; Kato and Ogawa 1994). In mammals, ATM and ATR are responsible for response to DNA double strand break (Elledge 1996; Weinert 1998; Osborn, Elledge et al. 2002; Wilson 2004). In humans, mutations in the ATM gene result in an autosomal recessive disease, ataxia telangiectasia (AT), a multi-system disorder associated with a high risk of cancer (Abraham 2003). In addition to response to DNA double strand breaks, ATR is also responsible for replication block responses (Osborn, Elledge et al. 2002; Yang, Xu et al. 2004). PIKK exists in stable

---

[2] *TEL1* is another PIKK in S.*cerevisiae* that has redundant role as *MEC1*, and it is the homolog of mammalian ATM. Elledge, S. J. (1996). "Cell cycle checkpoints: preventing an identity crisis." Science **274**(5293): 1664-72.

[3] By convention a yeast gene's name is abbreviated in italic four capital letters (e.g. *MEC1)*. The mutant of the gene is spelled with italic four lower case letters (e.g. *mec1)*. Protein is spelled with four letters starting with capitals followed by lower case letters (e.g. Mec1). Although "p" is often added to the end of the protein name (e.g. Mec1p), I decided not to follow this convention throughout this thesis.

Table 1. Conservation of PIKK and CHK in Eukaryotes.

Family	Biochemical activity	Budding yeast	Fission yeast	Vertebrates
PIKK	Protein kinase	Mec1 Tel1	Rad3 Tel1	ATR ATM
PIKK associated protein	Recruit PIKK to site of DNA damage	Ddc1	Rad26	ATRIP
CHK	Protein kinase Binds to phosphoprotein	Rad53 Chk1 Dun1	Cds1 Chk1	Chk2 Chk1

complexes with other proteins that recognize DNA damage (Table 1). In *S. cerevisiae*, Ddc2 is a protein that forms a complex with Mec1 (Paciotti, Clerici et al. 2001). The Mec1-Ddc2 complex is localized to a double strand breaks induced by HO endonuclease[4] (Kondo, Wakayama et al. 2001; Melo, Cohen et al. 2001) and to single stranded DNA at the teleomere caused by a mutation in the teleomere binding protein Cdc13 (Melo, Cohen et al. 2001). These studies suggest a possibility that the localization of Mec1 to the site of DNA damage brings about the activation of Mec1 kinase. The link between localization and activation of Mec1 is not well understood. Once activated, Mec1 phosphorylates several proteins in pathways responsible for cell cycle arrest, DNA repair, transcription of DNA repair genes, and up-regulation of RNR activity (Weinert 1998). Furthermore, in mammalian systems, ATM and ATR induce apoptosis by phosphorylating the tumor suppressor, p53 (Banin, Moyal et al. 1998; Canman, Lim et al. 1998; Yang, Xu et al. 2004).

A downstream target of PIKK during DNA damage and replication block responses is CHK kinase (Table 1). CHK kinase consists of a serine/threonine protein kinase domain and an forkhead associated (FHA) domain, which recognizes specific phosphoproteins (Durocher, Henckel et al. 1999; Durocher and Jackson 2002) In *S. cerevisiae*, Rad53 is a CHK kinase responsible for majority of the signaling in radiation damage (RAD) pathways during DNA damage and replication block responses. (Sanchez, Desany et al. 1996; Sun, Fay et al. 1996). Rad53 is activated in a Mec1

[4] In haploid yeast cells in nature, HO nuclease creates a double strand break at the MAT locus for mating type switching. Many laboratory strains lack HO nuclease, and their mating type is fixed. However, by introducing inducible HO nuclease, it is possible to artificially induce a double strand break at MAT locus in a controlled manner.

dependent manner through its hyperphosphorylation. Rad9 is a protein that acts as a mediator between Mec1 and Rad53. It has been demonstrated that Mec1 phosphorylates Rad9, and the phosphorylation of Rad 9 is necessary for the Rad53 phosphorylation (Sanchez, Desany et al. 1996). Furthermore, Rad53 specifically interacts with phosphorylated Rad 9 via its FHA domain (Schwartz, Duong et al. 2002). Rad9 is also hyperphsphorylated after DNA damage by Mec1, and the phosphorylated Rad9 alone can catalyze the autophosphorylation of Rad53[5] (Gilbert, Green et al. 2001). The activation of Rad53 kinase results in the release of Rad53 from Rad9 (Gilbert, Green et al. 2001). These observations can be formulated as follows: 1) Mec1 kinase directly phosphorylates Rad9; 2) phosphorylated Rad9 acts as scaffold to bring two Rad53 molecules to a proximity and catalyze trans-autophosphorylation of Rad53; 3) autophosphorylation of Rad53 causes activation of Rad53 to phosphorylate other proteins as well as the release of Rad53 from Rad9 (Toh and Lowndes 2003). Once activated, Rad53 phosphorylates proteins in the RAD pathway that induce G1 and G2/M cell cycle arrest, escape from the DNA damage response, up-regulation of DNA repair gene transcriptions and up-regulation of RNR (Elledge 1996; Weinert 1998; Osborn, Elledge et al. 2002; Wilson 2004).

DNA damage and replication block response is a large-scale cellular process composed of sensors, signal transducers, and effectors. PIKK and CHK play central roles

[5] Rad9 extracted from the yeast cells that are irradiated by UV exists as two forms, a hypophosphorylated form of ≥850kDa oligomer and hyperphosphorylated form of 560kDa tetramer. These two forms can be separated by gel filtration chromatography. When recombinant Rad53 expressed in *E.coli* is incubated with the hyperphosphorylated Rad9 and ATP, Rad53 is autophosphorylated, while incubation of Rad53 with hypophosphorylated Rad9 (and ATP) does not cause Rad53 phosphorylation (Gilbert, C. S., C. M. Green, et al. (2001). "Budding yeast Rad9 is an ATP-dependent Rad53 activating machine." Mol Cell **8**(1): 129-36.

in the activation of RAD pathways that lead to elevation of the cell's capacity to repair damaged DNA and to cause replication blocks. It should be noted that both PIKK and CHK are protein kinases, and the activation of the pathways is achieved by phosphorylation of other proteins on the pathway. Mec1 and Rad53 are the homologues of PIKK and CHK in yeast *S. cerevisiae.* One of the targets for Rad53 is the Dun1 kinase, which is involved in the up-regulation of the transcription and the activity of ribonucleotide reductase (Figure 2).

The DUN1 kinase

Cellular levels of deoxynucleotides are elevated both in the presence of DNA damage and during S-phase, reflecting up-regulation of ribonucleotide reductase (RNR) (Reichard 1988; Chabes, Georgieva et al. 2003). Regulation of RNR is an essential[6] function of Mec1 and Rad53 in *S.cerevisiae* (Elledge, Zhou et al. 1993). For example, over-expression of the genes for the large subunit of RNR, *RNR1* and *RNR3*, rescues lethality caused by deletion of *MEC1* or *RAD53*, although it does not rescue deletion-induced defects in cell cycle arrest (Desany, Alcasabas et al. 1998). Dun1[7] acts a mediator for Mec1 and Rad53 to up-regulate RNR. In *S.cerevisiae*, transcription of all RNR genes (*RNR1, RNR2, RNR3* and *RNR4*) is up-regulated by DNA damage and replication blocks in a *MEC1 / RAD53* dependent manner

[6] "Essential" in this context means that the function (regulation of RNR) is indispensable for survival of the organism.

[7] Although Dun1 is important for regulation of RNR by Mec1 and Rad53, deletion of the *DUN1* gene is not lethal indicating that there are other pathways that partially mediate regulation of RNR by Mec1 and Rad53.

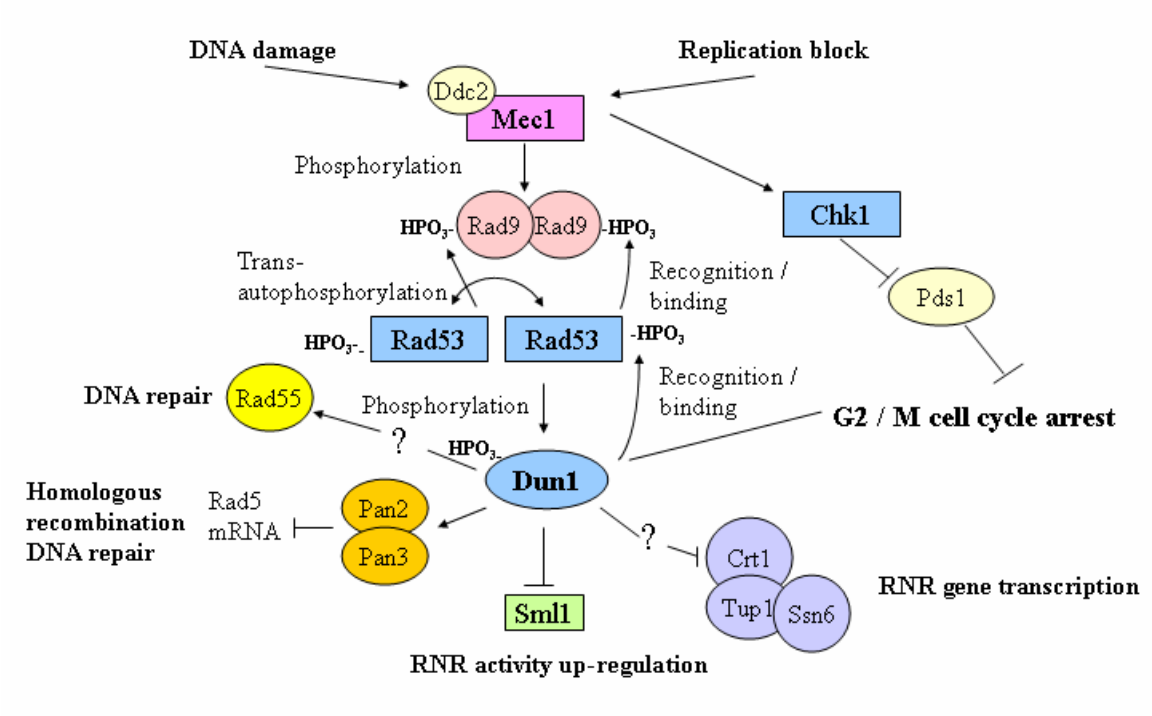

Figure 2. Mec1 / Rad53 / Dun1 dependent pathways in *S.cerevisiae* in DNA damage and replication block responses.

Arrowheads represent activating events and perpendicular ends represent inhibitory events. Lines without arrowheads or perpendicular ends represent pathways that are not studied in depth. It is not known whether these pathways are operated by inhibitory or activating mechanism. This figure focuses on pathways regulated by the Dun1 kinase, and most other pathways regulated by Mec1 and Rad53 are omitted.

(Elledge and Davis 1987; Elledge and Davis 1990; Zhou and Elledge 1993; Huang and Elledge 1997). Elledge and coworkers (Zhou and Elledge 1993) isolated "damage un-inducible" (*dun*) mutants that are incapable of inducing *RNR3* transcription in the replication blocks caused by hydroxyurea (HU). These mutants are recessive and segregate as a single gene mutation named *dun1*. The *dun1* mutant is more sensitive to DNA damage induced by methyl methane sulfonate (MMS) and UV or replication blocks induced by HU. The *dun1* mutant strains are also incapable of induction of other RNR genes (*RNR1, RNR2* and *RNR4*) in response to replication blocks and DNA damage (Zhou and Elledge 1993; Huang and Elledge 1997; Huang, Zhou et al. 1998). Furthermore *dun1* mutant strains show defects during G2 cell cycle arrest (Zhou and Elledge 1993), indicating that Dun1 has a function besides transcriptional regulation of RNR genes since over-expression of RNR genes does not have an apparent role in cell cycle arrest. The *DUN1* gene was cloned by complementation of the *dun1* mutant and identified as an ORF encoding a protein of 513 amino acid residues. The C terminus of the Dun1 amino acid sequence displays significant homology with the catalytic domains of Serine/ Threonine kinases (Zhou and Elledge 1993), while its N terminus shows homology to FHA domain, a module that recognizes specific phosphoproteins (Bashkirov, Bashkirova et al. 2003). The kinase activity of Dun1 is essential for its function to activate transcription of RNR genes, and a mutation that abolishes kinase activity of Dun1 causes a defect in the induction of RNR genes (Zhou and Elledge 1993).

On the other hand, the FHA domain of Dun1 is important for the activation of the Dun1 kinase in response to DNA damage. Dun1 is hyper-phosphorylated after DNA

damage in a *MEC1* and *RAD53* dependent manner (Huang, Zhou et al. 1998; Bashkirov, Bashkirova et al. 2003). A direct trans-phosphorylation of Dun1 by Rad53 *in vitro* and physical interaction of Dun1 and Rad53 *in vivo* has been demonstrated (Bashkirov, Bashkirova et al. 2003). Furthermore, this study demonstrated that interaction of Rad53 and Dun1 can be abolished by deletion of Dun1's FHA domain. In addition, Dun1's FHA domain is required for the *in vivo* function of Dun1 in RNR gene transcription, G2/M cell cycle arrest and phosphorylation of a DNA repair protein Rad55 (Bashkirov, King et al. 2000; Bashkirov, Bashkirova et al. 2003). Based on these observations, Heyer and coworkers (Bashkirov, Bashkirova et al. 2003) proposed the following model; 1) Upon DNA damage, Rad53 is phosphorylated; 2) Dun1 specifically binds to phosphorylated form of Rad53 via its FHA domain; 3) The interaction between Rad53 and Dun1 triggers the trans phosphorylation of Dun1 by Rad53, which up-regulates the Dun1 kinase activity. After its activation, Dun1 phosphorylates its substrates on the pathways leading to DNA repair, G2/M cell cycle arrest, and up-regulation of RNR.

By screening mutants that constitutively induce RNR transcription (constitutive RNR transcription [CRT]), Elledge and co-workers (Huang, Zhou et al. 1998) isolated the *CRT1* gene. The *crt1* mutant is epistatic[8] to the *dun1* mutant. They also found that *tup1* and *ssn6* mutants are also epistatic to *dun1* mutants. TUP1 and SSN6 encode general gene repressors that are recruited to a specific promoter by other proteins. These

[8] Generally epistasis refers to an interaction between genes in which an allele or mutation of the first gene (e.g. *CRT1*) masks a trait (e.g. inducibility of RNR3 gene) dependent on an allele or mutation of a second gene (e.g. DUN1). In this case, *crt1, tup1* or *ssn6* mutations eliminates defect in RNR3 expression caused by the *dun1* mutation. (i.e. mutation of the *DUN1* gene causes a defect in RNR3 expression. However, this defect is masked by a mutation in *CRT1, TUP1* or *SSN6* genes).

observations indicates that *CRT1, TUP1* and *SSN6* act downstream of *DUN1* in the pathway. ORF of *CRT1* encodes a protein of 771 amino acid residues that is homologous to the mammalian RFX family of DNA binding proteins. The mutation of *CRT1* rescues the lethality of the *mec1* mutation and result in constitutive expression of *RNR2, RNR3* and *RNR4* genes; over-expression of wild type Crt1 causes lethality presumably due to the suppression of RNR genes. The Crt1 protein binds to a conserved DNA sequence, the X-box, which is found in the promoter region of RNR and other genes. Crt1 also physically interacts with the general co-repressor protein complex of Tup1 and Ssn6. Furthermore, Crt1 is hyper-phosphorylated in a *DUN1* dependent manner upon DNA damage induced by MMS and upon replication block induced by HU. In addition, in the presence of DNA damage and replication block, Crt1 dissociates from the X-box. Based on these observations, the following model has been proposed (Huang, Zhou et al. 1998): In unperturbed cells, Crt1 represses RNR gene transcription by binding to the X box in the promoter region of the RNR genes and recruiting the co-repressors Tup1 and Ssn6. Once the Dun1 kinase is activated during the DNA damage response, Crt1 is hyperphosphorylated, and dissociates from the X box resulting in expression of the RNR genes. Once the Dun1 kinase is activated in response to DNA damage, Crt1 is hyperphosphorylated, causing it to dissociate from the X box, inducing the expression of the RNR genes. While direct phoshorylation of Crt1 by Dun1 has not been demonstrated, it is likely that there is a substrate of the Dun1 kinase that causes hyperphosphorylation of Crt1. Dun1 also up-regulates RNR activity by the phosphorylation of the RNR inhibitor Sml1 leading to its removal (Zhao and Rothstein 2002). It is remarkable that Dun1 up-

regulates RNR at transcriptional levels and post-translational levels through different pathways.

Apart from its essential role in RNR regulation, Dun1 plays other important roles in DNA damage and replication block responses. The Mec1 / Rad53 / Dun1 pathway is also involved in G2/M cell cycle arrest in the presence of DNA damage caused by expression of HO endonuclease, the *cdc13* mutation, or UV irradiation, while another parallel pathway dependent on Mec1, Chk1 and Pds1 contribute significantly to the G2/M cell cycle arrest (Pati, Keller et al. 1997; Gardner, Putnam et al. 1999; Sanchez, Bachant et al. 1999). On the other hand, Rad55, a double strand break repair protein of the recombination repair pathway, is specifically phosphorylated in a Mec1 / Rad53 / Dun1 dependent manner in the presence of DNA damage and Rad55 is likely to be a terminal target of the DNA repair pathway (Bashkirov, King et al. 2000). Furthermore, in response to replication blocks, Dun1 also down-regulates Rad5 mRNA levels through interaction with the Pan2/Pan3 protein complex leading to homologous recombination DNA repair (Hammet, Pike et al. 2002). These studies independently showed that Dun1 is involved in G2/M cell cycle arrest and DNA repair in addition to up-regulation of RNR (Figure 2).

Regulation of ribonucleotide reductase

A sufficient and balanced level of cellular deoxynucleotide triphosphates (dNTPs) is essential for DNA repair and DNA replication (Reichard 1988; Jordan and Reichard 1998). Insufficient or unbalanced levels of dNTP can lead to genome instability, growth

defects and hyper-sensitivity to DNA damage (Reichard 1988; Zhao, Muller et al. 1998; Ouspenski, Elledge et al. 1999), while unusually high levels of dNTPs can cause infidelity in DNA replication (Chabes, Georgieva et al. 2003). In addition, high levels of dNTP due to over-expression of the small subunit of RNR increases the malignancy of tumor cells by enhancing activation of the Ras/Raf1 pathway leading to expression of oncogenes (Fan, Villegas et al. 1996). Among enzymes responsible for synthesis of dNTPs, ribonucleotide reductase (RNR) is particularly important. RNR catalyzes the conversion of all four nucleotide diphosphates (NDPs: ADP, GDP, CDP and UDP) to deoxynucleotide diphosphates (dNDPs), and this is the rate-limiting step for the *de novo* synthesis of dNTPs (Thelander and Reichard 1979; Jordan and Reichard 1998). RNR plays the major role in controlling amount and balance of dNTPs, and therefore, regulation of RNR is critical for the cell.

RNR is a multimeric enzyme consisting of the large subunit, a dimer of Rnr1, and the small subunit, which is in yeast a heterodimer of Rnr2/Rnr4. The large subunit constitutes the active site (where the catalysis takes place) and the allosteric regulation sites. The small subunit stores a stable radical. This radical is transferred to the large subunit upon substrate binding and plays an essential role in catalysis. Either of these subunits alone is not catalytically active, and the interaction of the large and small subunit is absolutely necessary for activity (Thelander 1973). In budding yeast, *S. cerevisiae*, the *RNR1* gene encodes the large subunit, while the *RNR2* and *RNR4* genes encode the small subunit (Elledge and Davis 1987; Elledge and Davis 1990; Huang and Elledge 1997). Deletion of any of these genes is lethal. Rnr2 and Rnr4 form a hetero-

dimer required for their function. Rnr2 contains the essential tyrosyl radical and di-iron co-factor that generates and maintains the radical (Nguyen, Ge et al. 1999; Chabes, Domkin et al. 2000). Although it lacks residues responsible for the generation of the tyrosyl radical and di-iron co-factor, Rnr4 is indispensable for loading the iron and stabilizing the radical in Rnr2 (Ge, Perlstein et al. 2001; Sommerhalter, Voegtli et al. 2004). *S. cerevisiae* also has a homologous large subunit, Rnr3, although its biological role is unknown. Deletion of the *RNR3* gene does not show an apparent phenotype even in the presence of DNA damage, and the specific activity of Rnr3 protein is less than 1% of that of Rnr1 (Domkin, Thelander et al. 2002). However, in the presence of DNA damage, the *RNR3* gene expression increases 100 fold (Elledge and Davis 1990), and Rnr3 shows a significant activity when it is combined with Rnr1(Domkin, Thelander et al. 2002).

Due to its essential role in DNA damage response and the cell cycle, RNR is tightly controlled in multiple fashions (Figure 3). Generally, RNR is up-regulated during the S-phase of the cell cycle and in the presence of DNA damage.

First, the cellular level of RNR proteins is up-regulated in response to DNA damage (Desany, Alcasabas et al. 1998) and during S-phase (Eriksson, Graslund et al. 1984; Chabes and Thelander 2000). In *S. cerevisiae*, DNA damage and replication blocks activate the Mec1 / Rad53 / Dun1 pathway leading to transcriptional up-regulation of all RNR genes (Zhou and Elledge 1993; Huang, Zhou et al. 1998). Similarly, in mammalian cells, mRNA levels of both the large (R1) and small (R2) subunits significantly increase during S phase and decline when cells progress to G2 and M phases

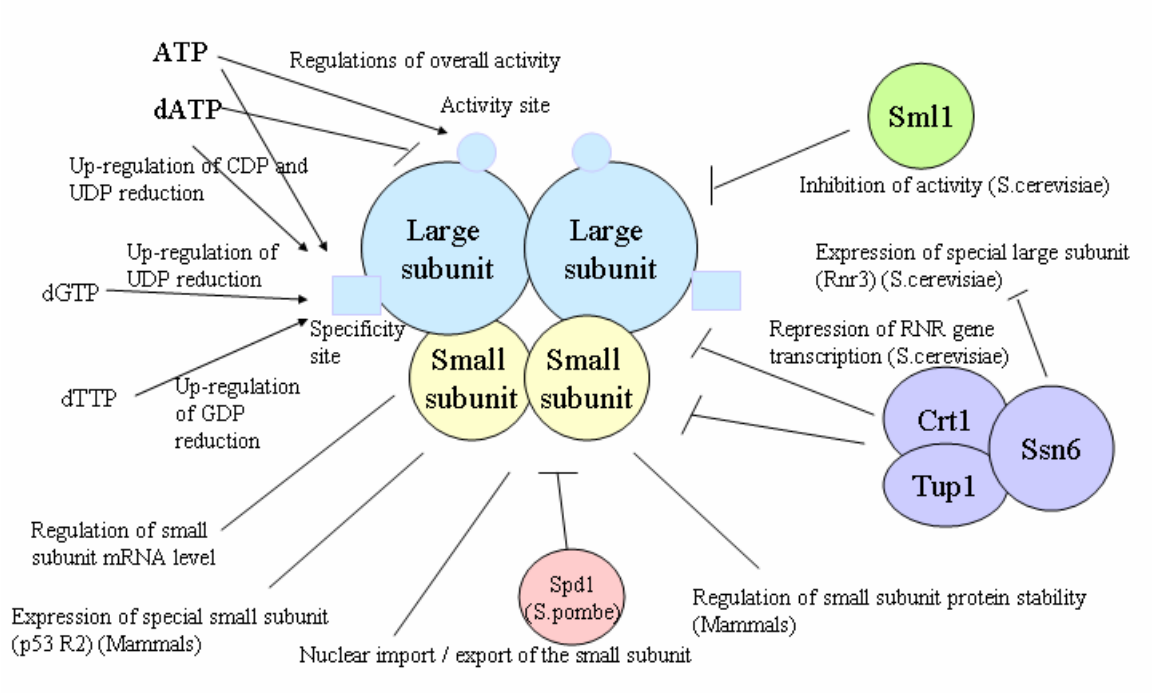

Figure 3. Regulation of ribonucleotide reductase at multiple levels.

Arrowheads represent activating events and perpendicular ends represent inhibitory events. Lines without arrowheads or perpendicular ends represent pathways that are not studied in depth. It is not known whether these pathways are operated by inhibitory or activating mechanism.

(Bjorklund, Skog et al. 1990). Furthermore, in mammalian cells, transcription of a special type of RNR small subunit, p53R2, is up-regulated by the tumor suppressor p53 in response to DNA damage (Tanaka, Arakawa et al. 2000). In mammalian cells, the half life of R2 protein is also increased during S phase and in response to DNA damage or replication blocks (Chabes and Thelander 2000).

Secondly the RNR's enzymatic activity is allosterically regulated by several nucleotide triphosphates (Jordan and Reichard 1998). Two binding sites of the nucleotide triphosphates, the activity site and the specificity site, are located on the large subunit of RNR. Binding of ATP to the activity site increases overall activity of the enzyme, while dATP binding to the same site reduces the overall activity. Binding of nucleotides to the specificity site controls the enzymatic activity toward specific substrates: A widely accepted model is that binding of ATP and dATP increases the reduction of pyrimidine nucleotide diphosphates (CDP and UDP), dTTP increases the GDP reductase activity and dGTP increases the ADP reductase activity (Jordan and Reichard 1998). Furthermore, a recent study indicates that murine ribonucleotide reductase possesses another binding site for ATP that controls activity of the enzyme by changing oligomeric state of the large subunit (Kashlan, Scott et al. 2002). A similar mode of regulation may exist in *S. cerevisiae*, since the oligomeric state of the yeast RNR large subunit (Rnr1) is also shifted from a monomer/dimer mixture to a dimer/tetramer mixture by binding of dTTP (Chabes, Domkin et al. 1999).

Thirdly in eukaryotic systems, cellular RNR activity is also regulated by the sub-cellular localization of the small subunit (Tanaka, Arakawa et al. 2000; Liu, Powell et al.

2003; Yao, Zhang et al. 2003). In the fission yeast (*S. pombe*) and budding yeast (*S. cerevisiae*), during interphase of the unperturbed cell cycle, the small subunit is mostly sequestered within the nucleus, but it is transferred to cytoplasm in response to DNA damage and before the S phase (Liu, Powell et al. 2003; Yao, Zhang et al. 2003). In *S. pombe*, it has been further demonstrated that the protein Spd1 binds to the RNR small subunit in the nucleus to block nuclear export of the small subunit. During S phase or after DNA damage, Spd1 is phosphorylated, which causes ubiquitination and proteasomal degradation of Spd1 leading to nuclear export of the small subunit (Liu, Powell et al. 2003). On the other hand, in human cells, p53R2 is mostly localized in the cytosol during the unperturbed cell cycle, and it is sequestered to nucleus in the presence of DNA damage, where it is believed to perform its function (Tanaka, Arakawa et al. 2000).

Finally specifically in *S. cerevisiae*, a novel protein, Sml1, binds to the large subunit (Rnr1) and inhibits its activity. The focus of this dissertation research is on Sml1, and a detailed background of Sml1 is given in the next section.

Regulation of ribonucleotide reductase by Sml1 and the molecular nature of Sml1

During screening of a recessive lethal *mec1-1* mutant, Rothstein and co-workers (Zhao, Muller et al. 1998) isolated a *mec1-1* homozygous strain that can still survive. The survival of this strain was due to mutation of another gene that suppresses lethality of the *mec1* mutation, so the suppressor mutation was named *sml1* (suppressor of *mec1* lethality). The *SML1* gene was mapped on chromosome XIII at the YML058 ORF, which encodes a small protein of 104 amino acids (Zhao, Muller et al. 1998).

Rothstein and co-workers (Zhao, Muller et al. 1998) concluded that Sml1 is an inhibitor of ribonucleotide reductase based on the following evidence:

1. A strain that carries two copies of *SML1* gene shows a high degree of petite colony formation, which is caused by the insufficiency of deoxynucleotide triphosphates (dNTPs).
2. *sml1Δ* alleviates the reduced levels of dNTPs in a *dun1Δ* strain after DNA damage.
3. The cellular dNTP levels of the *sml1* mutant are higher than those of the wild type strain.
4. Sml1 and the large subunit of ribonucleotide reductase (Rnr1) physically interact *in vivo*.

Interestingly, Sml1 also interacts with Rnr3 *in vivo*, although the biological significance of this interaction has not been investigated (Zhao, Georgieva et al. 2000). No interaction of Sml1 and the small subunit of ribonucleotide reductase (Rnr2/Rnr4) has been observed (Zhao, Georgieva et al. 2000). Later, Thelander and co-workers (Chabes, Domkin et al. 1999) demonstrated that Sml1 inhibits ribonucleotide reductase activity *in vitro*. Based on surface plasmon resonance assay using immobilized Sml1 on a solid surface, they also demonstrated that Sml1 interacts with Rnr1 with a dissociation constant (Kd) of 0.25 to 0.4±0.1μM. (the former value was determined by the k_{on}= 153,000$M^{-1}S^{-1}$ and k_{off}=0.04S^{-1} measured by the assay, while the latter was determined by plotting the intensity of surface plasmon resonance to Rnr1 concentration during the assay. It is interesting to note that k_{on} value of Sml1 binding to Rnr1 is ~10 fold less than rate of diffusion, possibly

suggesting that Sml1 undergoes structural change prior to binding to Rnr1.) In this assay, the stoichiometry of the Sml1-Rnr1 interactions was estimated to be roughly one to one, based on the amount of Sml1 immobilized per mm^2 of the surface and the amount of Rnr1 bound to it. The ribonucleotide reductase activity assay at various concentrations of the Rnr2/Rnr4 complex in the presence of Sml1 and the binding assay of Sml1-Rnr1 interactions in the presence of Rnr2/Rnr4 complex showed no indication of competition between Sml1 and Rnr2/Rnr4 for binding to Rnr1. Interestingly, they also demonstrated that Sml1 can also bind to mouse and human ribonucleotide reductase large subunit (R1) and inhibit RNR activity. In these cases, Sml1 and the small subunits of mouse or human ribonucleotide reductase (R2) did compete for binding to R1. In addition, another study independently showed an interaction of Sml1 and Rnr1 *in vivo* based on co-immunoprecipitation followed by mass spectrometric analysis (Ho, Gruhler et al. 2002).

So far only a few studies have characterized the structure of Sml1. One such study was conducted by scanning mutagenesis of Sml1 followed by a yeast two-hybrid screen (Zhao, Chabes et al. 2001). In this study, the authors recovered over ten point mutants of Sml1 that interact with Rnr1 very poorly. These mutations were located within the last C-terminal 33 residues. Inhibition of ribonucleotide reductase by some of these mutants was also tested and showed results consistent with those of the two-hybrid screen. In addition, deletion of residues 2 to 39 or 28 to 50 did not affect the ability of Sml1 to inhibit ribonucleotide reductase, demonstrating that this region is not necessary for binding of Sml1 to Rnr1. Based on ^{15}N relaxation and secondary $C\alpha$ chemical shift monitored by NMR spectroscopy, this study also showed that three regions of Sml1 have a high degree

of backbone order(Zhao, Georgieva et al. 2000). These regions are residues 4 to 14, 20 to 35, and 61 to 80, and residues 4 to 14 and 61 to 80 adopt an alpha helical backbone conformation[9], while the rest of the molecule does not have a defined structure.

The region of Rnr1 that is responsible for interaction with Sml1 has not been fully identified. A scanning mutagenesis study of Rnr1 showed that the W688G mutant has a higher affinity to Sml1 than wild type Rnr1, suggesting that the region containing W688 may be responsible for the interaction (Georgieva, Zhao et al. 2000).

Based on sedimentation equilibrium and gel filtration chromatography, Dealwis and co-workers (Gupta, Peterson et al. 2004) showed that Sml1 forms a dimer in solution. The presence of reducing agent made no significant difference regarding the oligomeric states of Sml1[10]. Furthermore, the C14S mutant of Sml1, in which sole cysteine was replaced with serine, showed a similar profile in the experiment. In addition, sedimentation equilibrium showed that Δ8 Sml1 was a dimer while Δ20 Sml1 was a monomer. These data indicated that Sml1 exists as a dimer via non-covalent interactions through the region spanning residues 8 and 20. In this study, a computational model of Sml1 predicts that Sml1 has four alpha helices (residues 6 to18, 22 to 32, 59 to 76 and 88 to 95). To date, only a few protein sequences from predicted ORFs in other species of yeast showed weak sequence homology to Sml1, and the structures of these putative proteins have not been investigated. Presumably, Sml1 has unique structural features not previously found in other proteins.

---

[9] Although authors reported that residues 20 to 35 have a high degree of backbone order, structure of this region has not been well characterized.

[10] This result was slightly different from observation in gel filtration chromatography study shown in Chapter 4.

Phosphorylation and degradation of Sml1

The cellular level of Sml1 is regulated by the DNA damage response and the cell cycle (Zhao, Chabes et al. 2001). During the S-phase of the cell cycle, the Sml1 level decreases along with the appearance of phosphorylated Sml1. The removal of Sml1 during S phase requires *MEC1* and *RAD53*, and the failure to remove Sml1 in *mec1* and *rad53* mutants[11] causes defective mitochondrial DNA propagation, reduced dNTP levels, and cell death. *mec1* and *rad53* mutants also show incomplete DNA replication possibly due to the failure to remove Sml1. The phosphorylation and removal of Sml1 occurs when the cells are exposed to DNA damaging and replication block inducing agents such as MMS, γ-rays, UV, and HU. Phosphorylation and removal of Sml1 in response to these agents requires *MEC1* and *RAD53*, but it is additionally dependent on other genes encoding components upstream of Mec1 or Rad53 in DNA damage response pathways.

Rothstein and co-workers (Zhao and Rothstein 2002) demonstrated that the removal of Sml1 during S phase and in the presence of DNA damage is triggered through its phosphorylation by the Dun1 kinase. Evidence for this came from the following observations:

1. Deletion of the *SML1* gene suppresses several phenotypes of the *DUN1* null mutant, including its prolonged S-phase.
2. In the null mutant of *DUN1*, both phosphorylation and degradation of Sml1 in response to DNA damage are significantly diminished.

[11] *mec1* and *rad53* mutations caused lethality. In this study, Rnr1 was over-expressed to maintain *mec1* and *rad53* mutants strains.

3. Sml1 and Dun1 physically interact *in vivo*.

4. Sml1 is phosphorylated by Dun1 *in vitro*.

These observations, combined with the previous observation that showed activation of the Dun1 kinase by Mec1 and Rad53, support the following model (Figure 4). In the presence of DNA damage or during S phase, Rad53 activates the Dun1 kinase, and the Dun1 kinase phosphorylates Sml1, leading to its degradation. Nevertheless, Dun1 may not be solely responsible for phosphorylation of Sml1 because Sml1 is weakly phosphorylated in the presence of DNA damage in *Dun1Δ* strain. It is possible that there are other kinases which have a partially redundant role with Dun1.

In *S. cerevisiae*, Dun1 is a particularly important serine/threonine kinase in that it acts upstream of multiple pathways such as transcriptional activation of RNR genes (Zhou and Elledge 1993) and the cell cycle arrest at G2/M phase (Pati, Keller et al. 1997; Gardner, Putnam et al. 1999; Sanchez, Bachant et al. 1999). Furthermore, over 10 different proteins physically interact with Dun1, suggesting that Dun1 may be involved in multiple pathways (Ho, Gruhler et al. 2002). Although the mechanism by which Dun1 transmits signals to its down stream effectors during cell cycle checkpoints is not well understood, its kinase activity is crucial for its biological function. For instance, the kinase-deficient mutants D328A and K229R, unlike wild type Dun1, cannot respond to DNA damage and do not induce the expression of the *RNR3* gene (Zhou and Elledge 1993). In addition to the regulation of Sml1, the phosphorylation of cell cycle checkpoint proteins such as the Crt1 repressor (Huang and Elledge 1997) and the DNA repair protein

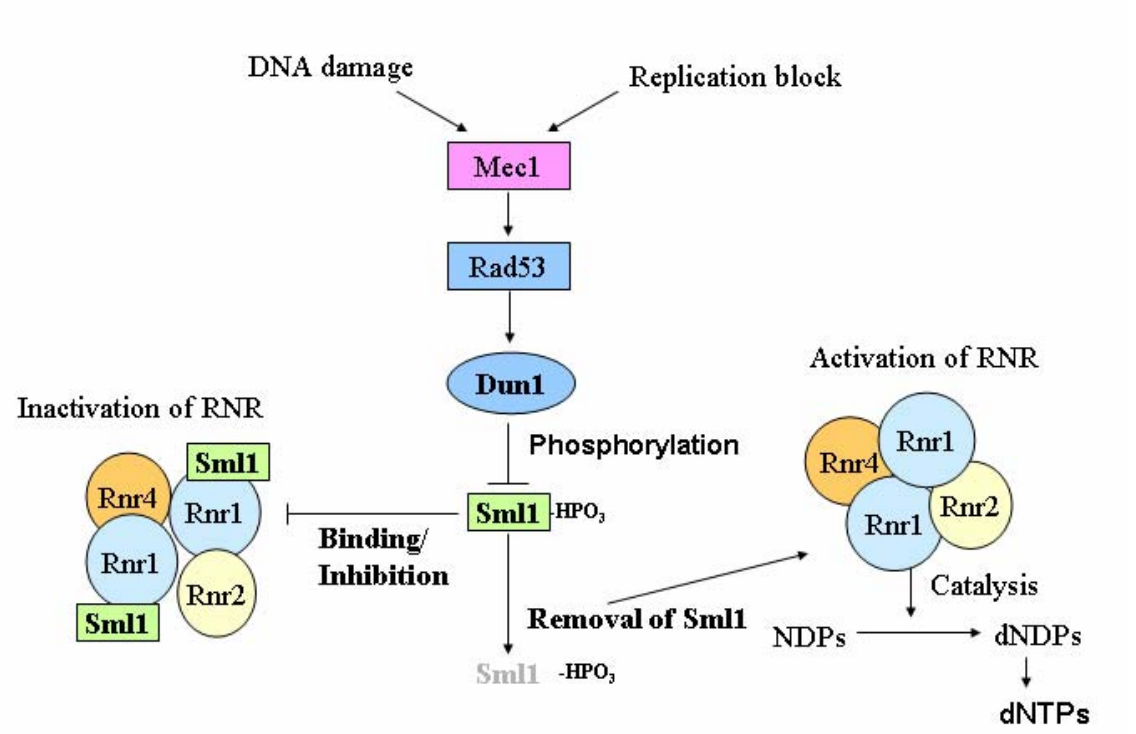

Figure 4. Pathway in *S.cerevisiae* that activates RNR through Sml1 degradation upon DNA damage and replication block.

A line with perpendicular end represents inhibition of the components or events. Arrowheads represent either activation or recognition that results in activation of downstream. The diagram shows removal of RNR inhibitor Sml1 after its phosphorylation by Dun1. Other pathways of RNR activation such as transcriptional activation by Crt1 and nuclear export of the RNR small subunit are omitted.

Rad55 (Bashkirov, King et al. 2000; Bashkirov, Bashkirova et al. 2003) depends on Dun1 via the pathways independent of Sml1. Therefore, it is likely that there are other unidentified substrates of Dun1. So far, Sml1 is the only known natural substrate of the Dun1 kinase. Prior to this study, the sites of Sml1 phosphorylated by the Dun1 kinase were unknown. By using an *in vitro* screen with a combinatorial library of 70 synthetic peptides that were known to be substrates of other serine/threonine kinases, Elledge and co-workers (Sanchez, Zhou et al. 1997) showed that Dun1 phosphopeptides have a basic residue at the –3 position from the phospho Ser/Thr site. Furthermore, they found that Dun1 phosphorylates the consensus cAPK recognition sequence, and that Dun1 and cAPK have similar substrate specificity. Nevertheless, in order to determine the Dun1 recognition sequence, it is necessary to identify phosphorylation sites in a natural substrate of Dun1.

An intriguing question is how phosphorylation of Sml1 causes degradation of Sml1. To date, very little is known about the degradation mechanism of Sml1. During the unperturbed cell cycle, the half-life of Sml1 in the cell is approximately 16 minutes during G2/M phase, in which the cellular Sml1 level peaks (Zhao, Chabes et al. 2001). In unperturbed cells, over-expression of Rnr1 stabilizes Sml1 (Zhao, Chabes et al. 2001). This is due to interaction between Sml1 and Rnr1, since overexpression of Rnr1 does not affect the stability of Sml1 mutants which do not interact with Rnr1. These observations indicate that Sml1 is a short-lived protein and unbound Sml1 is more prone to degradation. A possible mechanism of Sml1 degradation upon its phosphorylation is that phosphorylation might dissociate Sml1 from Rnr1 making it more prone to degradation.

Generally, proteins that adopt a partially folded structure are prone to a rapid degradation in the cell (Wright and Dyson 1999; Goldberg 2003; Imai, Yashiroda et al. 2003). The short half-life of Sml1 in the cell is probably due to its loosely folded structure. Although it is not known, Sml1 may adopt more rigid structure when it is bound to Rnr1, and this may explain the stabilization of Sml1 by overexpression of Sml1 *in vivo*. Some proteins that are loosely folded adopt more rigid structure when they bind to other proteins (Kriwacki, Hengst et al. 1996; Kim, Kakalis et al. 2000; Dyson and Wright 2002). In addition, some proteins such as the cell cycle inhibitory protein $p21^{Cip1}$ dissociate from other proteins when they are phosphorylated (Rossig, Jadidi et al. 2001). Another possible scenario of Sml1 degradation is that there are some cellular components for protein degradation that specifically recognize phosphorylated Sml1 as a target. For example, like Sml1, the G1/S cyclin-CDK inhibitor Sic1 is phosphorylated and degraded during the G1-S transition of the cell cycle (Feldman, Correll et al. 1997). In this process the F box protein of the SCF ubiquitin ligase Cdc4 specifically recognizes phosphorylated Sic1 for its ubiquitination (Nash, Tang et al. 2001). These two models are not completely exclusive and need to be further addressed.

Impact of this dissertation research

While many studies have revealed the pathways of the DNA damage and replication block responses, the biochemical events taking place in these pathway have not been studied in detail. This dissertation research is one of the few studies to unequivocally identify the phosphorylation sites of a protein involved in the DNA damage and

replication block responses. Although the Dun1 kinase is involved in several processes such as transcriptional regulation of RNR genes and regulation of G2/M cell cycle arrest, Sml1 is the only natural substrate of Dun1 identified to date. The major focus of this dissertation study is to identify the phosphorylation sites on Sml1 and to investigate the mechanism by which Dun1 phosphorylates Sml1. The data reported in this thesis provide an idea of the substrate specificity of the Dun1 kinase, which might be applied to study other substrates of the Dun1 kinase which are yet to be identified. In addition, knowledge of phosphorylation sites on Sml1 will be useful for further analyzing how Sml1 is regulated. For example, such knowledge can be used to generate antibodies specific to phospho-Sml1 to quantitatively analyze the relationship between phosphorylation and cellular levels of Sml1.

In the DNA damage and replication block responses, phosphorylation is a key for regulation of cellular processes. We have addressed the effect of Rnr1 on phosphorylation of Sml1. While many proteins that become phosphorylated in the cell interact with other non-kinase like proteins, there are relatively small numbers of studies that have shown whether such interaction influences phosphorylation of the target protein. We found that Rnr1 can reduce the degree of Sml1 phosphorylation *in vitro*. This finding indicates a possibility that protein-protein interactions can potentially influence rates of phosphorylation of proteins in general.

Finally, we addressed how phosphorylation of Sml1 affects the ability of Sml1 to inhibit RNR. The data suggest that phosphorylation of Sml1 weakens RNR inhibition by Sml1. However, further experiments are necessary to conclude the issue. In addition, it is

also important to test whether binding of Sml1 to Rnr1 is affected by phosphorylation. Although this issue was not addressed experimentally in depth, potential experiments to test binding of phosphorylated Sml1 to Rnr1 will be described in Chapter 6.

Chapter 2. Experimental design

Introduction

Overall experimental design of this research consists of two general types of experiments: 1) characterization of Sml1 and its modifications by mass spectrometry and 2) measurement of the properties of Sml1, Rnr1 and Dun1 in solution by conventional biochemical techniques. Considering the limitation in the amount of proteins available for these studies, mass spectrometry was the best approach to probe chemical modification (e.g. phosphorylation) of Sml1. Mass spectrometry was also used to verify the protein species used in this research. After verification and characterization of proteins by mass spectrometry, the effects of the modifications on biochemical activity / properties of the proteins were investigated.

In this chapter, the experimental design of this dissertation research will be discussed. First, protein expression and purification procedures will be described. It should be noted that all experiments in this research started with expression and purification of proteins, and the importance of these procedures cannot be underscored. Second, mass spectrometric techniques will be discussed. Mass spectrometry, specifically Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), was the major technique indispensable for this research. The rationale and advantage of using FTICR-MS, as well as the basic principles of FTICR-MS and the experimental procedures involving FTICR-MS will be discussed. In addition, since Sml1 phosphorylation is the major topic of this thesis, the background for mass spectrometric techniques used for studying protein phosphorylation is provided. Third, assays to

measure biochemical activity of RNR, Sml1 and the Dun1 kinase will be described along with the rationale behind these procedures. Fourth, a fluorescence based method to monitor the Sml1-Rnr1 interactions and gel filtration chromatography to estimate size of Sml1 oligomer will be described.

The data interpretation procedures required for the RNR activity assay, the Dun1 kinase assay, mass spectrometric identification of Sml1 phosphorylation site and the fluorescence based assay of the Sml1-Rnr1 interactions are described in the appendix of this thesis.

Preparation of proteins

Expression plasmid and yeast strain

Wild type Sml1 expression plasmid (pWJ-750-2), Rnr1 expression (pWJ751-3) plasmid, GST-Dun1 expression plasmid (pWJ772-11)(Zhao and Rothstein 2002) and the *S.cerevisiae* strain U952-B (MATa, sml1Δ::HIS3. RAD5 in W303) (Zhao, Muller et al. 1998) were kindly provided by Dr. Rodney Rothstein at Columbia University (New York, NY). The Sml1-histag expression plasmid was constructed by inserting the Sml1 ORF into pQE60 vector (Qiagene, Valencia, CA). Expression plasmids of Hisx6Rnr2 (p$(His)_6$-Y2) and Rnr4 (pY4J) (Ge, Perlstein et al. 2001) were kindly provided by Dr. JoAnne Stubbe at MIT (Boston, MA). For constructing the Sml1 mutants, the wild type Sml1 expression plasmid was used as a template to create various Sml1 mutants using the Quick Change Site-directed Mutagenesis Kit (Stratagene, La Jolla, CA). Detailed descrioption for construction of the Sml1 mutants can be found in (Dice 2003).

Preparation of Sml1

Expression and purification of Sml1-histag

E.coli M15 cell strain was transformed with the Sml1-histag expression plasmid. To express Sml1-histag, the cells were grown in 1L of TB media at 37°C to an optical density of 0.6 at 600nm, followed by induction with 1mM of isopropyl-1-thio-β-D-galactopyranoside (IPTG) for 3 hours. The bacterial cells then were harvested by centrifugation and the resulting cell pellet was stored at –20°C. Each purification procedure was performed at 4°C. Cells were re-suspended in 20mL of buffer A (50mM Tris pH 8.0, 300mM NaCl, 5mM imidazole, 5mM 2-mercaptoethanol and 1mM alpha-toluenesulfonyl fluoride) and lysed by passage through a French Press. The cell lysate was centrifuged at 10,000 x g for 30 minutes and the supernatant was incubated with 1 to 2ml of Ni/NTA agarose (Qiagen) for 3 hours on a gently moving orbital shaker. The resin was packed into a 0.5 x 2 cm column and washed with 15 bed volumes of the buffer A containing 20mM of imidazole. The protein was eluted from the resin using the buffer A with 1M imidazole added. The protein concentration in the eluate was determined by Bradford assay, using the Commassie® Plus Protein Assay Reagent Kit (PIERCE Rockford, IL). The purity of the protein sample was examined with SDS-PAGE (15% polyacrylamide gel). The bands on the gel were analyzed by an image acquisition and analysis software, Lab Works (UVP, Upland, CA). The purity of Sml1-histag in the sample was calculated based on the image optical density (IOD) of all the bands in the lane of the gel. The IOD of the band corresponding to Sml1-histag was plotted against

the amount of the total protein (3-20 μg) loaded in the lane. The range from 3 to 15 μg of total protein loaded onto the gel gave a linear plot with R^2 of 0.97, and was used to calculate the average percentage of the Sml1-histag protein in the sample. To prepare the protein for either the enzymatic digestion or the mass spectrometry experiments, the protein sample was dialyzed with a membrane of MWCO 3500 (Fisher Scientific, Pittsburgh, PA) against 16 volumes of pure water five times. No attempt was made to quantify the amount of protein recovered from the dialysis procedure; however, based on other dialysis experiments in our laboratory on a variety of proteins with similar molecular masses and amino acid sequences, we estimate a recovery of about 80-90% of the protein from the dialysis chamber. In all cases, the amount of protein recovered was sufficient for the *qualitative* mass spectrometry experiments described in Chapter 3.

Expression and purification of wild type and mutant Sml1

An overnight culture of *E.coli* BL21(DE3)pLys cells transformed with the Sml1 expression plasmid was grown in Terrific Broth media (TB) containing 100 mg/liter of ampicillin and 34 mg/liter of chloramphenicol at 37ºC. The culture was diluted 100-fold in fresh TB, and grown to an OD_{600} of 0.6. Expression of the protein was induced by the addition of IPTG to a final concentration of 0.5mM followed by three hours of incubation in the shaker at 37ºC. The cells were harvested by centrifugation and re-suspended in buffer B (50 mM Tris-HCl, 1 mM EDTA, 5 mM dithiothreitol (DTT), containing 1tablet/50ml COMPLETE™ protease inhibitor cocktail (Roche Molecular Biochemicals, Mannheim, Germany) pH 7.4) and were frozen in liquid nitrogen.

The following procedures were all carried out at 4°C. After thawing the frozen cell suspension, the lysate was centrifuged at 150,000 x g for 1 hour. The protein in the supernatant fraction was precipitated by addition of ammonium sulfate to 25% saturation. The precipitate was harvested by centrifugation at 12,000 x g for 30 minutes and dissolved in buffer B before applying to a HiLoad 26/60 Superdex 75 gel filtration chromatography column (26 mm I.D. and 60 cm height. Amersham Biosciences, Uppsala, Sweden), which was pre-equilibrated with buffer B or buffer C (50mM Hepes-KOH, 5mM $MgCl_2$, 5% (v/v) glycerol, 5mM DTT pH 7.0) containing 100mM KCl[12]. The eluted fractions containing Sml1 were identified by SDS-PAGE using a 15% polyacrylamide gel followed by ESI-FTICR-MS as described below. These fractions were pooled and used for phosphorylation assays. The concentration of Sml1 was determined by Coomassie Protein Assay Kit (PIERCE, Rockford, IL) using bovine serum albumin (BSA) as a standard.

Preparation of Rnr1

Preparation of Rnr4 C terminus 9mer conjugation for Rnr1 purification

Studies reported that peptides with the sequence consisting of C terminus sequence of the RNR small subunit can specifically binds to the RNR large subunit and inhibit RNR activity (Climent, Sjoberg et al. 1991; Filatov, Ingemarson et al. 1992; Davis, Thelander et al. 1994). Exploiting this property of such peptides as an advantage, Cooperman and

[12] For analyzing oligomeric state by gel filtration (Chapter 4), buffer B was used. For studying Sml1 phosphorylation, buffer C was used.

co-workers (Yang, Spanevello et al. 1990) reported that murine ribonucleotide reductase large subunit (R1) can be effectively purified by resin coupled to the peptide consisting of the C-terminal 9 amino acid residues of murine ribonucleotide small subunit (R2). In *S.cerevisiae* peptides consisting of the C-terminal 9 amino acid residues of Rnr2, Rnr4 and Sml1 can inhibit RNR activity (Chabes, Domkin et al. 1999). Among these three, the peptide consisting of the Rnr4 sequence has higher efficacy than the peptides corresponding to the Rnr2 or Sml1 sequences (Chabes, Domkin et al. 1999).

Slightly modifying the procedure reported by Cooperman and co-workers (Yang, Spanevello et al. 1990), we utilized a resin coupled with a peptide consisting of C terminal residues of a small subunit of yeast ribonucleotide reductase to purify Rnr1.

Expression and purification of Rnr1

An overnight culture of *E.coli* BL21(DE3)pLys cells transformed with the Rnr1 expression plasmid was grown in Terrific Broth media (TB) containing 100 mg/liter of ampicillin and 34 mg/liter of chloramphenicol at 37ºC. The culture was diluted 100-fold in fresh TB, and grown to an OD_{600} of 0.6. After chilled on ice water for 15 minutes, expression of the protein was induced by the addition of IPTG to a final concentration of 0.5mM followed by incubation for 16 to 20 hours with shaking at 15°C. The cells were harvested by centrifugation and stored at -80°C.

The following procedures were all carried out at 4°C. After being thawed, 10g cells were re-suspended in buffer C containing 100mM KCl, 1mM phenylmethylsulfonyl fluoride (PMSF), 1tablet/50ml COMPLETE™ protease inhibitor cocktail, 2μM

pepstatin pH 7.0 and lysed by one passage through a French pressure cell (12,000 psi; 1 psi = 6.89 kPa). Cell debris was removed by centrifugation (12,300 × *g*, 30 min). Streptomycin sulfate (10% wt/vol) was added gradually to the extract to a final concentration of 1.5% wt/vol, and the DNA was removed by centrifugation (12,300 × *g*, 30 min). Solid $(NH_4)_2SO_4$ was added to the supernatant to 40% saturation (0.29 g/ml) over a period of 30 min, and precipitated proteins were collected by centrifugation (12,300 × *g*, 40 min). The pellet was dissolved in 2 to 2.5 ml of buffer C containing 1mM PMSF, 1tablet/50ml COMPLETE™ protease inhibitor cocktail, 2μM pepstatin pH 7.0 and desalted by passage through a PD10 column (Amersham Bioscience, Uppsala, Sweden) with buffer C as eluent. COMPLETE™ protease inhibitor cocktail (1 tablet/ml stock) was added to the sample to a final concentration of 1tablet / 50ml. This solution was incubated with 2ml of Y4 9mer peptide coupled resin (described above) for one hour at 4°C on a gently moving rocker. After the incubation, the resin was packed in a column (1cm I.D. Econo-column: BioRad, Hercules, CA) and manually washed with 10ml of buffer C. Using a flow adapter (1cm plunger diameter: BioRad), the column was washed with 15ml of buffer C followed by KCl gradient from 100mM to 200mM in buffer C over 60ml elution volume at a flow rate of 1ml/min. The bound protein was then eluted with 20 to 30ml of 1M KCl in buffer C. Protein-containing fractions were pooled, concentrated by ultrafiltration with a YM30 membrane (Millipore, Bedford, MA) once and the KCl concentration was adjusted to 100mM by dilution in buffer C. The diluted sample was concentrated again to >5mg/ml and frozen in liquid nitrogen, and stored at

−80°C. Typically ~5–10 mg of Rnr1 with ~95% purity (based on densitometry analysis of Coomassie Blue stained SDS/PAGE gel) was obtained.

Expression and purification of histag Rnr2/ Rnr4

E. coli BL21 CodonPlus (DE3) RIL cells were co-transformed by the expression plasmid of Hisx6Rnr2 and Rnr4. The overnight culture was grown in Terrific Broth media (TB) containing 100 mg/liter of ampicillin and 25 mg/liter of Kanamycin at 37ºC. The culture was diluted 100-fold in fresh TB and grown to an OD_{600} of 0.6. After chilled on ice water for 15 minutes, expression of the protein was induced by the addition of IPTG to a final concentration of 0.5mM followed by 16 to 20 hours shaking at 15°C. The cells were harvested by centrifugation and stored in -80°C.

Purification of Hisx6Rnr2/Rnr4 complex was carried out following the procedure described by Thelander and co-workers (Chabes, Domkin et al. 2000). Frozen bacterial cell paste obtained from 2L cell culture was re-suspended in approximately 4 wet cell volumes of buffer D (50 mM Hepes/5% glycerol/1.0 mM PMSF, 1x COMPLETE™ protease inhibitor cocktail, pH 7.4 at 4°C). The cells were lysed by passage through a French pressure cell (15,000 psi; 1 psi = 6.89 kPa), and cell debris was removed by centrifugation at 19800 x g for 30 minutes at 4°C. Fe(II) ammonium persulfate was added to the supernatant to a final concentration of 170 μM. The supernatant was then incubated with 1ml bed volume TALON cobalt affinity resin at 4˚C for 1 hour on a gently rotating shaker. The column was subsequently washed with 40 column volumes (CV) of buffer D containing 100 mM NaCl, followed by 16 CV of buffer D containing 10

mM imidazole. The protein was eluted first with 10 CV of buffer D containing 100 mM imidazole, then with 10 CV of buffer D containing 200 mM imidazole. Fractions containing Hisx6 Rnr2 and Rnr4 were identified by 12% SDS-PAGE, and they were combined. The imidazole concentration was reduced to less than 1mM by repeated ultrafiltration with a YM30 membrane (Millipore, Bedford, MA) or dialysis. Proteins were concentrated to >5 mg/ml, frozen in liquid nitrogen, and stored at −80°C.

Expression of GST-Dun1 and preparation of GST-Dun1 bound resin

Expression of GST-Dun1 was carried out as described by Rothstein and co-workers (Zhao and Rothstein 2002). Yeast cells (U952-B) transformed with the GST-Dun1 expression plasmid (pWJ772-11) were grown in SC-URA raffinose, and the expression of Dun1 was induced at mid-log phase (5~6 x 10^6 cells /ml) by addition of solid galactose to a final concentration of 2% (w/v). After two to three doubling times (3 to 3.5 x 10^7 cells /ml), the cells were harvested, washed with PBS, and stored at –80ºC. For the experiment involving induction of DNA damage response, methyl methane sulfonate (MMS) was added to a final concentration of 0.1% when cell density reached ~1.5 x 10^7cells / ml, and cells were grown for two more hours. To compare the activity of GST-Dun1 from the cells treated with and without MMS, a fraction of cells were set aside before addition of MMS, and these cells were grown without MMS for two more hours. The purification of GST-Dun1 was carried out as described by Zhao and Rothstein and Elledge and co-workers (Sanchez, Zhou et al. 1997; Zhao and Rothstein 2002). Cells were re-suspended in ice cold NP40 buffer (50 mM Tris-HCl,150 mM NaCl, 50 mM

KCl, 5 mM $MgCl_2$, 1% (v/v) Igepal 630, 10% (v/v) glycerol, 10 mM Dithiothreitol, 0.1 mM $NaVO_4$, 30 mM NaF, 1 mM PMSF, 1x COMPLETE^{TM} protease inhibitor cocktail, 2 μM pepstatin pH 8.0) to a cell density of approximately 3~4 x 10^9 cells /ml. An equal volume of glass beads (0.45~0.50mm particle size) was added to the cell suspension. The cells and glass beads are agitated by vigorously vortexing for 30 second and then placed on ice water for 30 second to cool. The cycles of agitation and cooling were repeated seven times, and the sample was centrifuged for 30 seconds. The soluble fraction was taken into another tube and set aside. The same volume of NP40 buffer was added to the mixture of cells and glass beads again and the cycles of agitation and cooling were repeated three times. After centrifuged for a 30 seconds, the soluble fraction was pooled. This procedure was repeated once more. The soluble fraction was further centrifuged at 12,000 x g for 20 minutes at 4ºC. Total amount of protein in the supernatant was estimated based on concentration as determined by the Coomassie Plus Protein Assay. For every 1mg of total protein in the supernatant, 10μl of glutathione resins (Glutathione-4 Sepharose Superflow: Amersham Bioscience, Piscataway, NJ) was added and allowed to incubate with the protein for one hour at 4ºC on a gently moving rocker. After incubation, the resin was packed into a column (1cm I.D. Econo-column: BioRad, Hercules, CA) and washed with 50 bed volumes of 50mM Tris-HCl, 500mM NaCl, 1mM EDTA, 1mM DTT pH 7.5. Then, the column was equilibrated with 50mM Hepes-KOH, 10% glycerol, 5mM $MgCl_2$, 5mM DTT pH 7.0. The resin was suspended in 3 bed volumes of the same buffer as 25% slurry and stored in –80ºC. Examination of the glutathione-beads by SDS-PAGE revealed pure GST-Dun1.

Mass spectrometry

Rationale for using Fourier Transform Ion Cyclotron Resonance mass spectrometry (FTICR-MS) for the characterization of Sml1

Mass spectrometry has become an indispensable technique useful for characterization of protein and peptides for biological research (Mann and Pandey 2001). When studying the biochemical properties of a protein, it is important to confirm the identity and purity of a protein species and ensure the integrity of the proteins in the sample. By combining specific proteolytic digestion and protein sequence database searches, mass spectrometry has become one of the best methods for protein identification for both purified proteins and mixtures of several proteins. Furthermore, mass spectrometry provides information concerning presence of protein variants and post-translational modifications.

Significant progress has made in both the instrumentation and application of several types of mass spectrometers for identification and characterization of biological macromolecules (Mann and Pandey 2001). Nonetheless, each mass spectrometer has its own strengths and weaknesses. FTICR-MS provides mass resolution (FWHM of 100,000 to 150,000) far superior to other types of instruments and also provides high mass accuracy (1 to 10 ppm for molecules of 100 to 30,000Da) with proper calibration (Marshall, Hendrickson et al. 1998; Hendrickson and Emmett 1999). In addition, its ability to comprehensively measure a wide dynamic range (up to 10^5)[13] minimizes the

[13] Dynamic range is difference in abundance of two or more species. It is defined as intensity ratio of the base peak (the peak of the most abundant species) to the smallest peak that is comprehensively detected.

requirement of extensive sample separation, making it possible to analyze relatively complex mixtures of peptides and proteins. Furthermore, FTICR-MS has capability of tandem mass spectrometry to probe ion fragmentation.

The high performance that can be achieved only by FTICR-MS was particularly crucial for analyzing intact Sml1 and its modified forms. Because of multiple carbon atoms in the molecule, the molecular region of this protein exists as a population of numerous isotopic species. The mass of a protein is determined most accurately if different isotopic species are resolved. Even for a protein with the size of Sml1 (12kDa), FTICR-MS is only the instrument that can comprehensively resolve all of these isotopic species. Resolution of isotopic species is even more important when analyzing modified proteins. For example, as described in Chapter 3, Sml1 forms a disulfide-linked dimer in non-reducing conditions. The disulfide-linked dimer of Sml1 has mass 2.0156Da less than twice the mass of Sml1 due to loss of two hydrogen atoms. Such mass difference of *intact proteins* can be probed only when isotopic species are comprehensively resolved. Due to the relatively low complexity of the Sml1 sample, LC separation was not absolutely necessary for analyzing Sml1. Therefore, in every experiment, electrospray ionization (ESI) was carried out in direct infusion mode. Nevertheless, direct infusion ESI still provided high quality data and sufficient information content to achieve the goal of characterizing Sml1 and its phosphorylation.

In the next three sections, the fundamental principles of FTICR-MS will be illustrated. A more detailed description of FTICR-MS can be found in (Marshall, Hendrickson et al. 1998; Hendrickson and Emmett 1999).

Basic principles of FTICR

Detection and mass determination of ions by FTICR-MS

FTICR-MS relies on the cyclotron motion of ions in a magnetic field. When an ion moves at a velocity ***v*** in the presence of a spatially uniform magnetic field, **B**=B_0k (**B** is a vector with amplitude B_0 and a direction defined by k parallel to Z axis[14]), the ion is subjected to a force defined as

$$Force = mass \bullet accerelation = m\frac{dv}{dt} = qv \times B \qquad \text{(Equation 1)}$$

In this equation, *m* and *q* are ionic mass and charge, respectively. Note that ***v*** and **B** are vectors, and their cross product is perpendicular to ***v*** and **B** (Figure 5). Equation 1 shows that an ion with velocity perpendicular to magnetic field experiences an inward directed force, which is called the Lorentz force. Due to the Lorentz force, if the ion maintains a constant speed, the ion path is bent into a circle (Figure 5).

Further development of Equation 1 allows expression of the angular velocity of the ion cyclotron motion about the z-axis in terms of mass, charge, and magnetic strength. For this conversion, consider only the magnitude of the force in Equation 1. Let $v_{xy} = \sqrt{v_x^2 + v_y^2}$ denote the magnitude of the ion velocity in the x-y plane (plane of the paper in Figure 5), which is perpendicular to **B**. In this case, all the vector quantities in Equation 1 should be converted into scalar quantities. First, $v \times B$ on the right side of

[14] Z-axis is direction in which ion is injected into the mass analyzer, and it is parallel to the magnetic field, **B**. Please see Figure 6B.

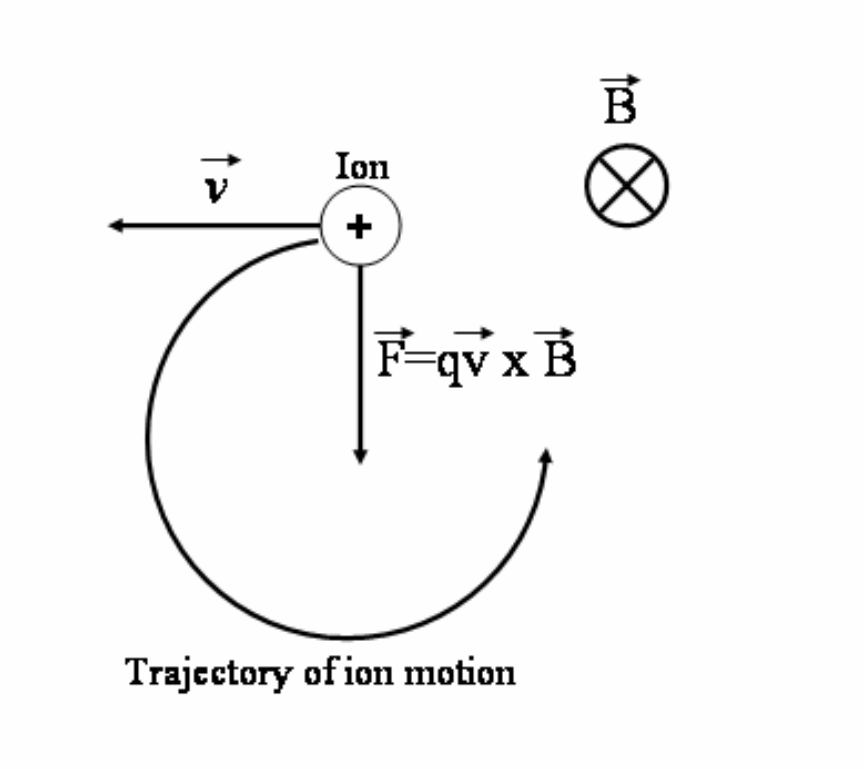

Figure 5. Ion cyclotron motion.

In this diagram, magnetic field **B** is perpendicular to plane of the paper. For an ion moving at a velocity v, it experiences a force perpendicular to ***v*** and **B**. Then motion of ion is bent into a circle.

Equation 1 should be converted into $|v \times B|$, and that is equal to product of v_{xy} and B_0. Second, d***v***/dt in equation 1 should be converted into $d\boldsymbol{v}_{xy}/dt$, and that is equal to v_{xy}^2 / r where r is radius of the circular motion.[15] Then the magnitude of the force defined in equation 1 is expressed as

$$\text{Force} = mv_{xy}^2 / r = qv_{xy}B_0 \qquad \text{(Equation 2a)}$$

The angular velocity, ω (radian / sec) about z-axis is defined as

$$\omega = v_{xy} / r \qquad \text{(Equation 2b)}$$

In this particular case, ω is angular velocity of ion cyclotron and can be written as ω_c. By replacing v_{xy} with $\omega_c r$, Equation 2a can be expressed as

$$m\omega_c^2 r = qB_0\omega_c r \qquad \text{(Equation 2c)}$$

Simplifying the expression in Equation 2a;

$$\mathbf{\omega_c = qB_0 / m} \qquad \mathbf{\text{(Equation 3a)}}$$

---

[15] Generally, acceleration (dv/dt) of a circular motion can be written as v^2/r. In our case, we are concerned with only magnitude of acceleration, and v can be a scalar quantity.

Generally, the frequency of circular motion can be defined as $f = \omega/2\pi$. Likewise, the ion cyclotron frequency, f_c, is $f_c=\omega_c /2\pi$. By replacing ω_c with qB_0 / m (as in Equation 3a), f_c can be expressed in simplistic terms as below

$$\boldsymbol{f_c = \omega_c / 2\pi = qB_0 / 2\pi m = B_0 / 2\pi (m/q)} \qquad \textbf{(Equation 3b)}$$

These equations show that the frequency of the ion cyclotron is unique to a particular mass/ charge, and it is inversely proportional to mass/charge of the ion. On the other hand, the frequency of the ion cyclotron is independent of ion velocity and proportional to magnetic field strength. The frequency of the ion cyclotron motion is the quantity measured by FTICR-MS, and from which mass/charge is determined.

As evident in Equation 1, the movement of ions in the x-y plane is restricted by the magnetic field **B**. However, ions are still free to escape from the mass analyzer along the z-axis. To confine the movement of ions along the z-axis, an electrostatic potential can be applied on two plates positioned perpendicular to the z-axis ("Trapping" in Figure 6B). The application of an electrostatic potential on the trapping plates not only confines ions movement along the z-axis but also creates two additional motions of ions. The first type of motion, called trapping oscillation, is the oscillation of the ions along the z-axis. The second type of motion, called magnetron rotation, is a circular precession of ions in the x-y plane with much lower frequencies than ion cyclotron frequencies. However, the trapping oscillation and magnetron rotation are usually not detected due to their low

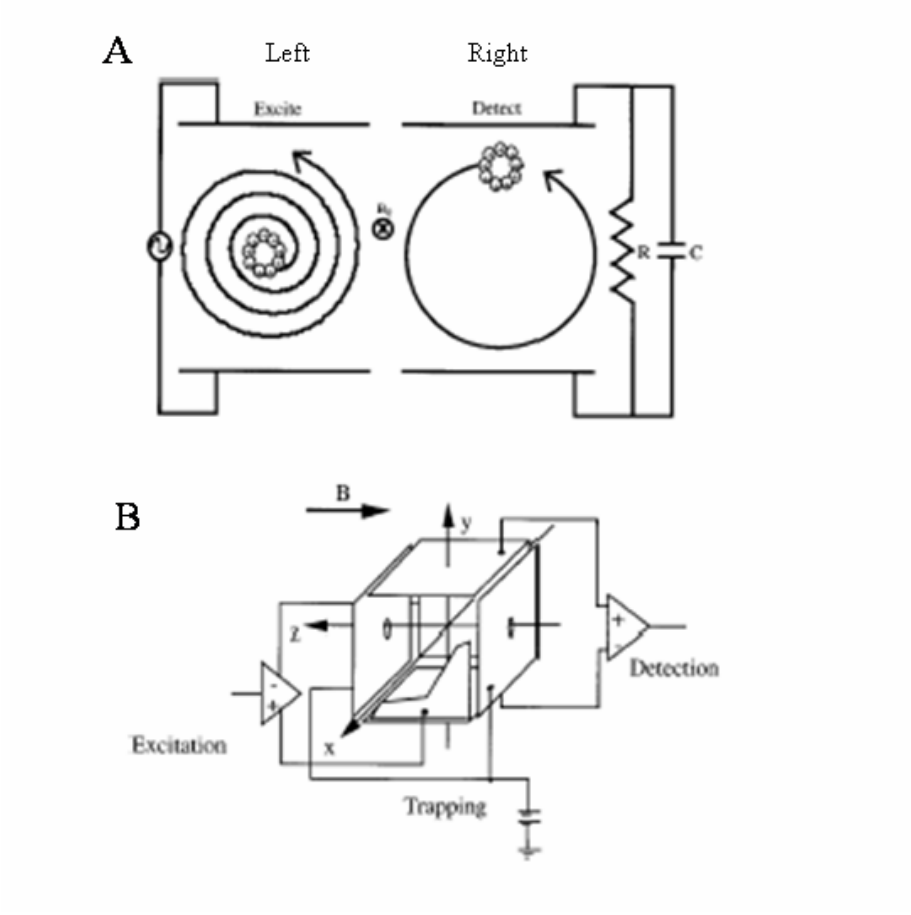

Figure 6. Mass analyzer of FTICR-MS.

(Images were taken from (Hendrickson and Emmett 1999))
(A) Ion acceleration and detection. In this picture, magnetic field, B_0, is perpendicular to the plane of the paper. Picture on the left shows application of rf to positive ions in the magnetic field to excite ICR. The picture on the right shows detection of ICR image current by detection plate. (R: resistor C: capacitor).
(B) A schematic diagram of FTICR-MS mass analyzer. The ion is injected through the small hole on the trapping plate and confined in the mass analyzer by an electric potential created between the trapping plates. Excitation, detection and trapping plates are orthogonal to each other. Excitation and detection is achieved as in (A).

frequencies, and these frequencies are also relatively independent of mass/charge. Because they do not have significant effect on excitation and detection of ions, trapping oscillation and magnetron rotation will be omitted from rest of the discussion.

Excitation and detection of ions

Ions trapped in a magnetic field generally have incoherent cyclotron motion (i.e. they are moving independent of each other). In this mode, it is impossible to detect their net motion. To force the ions to move coherently, an electric field at the appropriate frequencies oscillating along a direction perpendicular to the magnetic fields needs to be applied (Figure 6). For B=9.4 Tesla, these oscillating electric fields must be in the radio frequency (rf) range of 10 to 5000 kHz. If the rf frequencies matches the frequencies of the ion cyclotron motion, the ions continuously absorbs power, A(t), according to dot product

$$A(t) = \textit{Force} \bullet \textit{velocity} = q\mathbf{E}(t) \bullet \boldsymbol{v}_{xy} \qquad \text{(Equation 4)}$$

(q: Charge of ion, E(t): the oscillating electric field, velocity of ion on x-y plane)

While absorbing power, the ions are accelerated (Figure 6A left), and at the same time, all ions of the same mass/charge are forced to move in a phase coherent motion forming a packet of ions. The coherent ion cyclotron motion is called the *ion cyclotron resonance (ICR)*. Note that as the ions are accelerated, the frequencies of the ICR do not change (as in equation 3), but the radius of ICR increases. On the other hand, if the frequencies of rf

and ion cyclotron are different, the ions will not absorb power. In FTICR-MS, a chirp pulse of rf that covers a wide range of ICR frequency is applied to the ions confined in a magnetic field, and the ions absorb power when the applied rf has the same frequency as their ion cyclotron.

After controlled acceleration of the ions, the application of the rf pulse is terminated, but the ICR continues for a period of time until collisions of ions with neutral molecules quench the coherent motion. In the mass analyzer of FTICR-MS, the detection plates are placed orthogonal to the magnetic field (Figure 6 and Figure 7). During ICR, as the positively-charged ion packet approaches the upper detection plate (Figure 7), it induces a negative charge on the upper detection plate. After a short period of time, the ion packet moves toward the lower detection plate in Figure 7 and induces a negative charge on the lower plate. In short, ICR creates a charge differential between the two detection plates according to

$$\Delta Q = -2qy/d \qquad \text{(Equation 5)}$$

(ΔQ: charge difference between the two detection plates, q: the net charge of the ion packet, y: position of the ion packet with respect to detection plates, d: the distance between the detection palates. [See Figure 7])

It should be noted that the charge differential between the detection plates constantly changes over the period of ICR due to the coherent ion motion. Thus, the ICR creates an

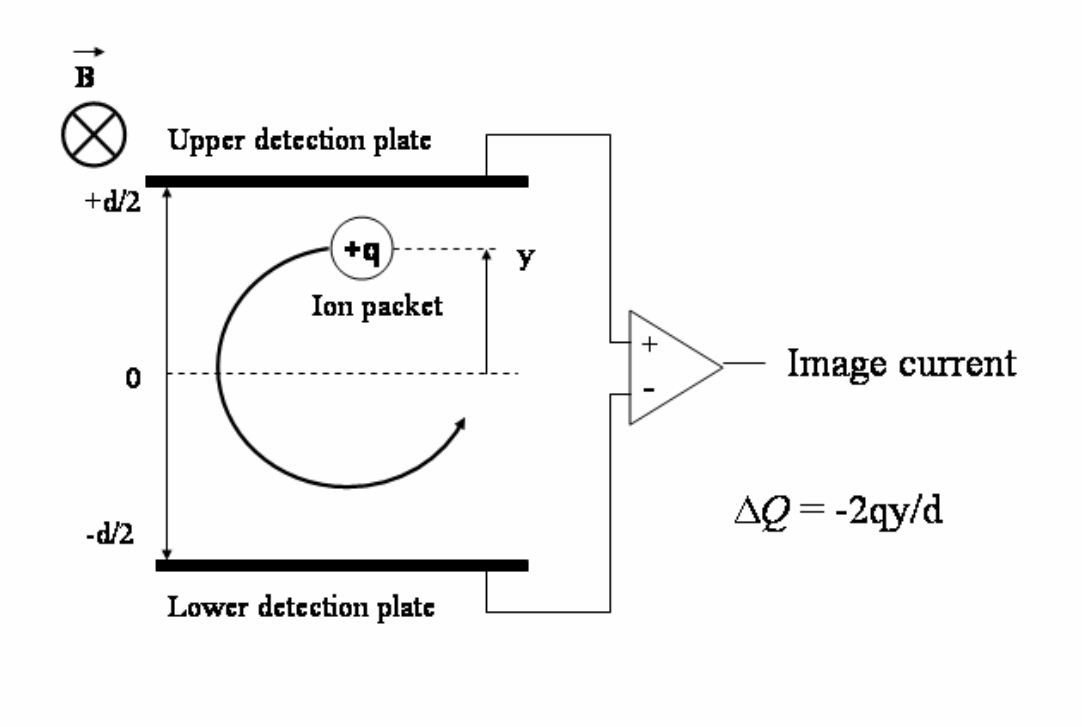

Figure 7. Schematic diagram of FTICR-MS ion detection system.

Y-axis is orthogonal to magnetic field, and the center of the mass analyzer is taken as the origin of y-axis. When the ion packet is at y=0, there is no charge difference between the detection plates. When the ion packet is closer to the upper detection plate, it crates negative charge to the upper detection plate. Thus, that creates a charge difference between upper and lower plate. Opposite charge different is created as the ion move toward the lower plate. Due to a circular motion of the ion packet during ICR, the image current of an ion packet resembles sine or cosine curve.

alternating current with ICR frequency on the detection plates, since electric current is defined as rate of change in charge difference over time ($d\Delta Q/dt$). This alternating current is usually called image current. The amplitude of the image current is proportional to the number of ions in the ion packet.

If there are only ions of single mass/charge in the mass analyzer cell, the image current will resemble a pure sine wave, and this can be expressed in the time domain as a function of voltage amplitude with respect to time (Figure 8). By a mathematical operation called Fourier transformation, the time domain image current can be converted into the frequency domain, generating a function of amplitude with respect to ICR frequency (Figure 8). In the frequency domain, amplitude is also proportional to abundance of ions trapped in the analyzer cell. As in equation 3, ICR frequency and mass/charge have an inverse relation. In other words, the mass spectrum is a mirror image of the frequency domain.

If there are ions with different (multiple) mass/charge, they create a complex waveform representing an addition of image currents from ion packets with different (multiple) mass/charge ratios (Figure 9). In order to Fourier transform this waveform, it needs to be expressed as a series of individual waveforms, called the Fourier series. In the Fourier series, the waveform is expressed as the sum of sine and cosine terms in which each sine or cosine term represents the image current from an individual ion packet. The image current expressed by the Fourier series is converted into the frequency domain, and the frequency domain is further converted to mass spectrum according to equation 3.

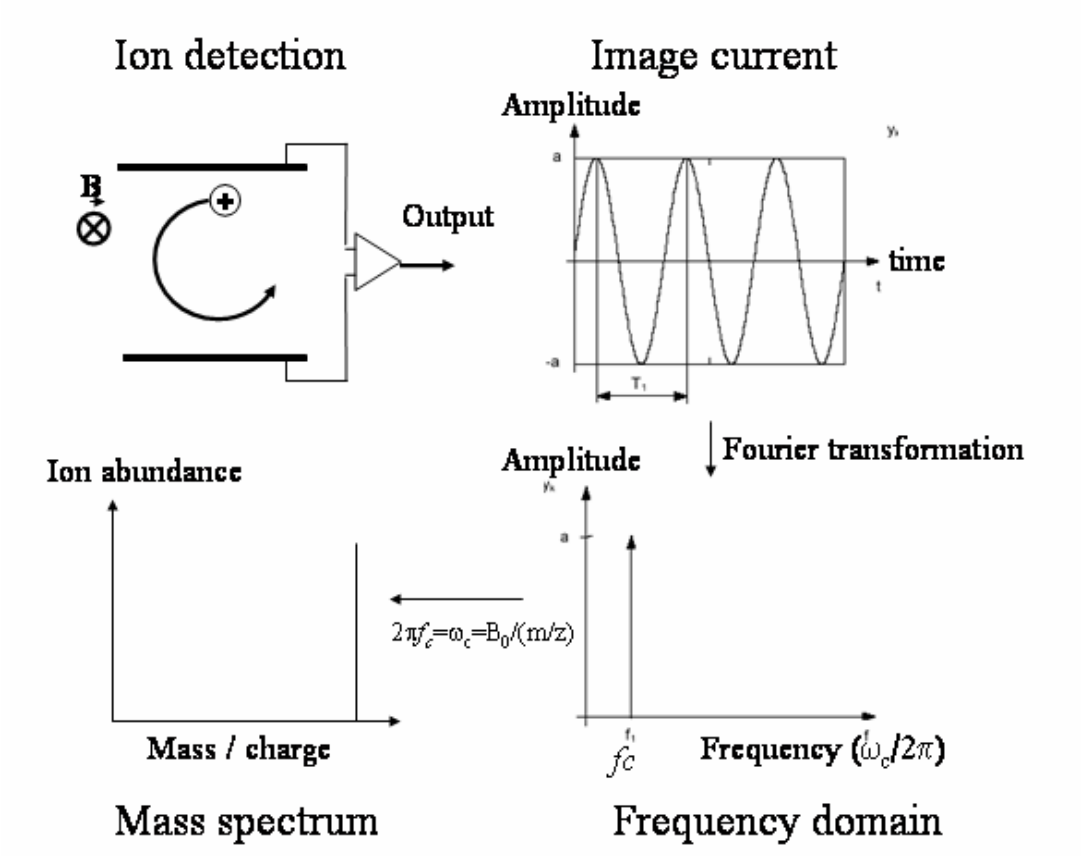

Figure 8. A simplified summary of data processing in FTICR-MS.

ICR creates image current that represents a function of time in terms of ICR frequency and amplitude (abundance) of ion packet. The image current is converted to frequency domain spectrum by Fourier transformation. The frequency domain spectrum is the function of frequency in terms of ion abundance, and it is a mirror image of mass spectrum. The frequency domain spectrum is further converted into mass spectrum based on the relation $2\pi fc=\omega_c=B_0/(m/q)$.

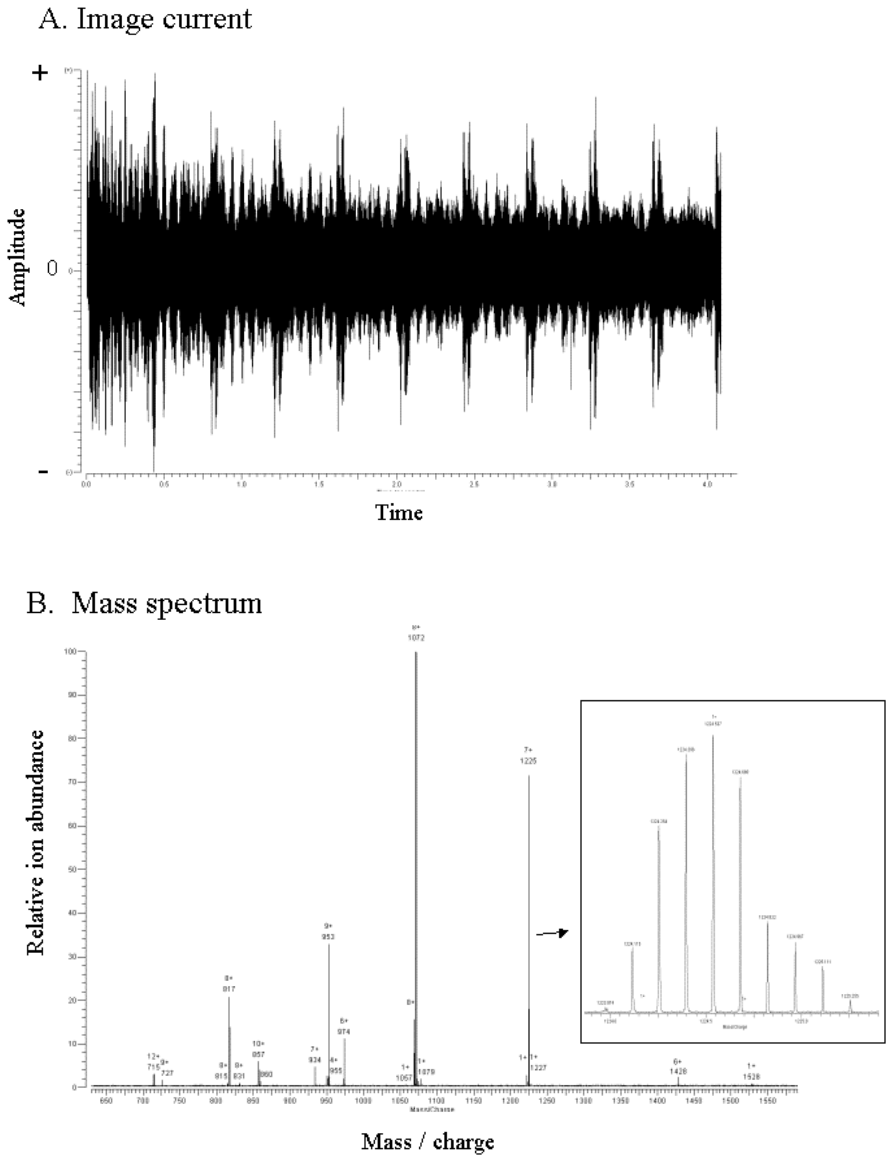

Figure 9. Image current and mass spectrum created by multiple ions of the protein ubiquitin (Mr=8565Da).

(A) Image current. ICR of more than 50 different species in different charge states and isotope compositions created a waveform representing addition of alternating currents from individual ion packets. The waveform was synthesized into a periodic function of time, f(t). (**B)** Mass spectrum. The spectrum was created by Fourier transformation of f(t) into F(ω) followed by conversion of ω to m/z. The inset is the zoom of $(M+7H)^{+7}$ charge state ions showing the different isotopic species.

Sustained off resonance irradiation (SORI) collisional activated dissociation

Fragmentation of selected ions followed by detection of fragment ions is called tandem mass spectrometry and provides rich information for determining chemical structure of the ions (Dass 2001). A common method of ion fragmentation is to increase kinetic energy of ion by accelerating it and colliding it with an inert gas molecule; this method is called collisional activated dissociation (CAD). CAD is performed by a series of sequential operations (Figure 10). First, only ions of interest are isolated in the mass spectrometer, and all the other ions are ejected from the region where the collisions will take place (Figure 10-2). These selected the ions of interest are called parent ions. In many cases, before isolating the parent ions, mass spectra are acquired to obtain information regarding the presence and the abundance of the parent ions. Second, the parent ions are accelerated by application of an appropriate electric field in the presence of inert gas such as nitrogen or helium. During this process, the parent ions pick up kinetic energy and collide with the gas molecules (Figure 10-4). As a result of consecutive collisions, the kinetic energy is converted to internal energy in the parent ions. When the ion's internal energy exceeds chemical bonding in the ions, fragmentation of ion occurs. Generally, the weaker bonds in the ion tend to be broken first. Third, the fragment ions are detected, and their mass/charge ratios are determined (Figure 10-6 and Figure 10-7).

Among different types of CAD, the most popular method of CAD by FTICR-MS, termed sustained off resonance irradiation (SORI) CAD, was employed in this research due to its versatility. Like the more conventional CAD, SORI-CAD is performed in the

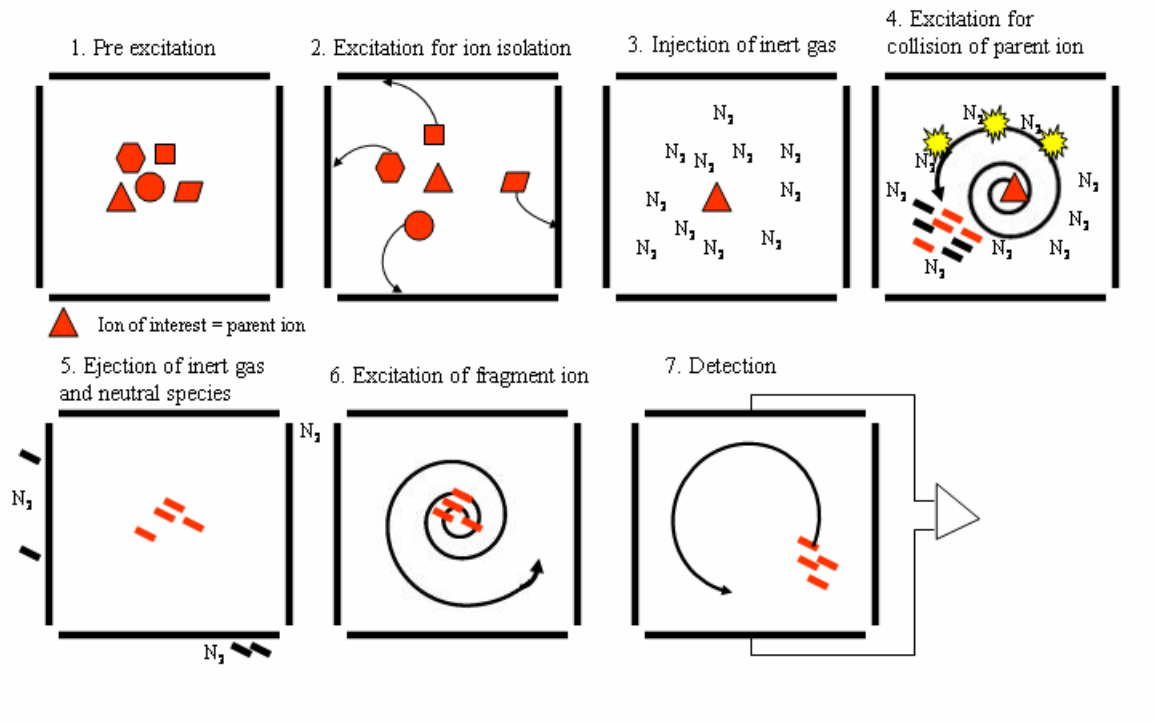

Figure 10. Collisional activated dissociation in FTICR-MS.

(1&2) Ion isolation. All the ions except the ions of interest are accelerated. The accelerated ions hit the wall of ICR cell, and they are neutralized. (**3)** Pre ion-gas collision. Inert gas (in this case, nitrogen) is injected to the ICR cell. (**4)** Ion-gas collision and ion fragmentation. Ions of interest (parent ions) are accelerated to collide into the inert gas. Kinetic energy is accumulated to the parent ions, and the ions are eventually broken into fragments. **(5)** Pre-ion detection. The inert gas and neutral fragments are ejected by the vacuum system, while the fragment ions are confined into ICR cell by the trapping electric potential and the magnetic field. (**6&7)** Ion detection. All the fragment ions are accelerated, and image current of fragment ions is acquired.

same sequence of ion isolation, ion-gas collision and ion detection (Gauthier, Trauman et al. 1991; Marshall, Hendrickson et al. 1998).

The method of ion isolation in SORI-CAD is similar to the ion detection process described in the previous section. Any ion that is accelerated above a certain threshold will hit the wall of the analyzer and become a neutral molecule (this process is called annihilation). Such neutral molecules are no longer detected. In ion isolation, all the ions *except the parent ions* are accelerated above the threshold. As a result, only the parent ion remains to be detected. For this process, first, the instrument creates a theoretical frequency domain that does *not* include the frequency for the parent ions. Second, the frequency domain is reverse Fourier transformed into a time domain. Based on the time domain, the instrument designs an rf pulse that covers a range of ICR frequencies except that of parent ion. Third, the rf is applied to the mass analyzer, and all the ions except the parent ions are annihilated (Figure 10-2). Before the ion-gas collision process, a pulse of the inert gas is injected into the mass analyzer (Figure 10-3).

In the ion-gas collision process, the parent ions are accelerated by application their cyclotron frequencies at well controlled amplitudes, so that they collide with the inert gas molecules under carefully controlled conditions. Thus, sustained off resonance irradiation (SORI) is a special type of CAD experiment. In this application, the rf frequency is slightly off from the ICR frequency of the parent ions so that the phase of ICR and rf differs slightly (~1 kHz). This partial phase matching is referred to as off-resonance. Because of off-resonance between the ICR and the applied rf, the parent ions are accelerated only temporarily even though rf is continuously applied (Figure 11). If the

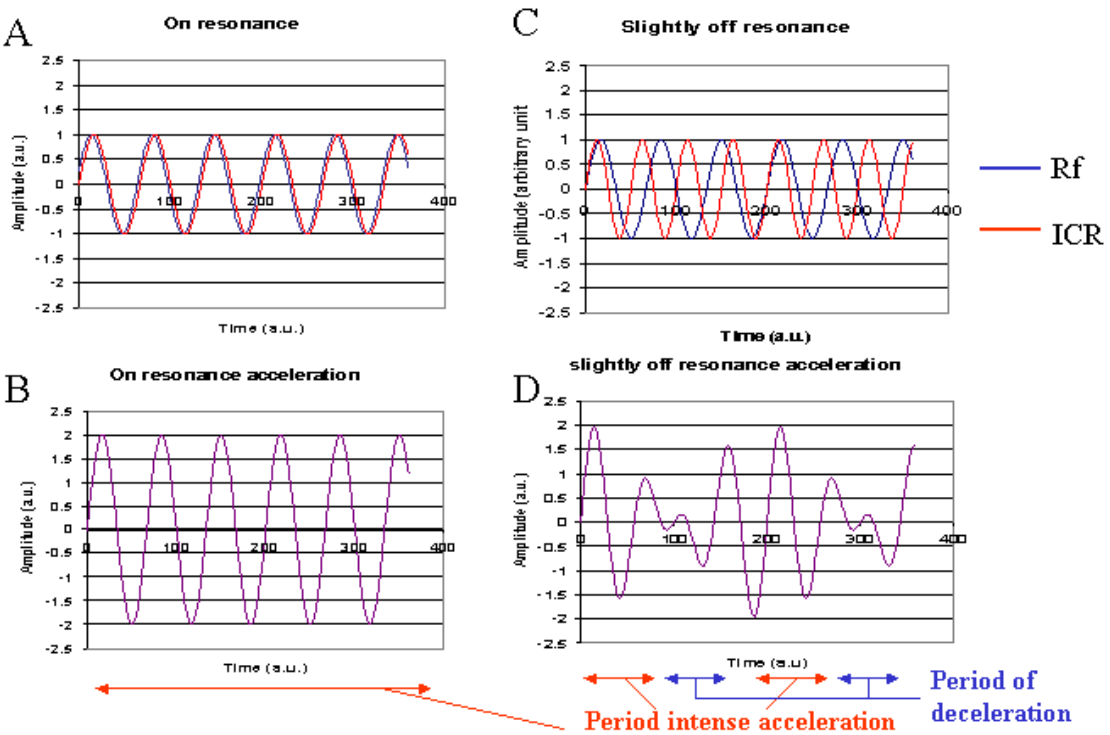

Figure 11. Conceptual waveforms of on and off resonance ion accelerations.

(In (A) and (C), blue line represents Rf and red line represents ICR.) **(A)** On-resonance acceleration. In this case, the frequencies of Rf and ICR are the same. **(B)** ICR in on-resonance acceleration. Addition of Rf and ICR waveforms in A shows that the ions are intensely accelerated throughout the period of Rf application. **(C)** Off-resonance acceleration In this case, the frequency of Rf is slightly lower than that of ICR. **(D)** ICR in Off-resonance acceleration. Addition of waveforms in C shows that the ions are intensely accelerated only in fractions of the period of Rf application.

ions were accelerated continuously, they would keep moving outward and eventually hit the wall of the analyzer. However, since the ions are accelerated only temporally, their outward movement stops whenever the phase of the applied rf does not match the ICR motion. When the phase of the applied rf is 180˚ off from the phase of ICR, ions are decelerated. Furthermore, while not being accelerated, the ions experience collisions with the gas molecules, which can cause them to lose energy or fragment. As a result of the deceleraton and energy loss, their ICR radius decreases, and the ions relax back to the center of the analyzer cell. Generally, ions are more efficiently confined in the analyzer cell when they are close to the center. In this way, the application of slightly off resonance rf minimizes potential loss of unfragmented parent ions during the collision process. After the collision process, all the ions in the mass analyzer are re-accelerated for detection as described in the previous section.

In summary, the ion acceleration processes takes place multiple times during SORI-CAD. The amplitudes and frequencies of rf in these accelerations are different, and they are designed for different processes namely ion isolation, ion-gas collision, and ion detection.

Experimental procedure for mass spectrometric analysis of protein and peptides employed in this study

ESI-FTICR-MS

All samples were desalted with C18 reverse phase ZipTips (Millipore, Bedford, MA) prior to MS analysis. For positive ion analysis, samples were prepared in a 50:50 mixture

of water and acetonitrile containing 0.1%(v/v) acetic acid. For negative ion analysis samples were prepared in 40:60 water to acetonitrile mixture containing 20mM piperidine. All mass spectra were acquired with an IonSpec (Lake Forest, CA) 9.4-Tesla HiRes electrospray Fourier transform ion cyclotron resonance mass spectrometer (ESI-FTICR-MS), as described previously for the Sml1-histag species (Uchiki, Hettich et al. 2002). The MS experiment consisted of the four following steps: (1) ions are generated in the electrospray source, (2) the ions are accumulated in an external hexapole ion guide, (3) the ions are transferred into the high vacuum region with a quadrupole lens system, and (4) the ions are detected in the cylindrical analyzer cell of the mass spectrometer. To enhance ion trapping, nitrogen gas was pulsed into the mass analyzer to cool the ion packet prior to detection. Ions were measured under broadband conditions with resolutions ranging from 50,000 – 150,000 (FWHM). External calibration was accomplished using the various charge states of bovine ubiquitin. The high-resolution mass measurement enables isotopic resolution of multiply charged ions. Thus, the charge state of multiply-charged ions can be determined solely by its isotopic spacing (Horn, Zubarev et al. 2000). The deconvoluted *molecular* mass spectra were generated with the IonSpec software from the electrospray mass spectra by multiplying the masses of the electrospray ions by their respective charge, and then subtracting the masses of the protons added. This "unfolds" the multiply charged ion mass spectrum into a more easily interpreted molecular mass spectrum. Errors in the mass measurement for the multiply charged ions were scaled proportional to the charge in the calculation of the molecular masses in the deconvoluted mass spectra. By calibrating on the calculated values of the

most abundant isotopic peaks of six different charge states (7^+ to 12^+) of bovine ubiquitin, the deconvoluted molecular mass spectrum yielded a measured molecular mass for ubiquitin that was within 0.030 Da of the calculated value.

For the SORI-CAD experiments, a pulsed valve was used to admit the nitrogen collision gas into the high vacuum region to a maximum pressure of about 5×10^{-6} Torr during the ion excitation step. A base pressure of about 1×10^{-9} Torr was re-established prior to ion detection.

Analysis of protein phosphorylation by mass spectrometry

In this section, the use of mass spectrometry for analysis of phosphoproteins or phosphopeptides will be discussed. Although several highly effective techniques for isolation, detection, and mass spectrometric fragmentation of phosphopeptide/protein have been developed in the field, the discussion will be limited to methodologies appropriate for FTICR-MS.

The majority of protein phosphorylation in eukaryotes occurs on the hydroxyl of serine, threonine or tyrosine residues (Hunter 1991), and phosphorylation of these residues results in mass increase of 79.966Da (in monoisotopic mass). Most of the mass spectrometers used today can easily resolve the 80Da mass shift, and the 80Da mass shift is often the first observation of phospho-protein/peptide. However, peaks of proteins or peptides spaced by 80Da do not always indicate phosphorylation. For example, sulfation on the side chain hydroxyl groups of serine, threonine, and tyrosine has been observed (Moore 2003; Medzihradszky, Darula et al. 2004), and adds 79.9568 Da (in monoisotopic

mass) to these residues. It is also probable that two protein or peptide masses may differ by 80 Da. For these reasons, an instrument that can provide high mass resolution such as FTICR-MS is useful for distinguishing mass shifts due to phosphorylation and other causes. Sometimes, to unambiguously determine that the mass shift of 80 Da is due to phosphorylation, a sample containing the corresponding unphosphorylated peptide is compared to a sample containing the putative phosphorylated peptide. For example, the incubation of phosphoprotein with protein phosphatases such as alkaline phosphatase dephosphorylates the phosphoprotein. The direct comparison of two sets of putative phosphoprotein samples, one treated and one untreated with a protein phosphatase, is a common method for confirming the phosphorylation. This method is especially useful for the analysis of proteins phosphorylated *in vivo.* On the other hand, for *in vitro* phosphorylation analysis, confirmation of the phosphorylation can be accomplished by comparison of substrate proteins incubated with active kinase and catalytically inactive kinase (or without any kinase if the kinase sample is pure).

CAD of phosphopeptides usually generates unique fragment ions. The most common CAD fragmentation of phosphopeptides is neutral loss of phosphoric acid (-H_3PO_4 Mr:-98Da) (Figure 12A) from phosphoserine or phosphothreonine and neutral loss of phosphate (-HPO_3 Mr:-80Da) from phosphotyrosine (Mann, Ong et al. 2002). In negative ion analysis, a phosphate ion (-PO_3^- m/z:79) is also generated. These fragmentation patterns are a strong indication of protein/peptide phosphorylation. However, because these types of fragmentations are energetically favored,

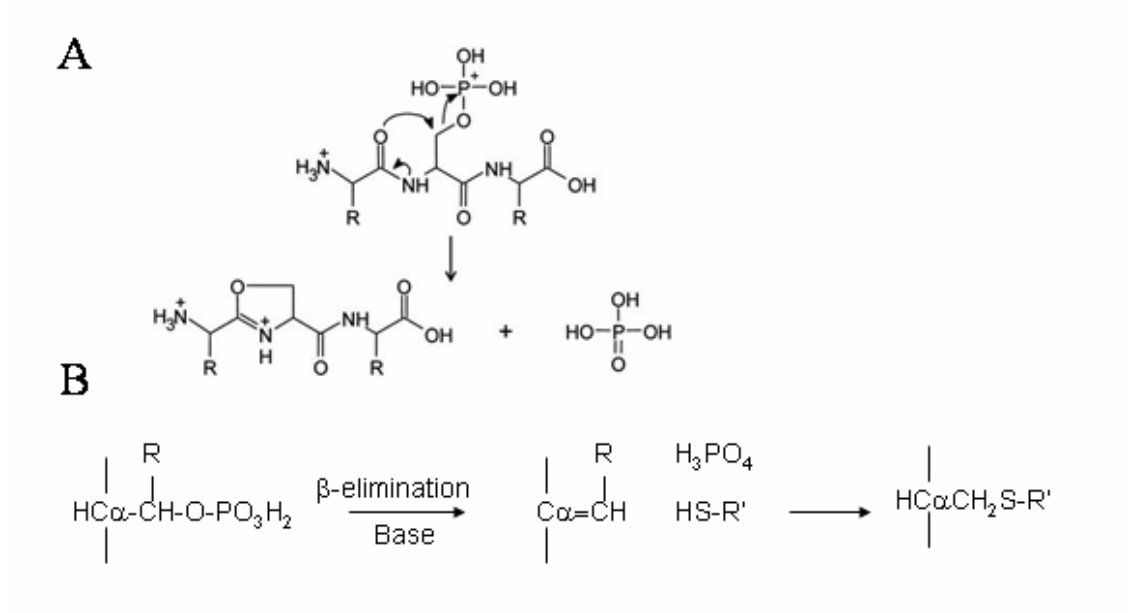

Figure 12. CAD of phosphoserine or phosphothreonine.

(A) (Images were taken from (Syka, Coon et al. 2004).) CAD Fragmentation scheme for loss of phosphoric acid from a multiply protonated phosphopeptide. In CAD of peptides containing phosphoserine or phosphothreonine residues, loss of phosphoric acid is a preferred fragmentation over fragmentation at peptide backbone. **(B)** Beta elimination of phosphoserine or phosphothreonine followed by reaction with sulfhydryl group in solution. In the first step, R is hydrogen (H) for phosphor-serine, while R is methyl group (CH_3) for phosphothreonine. In the presence of strong base such as LiOH, phosphor-serine or phpsphothreonine undergoes β-elimination and is converted into dehydroalanine or dehydroaminobutyric acid. These derivatives react with sulfhydryl group (HS-) at acidic or neutral pH.

fragmentations at the peptide backbone (which provides information for exact location of the phosphoacceptor residues) do not occur extensively on phosphopeptides. When analyzing phosphoserine or phosphothreonine residues, this problem can be solved by chemical derivatizations. For example, in the presence of a strong base such as LiOH, both phosphoserine and phosphothreonine undergo β elimination and are converted into dehydroalanine and dehydroamino-2-butyric acid (Figure 12B). The β carbon of these residues can react with the sulfhydryl group to form a covalent linkage. For example, ethanethiol has been used to convert phosphoserine to S-(2-mercaptoethyl) cysteine (Meyer, Hoffmann-Posorske et al. 1991). The side chain of S-(2-mercaptoethyl)cysteine is not easily fragmented in CAD, and fragmentation on the derivatized peptide mainly takes place at the peptide backbone, allowing identification of phosphoacceptor residues (Lapko, Jiang et al. 1997; Uchiki, Dice et al. 2004). Similar derivatizations using other compounds with sulfhydryl groups have been developed for biotinylation (Adamczyk, Gebler et al. 2001; Oda, Nagasu et al. 2001) or specific proteolysis (Knight, Schilling et al. 2003) specifically targeting phosphoserine or phosphothreonine.

A challenge for phosphopeptide analysis is to detect the phosphopeptides from mixture of other unphosphorylated peptides. In some cases, only a small fraction of the target protein is phosphorylated, making the analysis more difficult. Therefore, prior enrichment or isolation of phosphopeptides is necessary in many cases. The most common strategy for enrichment of phosphopeptide is immobilized metal affinity chromatography (IMAC) using metal ions such as Fe(III), Ga(III) or Al(III). It has been reported that phosphate or phosphopeptide binds to these metal ions at low pH and

dissociates at high pH or in the presence of competing phosphate groups (Andersson and Porath 1986). On the other hand, most of unphosphorylated peptides do not interact with these ions, although the carboxylic acid of aspartate and glutamate also binds to these metal ions. Esterification of the carboxylic acid eliminates the its binding to IMAC and significantly increases purity of phosphopeptides (Ficarro, McCleland et al. 2002).

Enzymatic digestion and Immobilized Metal Affinity Chromatography (IMAC)

After phosphorylation reactions, the Sml1 sample was dialyzed against 100mM NaCl to reduce the ATP concentration to less than 1μM. For tryptic digestion, a sample containing approximately 10 μg (83 pmol) of Sml1 was incubated with 0.2 μg of sequencing grade trypsin (Promega, Madison, WI) at 37ºC for 10 hours. To quench the reaction, acetic acid was added to a final concentration of 0.1% (v/v). For cyanogen bromide (CNBr) digestion, a sample containing approximately 10 μg of Sml1 was dried using a speed vacuum apparatus. 10 μl of 10 mg/ml CNBr (SIGMA, St Louis, MO) in 70% (v/v) formic acid was added to the sample and incubated for 24 hours. To quench the reaction, 50 μl of water was added, and the sample was dried using a speed vacuum apparatus. This step was repeated five times.

A 20 μl bed volume of Ga(III)IMAC column with Poros MC resin (PerSeptive Biosystems, Framingham, MA) was prepared as described in (Posewitz and Tempst 1999). The CNBr digested samples were dissolved in 10μl of 10% (v/v) of acetic acid and manually loaded to the column pre-equilibrated with 1% (v/v) acetic acid. The

column was washed with six bed volumes of 1% (v/v) acetic acid, six bed volumes of a mixture consisting of water, acetonitrile, and acetic acid in a 70:30:1 respective ratio, followed by six bed volumes of 1% (v/v) acetic acid in water. The bound peptide was eluted with three bed volumes of 200 mM sodium phosphate (pH 8.5).

Derivatization of phosphoserine

The phosphopeptides enriched by Ga(III) IMAC were desalted once with C18 reverse phase ZipTip columns (Millipore, Bedford, MA), and the sample volume was reduced to less than 5 µl using a speed vacuum apparatus. The phosphoserine residues of the peptides were chemically modified to S-ethylcysteine by a β-elimination reaction in the presence of ethanethiol as described in (Meyer, Hoffmann-Posorske et al. 1991; Oda, Nagasu et al. 2001). Briefly, 50 µl of H_2O/4M LiOH/ acetonitrile/ ethanol / ethanethiol mixed in a ratio of (5: 14: 5: 5: 2) were added to the sample and incubated in 37ºC for one hour. The reaction was quenched by addition of 25 µl of acetic acid. Since Sml1 contains only one cysteine residue (Cys14), we omitted the oxidation step, which is normally employed prior to the derivatization.

Quantitative assay for biochemical activity of Dun1, RNR and Sml1.

The Dun1 kinase assay

A 15µl bed volume of the resin-bound GST-Dun1 was mixed with 30µl of the kinase reaction mixture (50mM Tris-HCl, 10mM $MgCl_2$, 1mM DTT, 60 to 250µM ATP, 0.06 µCi/µl (1Ci=37GBq) of [γ-^{32}P]ATP (4500mCi/mmol: MP Biochemicals, Costa Mesa,

CA), 0.1 μg/μl (8.3μM monomer) Sml1, pH 7.5) and incubated in 30ºC water bath for thirty minutes to two hours. Phosphorylated Sml1 for mass spectrometric analysis was prepared in the same method except that [γ-^{32}P] ATP was omitted from the reaction mixture. Upon the completion of the reaction, 10μl of 400 to 900mM of cold ATP dissolved in the kinase reaction buffer was added and the reaction tubes were placed on ice to stop the reaction[16]. After a brief centrifugation, 30μl of the supernatant was placed on 1.7cm^2 P81 phosphocellulose filter (Whatman, Maidstone, England), and the filter was incubated at room temperature for 5 minutes. Then, the filter was washed with 300ml of 10mM phosphoric acid twice for 5 minutes each and then with 15ml of 95% (v/v) ethanol for three minutes. Through its basic amino acid residues, Sml1 is bound to the phosphocellulose filter, while [γ-^{32}P]ATP is washed off. The radioactivity (count per minutes (cpm) in energy channel for P32) that remains on the filter represents ^{32}P covalently attached to Sml1, and this was measured by a liquid scintillation counter (LS3801: Beckman, Fullerton, CA).

To determine the molar concentration of radioactivity, 1μl of reaction mixture was diluted to 1/100 prior to the reaction and, the radioactivity of 2, 4 and 8μl of the diluted reaction mixture was measured. Similarly, total concentration of ATP in the reaction mixture was measured based on absorbance at 259nm using an extinction coefficient of ε_{259}=14500cm^{-1}M^{-1}. Based on these measurements, radioactivity (cpm)

---

[16] In this method, phosphorylation still takes place after addition of cold ATP. However, incorporation of ^{32}P to Sml1 is significantly slowed down. Over 30 minutes on ice after addition of 500mM ATP, increase of ^{32}P incorporation to Sml1 was indistinguishable from fluctuation of background radioactivity, which is 100-500cpm.

given by 1μl of the reaction mixture and total moles of ATP in 1μl were obtained. The ratio of radioactivity (cpm/μl) and moles of ATP (moles/μl) represents specific activity of ATP or phosphate (how much radioactivity corresponds to a mole of γ-phosphate, which is expressed as cpm/pmol). Similarly, the total amount (moles) of Sml1 in reaction mixture applied to phosphocellulose paper was determined based on the initial concentration of Sml1 in the reaction mixture and a 30/44.4 dilution[17] due to mixing with the resin.

For autoradiography, 10 to 20 μl of the supernatant was mixed with SDS-PAGE gel loading buffer (62.5mM Tris-HCl pH 6.8, 2 % SDS, 10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol and 0.025 bromophenol blue), was separated on a 15 % polyacrylamide gel, and the gel was dried on cellophane membranes. Cerenkov radiation on the gel was detected by an electric autoradiography imager (Instant Imager: A Packard Bioscience Company, Ontario, Canada) and exposure of a scientific imaging film (BioMax Light film: Kodak, Rochester, NY) to the gel for 10 to 24 hours.

Ribonucleotide reductase (RNR) activity assay and inhibition of RNR by Sml1

The velocity of CDP reductase activity in steady state kinetics was determined following the methods described by Stubbe and co-workers (Ge, J., D. L. Perlstein, et al. 2001) except that DTT was used as the reducing agent of RNR. First, 1μM of Rnr1 dimer and

[17] According to manufacturer of the glutathione resin (Amersham Biosciences), approximately 96% of the bed volume of the resin is occupied with liquid by which the resin is equilibrated. 15μl of resin contains 14.4μl of buffer C. Mixing 30μl of the reaction mixture causes 30/44.4 dilution of components in the reaction mixture.

2.5-5μM of Histag Rnr2/Rnr4 dimer were prepared in 50mM Hepes-KOH pH 7.4, 200mM potassium-acetate, 20mM magnesium-acetate, 20μM $FeCl_3$, 3mM ATP, 20mM DTT, and the mixture was incubated in a 30ºC water bath for 3 minutes. In this mixture, ATP is an activator of ribonucleotide reductase. ^{14}C labeled CDP (Specific activity = 2000 to 3000cpm/nmol) was added to a final concentration of 1mM[18]. Immediately after addition of CDP, a 50μl aliquot was taken from the reaction mixture and placed in a boiling water bath to quench the reaction. Time point at the first quenching taken as time zero. The reaction was continued in the 30ºC water bath, and 50μl aliquots were taken every 3 or 4 minutes. A total of 4 to 5 aliquots were taken over 12 to 20 minutes. After boiling, the aliquots were centrifuged to sediment precipitated materials. 950μl of 10.5u/ml Alkaline phosphatase and 400μM deoxycytidine (as carrier), 50mM Tris-HCl pH 9.0, 0.1mM Zn-acetate was added to each aliquot, and was incubated at 37ºC for 3 hours. Alkaline phosphatase digests the reaction product, dCDP, into deoxycytidine, and this process is necessary to prevent binding of dCDP to the borate column via its phosphate groups. The samples digested by alkaline phosphatase were loaded onto the borate column (0.6cm I.D. x 8cm L =2.3cm^3). The cis-diol of cytidine binds to a borate

---

[18] Km of CDP reduction *by S.cerevisiae* RNR has not been determined. Km of CDP reduction by *R. typhi* (murine) RNR is 2μM Kashlan, O. B. and B. S. Cooperman (2003). "Comprehensive model for allosteric regulation of mammalian ribonucleotide reductase: refinements and consequences." Biochemistry **42**(6): 1696-706.. Assuming that Km of *S.cerevisiae* and *R. typhi* enzyme is similar, 1mM of CDP concentration was chosen to be saturating concentration of yeast RNR. Although this is not ideal method to choose CDP concentration, this concentration has been chosen by two groups (JoAnn Stubbe's and Lars Thelanders' groups) that have reported *S.cerevisiae* RNR activity (Nguyen, H. H., J. Ge, et al. (1999). "Purification of ribonucleotide reductase subunits Y1, Y2, Y3, and Y4 from yeast: Y4 plays a key role in diiron cluster assembly." Proc Natl Acad Sci U S A **96**(22): 12339-44, Chabes, A., V. Domkin, et al. (2000). "Yeast ribonucleotide reductase has a heterodimeric iron-radical-containing subunit." Proc Natl Acad Sci U S A **97**(6): 2474-9.

ion forming a cyclic ester (Kim, Faull et al. 2004). Therefore, un-reacted substrate derivatives are retained on the column. On the other hand, the product derivative, deoxycytidine, does not interact with borate and passes through the column freely. The column was washed with an additional 8ml of deionized water. All the flow-through during loading and washing was collected in a tube. After the flow-through was mixed thoroughly, 3 ml of this was taken into 15ml of an aqueous scintillation cocktail and radioactivity of the sample (counts per minutes (cpm) in energy channel of 110-610) was measured by liquid scintillation counter (LS3801: Beckman, Fullerton, CA). Moles of deoxycytidine present in the sample were determined based on the radioactivity. A plot of time (min) as independent variable and amount of dCDP (nmol) as dependent variable was made, and linear regression was performed. The slope of the curve represents velocity of the reaction (nmol dCDP formed/min), which is defined as activity. Based on the amount of Rnr1 present in each aliquot (10μg), specific activity (nmol dCDP/min/mg Rnr1) was also determined. (See appendix for a sample calculation).

The ability of Sml1 to inhibit RNR was measured by performing the RNR activity assay in the presence of 1.3 to 17μM Sml1 monomer. At the same time (using the same batch of RNR), an RNR activity assay was performed without Sml1. The ratio of RNR activity in the presence and in the absence of Sml1 was taken as the relative degree of RNR inhibition.

Preparation of materials for RNR activity assay

Preparation of borate column

200ml of anion exchange beads (AG-X8 50-100 mesh: BioRad, Hercules, CA) was packed in a column (2.5cm I.D. x 75cm L), and 10 column volumes of 1N NaOH was passed through the column to replace Cl^- on the beads to OH^-. After washing with 10 column volumes of deionized water, the beads were equilibrated with saturated potassium tetraborate (~0.6M) and incubated at room temperature overnight to load the borate ion. After a final wash with 10 column volumes of deionized water, the beads were stored in 20% (v/v) ethanol. To make the column for the activity assay of ribonucleotide reductase, bottom of a Pasteur pipette (0.6cm I.D.) was plugged with a small piece of glass wool, and then approximately 2.3ml of the borate resin was packed by gravity.

Preparation of CDP stock

100µl of 0.05µCi/µl of ^{14}C labeled CDP (50mCi/mmol. Shipped as 0.05µCi/µl in 50% ethanol: Moravek Biochemicals, Brea, CA) was lypholized to complete dryness as it is shipped in 50% ethanol. Approximately 30mM of non-radio labeled CDP (Sigma-Aldrich, St Louis, MO) was prepared in 20mM Hepes-KOH pH 7.0. 123 to 185µl of the non-radioactive CDP solution was added to the dried ^{14}C labeled CDP and mixed thoroughly. 2000 to 3000cpm / nmol of specific activity was expected from this volume ratio of the radioactive to the non-radioactive CDP. To accurately determine concentration and radioactivity of CDP in the CDP stock, the following procedures were

taken. After mixing thoroughly, 1/1000 dilution of the CDP stock was made in 20mM Hepes-KOH pH7.0. From the diluted solution, the concentration (mM) of CDP was determined based on absorbance at 270nm with an extinction coefficient of 9000 M^{-1} cm^{-1}. Similarly, a 1/100 dilution of the CDP solution was made in Hepes-KOH pH 7.0. 2, 4, 10µl of the diluted CDP stock (1/100 dilution) was mixed with aqueous scintillation cocktail and their radioactivity was measured. A plot of the volume as independent variable and cpm as dependent variable was created, and linear regression was performed on the plot. The slope of the curve represented concentration of radioactivity (cpm /µl) of the diluted CDP stock. This value was multiplied by 100 to obtain concentration of radioactivity in the undiluted CDP stock. Based on concentrations of CDP and radioactivity determined by the procedures above, specific activity (cpm/nmol) of the CDP stock was determined. (See appendix for sample calculation).

Other techniques

Fluorescence based method to monitor the Sml1-Rnr1 interactions

Conjugation of IANBD fluorescence probe to Sml1

Sml1 or its variants were prepared in the method described above except that PBS was used as running buffer for Superdex75 gel filtration chromatography. Sml1 concentration was determined using the Coomassie Plus Protein Assay kit (PIERCE, Rockford, IL) using a BSA standard supplied with the assay kit. The Sml1 concentration was kept at about 0.1 to 0.2mg/ml (8.3~16.6µM monomer) before conjugating the fluorescence probe. Throughout the procedures involving *N*,*N*'-dimethyl-*N*-(iodoacetyl)-*N*'-(7-

nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamine (IANBD-amide: Molecular Probe, Eugene, OR), exposure of the sample to light was minimized by wrapping the sample with aluminum foil and working under red light. IANBD-amide was dissolved in dimethyl sulfoxide (DMSO) to a final concentration of 5 to 15mM and centrifuged for 10 minutes at 20,000g at room temperature. A small amount of the solution was diluted by 1/10000 in methanol, and the concentration of IANBD-amide was determined from the diluted sample based on absorbance at 478nm and the extinction coefficient of IANBD-amide (ε_{478}=25000$cm^{-1}M^{-1)}$. IANBD-amide dissolved in DMSO was added to the Sml1 solution to at 10 to 15 molar excess of Sml1. Then, 1M Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) (pH ~7.0) was added to the mixture of Sml1 and IANBD at 10 molar excess of IANBD-amide. The mixture of Sml1, IANBD-amide and TCEP was kept in the dark and incubated at 4ºC for 12 to 16 hours. After incubation, β-mercaptoethanol was added at 15 molar excess of IANBD-amide. In this step the remaining IANBD-amide reacts with β-mercaptoethanol. This solution was incubated at 4ºC in the dark for an additional 12 to 16 hours. The solution was concentrated to a final volume of 2 to 5 ml by using a Centriprep unltrafiltration vessel with YM10 (Millipore Bedford, MA) membrane and then applied to Superdex75G gel filtration chromatography using 50mM Hepes-KOH, 100mM KOH, 5mM MgCl2, 5mM DTT, 5%(v/v) glycerol pH7.0 as the running buffer. IANBD conjugated with Sml1 and β-mercaptoethanol was separated on the column. Fractions containing Sml1 were further concentrated to 1~3mg/ml by Centriprep with a YM10 membrane, divided into aliquots and frozen in liquid nitrogen. The sample was wrapped with aluminum foil and kept in –80ºC until they were used. Concentration

of IANBD probe in the sample was determined based on absorbance at 478nm with the extinction coefficient of 25,000$M^{-1}cm^{-1}$. The concentration of Sml1 was determined by Coomassie Plus Protein Assay kit as described above. The molar ratio of IANBD to Sml1 was >0.9 demonstrating that over 90% of Sml1 was conjugated with the fluorescence probe.

IANBD used as a negative control was prepared by incubating 1mM IANBD-amide in DMSO and 10mM β-mercaptoethanol (BME) in 4°C for 24 hours in the dark. After the incubation, the solution was centrifuged at 20,000x g for 10 minutes to sediment precipitated materials. The concentration of IANBD conjugated with BME was determined based on absorbance at 478nm and the extinction co-efficient of IANBD-amide (ε_{478}=25000$cm^{-1}M^{-1}$). This solution was further diluted in 50mM Hepes-KOH, 100mM KOH, 5mM $MgCl_2$, 5mM DTT, 5%(v/v) glycerol (pH7.0) and stored in -20°C. The sample aliquots were wrapped with aluminum foil until being used. (The negative control IANBD will be called IANBD-BME in the following section).

Fluorescence spectroscopy to monitor the Sml1-Rnr1 interactions

5μM of Wild type or C14S/S60C Sml1 monomer conjugated with IANBD-amide (IANBD-Sml1 or IANBD-S60C Sml1) was prepared in 50mM Hepes-KOH, 100mM KOH, 5mM $MgCl_2$, 5mM DTT, 5%(v/v) glycerol (pH7.0). In a separate aliquot, a mixture of 5μM IANBD-S60C Sml1 and 30μM wild type Rnr1 monomer was prepared. 100μl of 5μM IANBD-S60C Sml1 was placed in a quartz cuvette (3 x 3 x 20mm) and emission scan was performed for 1 second/nm at an excitation wavelength of 478nm. The slit width was adjusted to give an emission intensity of 150 to 200 (arbitrary units). 50μl

of Rnr1/IANBD-Sml1 mixture (30μM wild type Rnr1 monomer, 5μM IANBD-Sml1 monomer) or Rnr1/IANBD-S60C Sml1 mixture (30μM wild type Rnr1 monomer, 5μM IANBD-S60C Sml1 monomer) were added to the cuvette. After mixing the solution thoroughly by pipette, the same emission scan was performed without changing the slit width.

Quantitative measurement of Sml1-Rnr1 interactions by fluorescence spectroscopy

5μM of C14S/S60C Sml1 monomer conjugated with IANBD-amide (IANBD-S60C Sml1) was prepared in 50mM Hepes-KOH, 100mM KOH, 5mM MgCl2, 5mM DTT, 5%(v/v) glycerol (pH7.0). In a separate aliquot, a mixture of 5μM IANBD-S60C Sml1 monomer and 30μM wild type Rnr1 monomer was prepared in the same buffer (Sml1 / Rnr1 mixture). 100μl of 5μM IANBD-S60C Sml1 monomer was placed in a quartz cuvette (3 x 3 x 20mm) and its fluorescent intensity was measured for 1 second at an excitation wavelength of 478nm and an emission wavelength of 541nm using the fluorometer (LS50B: Perkin Elmer, Wellesley, MA). The slit width was adjusted to obtain an emission intensity of 150 to 200 (arbitrary units). Then, the sample was titrated with the Sml1 / Rnr1 mixture at 3μl increments. After each addition of the Sml1 / Rnr1 mixture, the solution in the cuvette was mixed by aspirating and dispensing the solution by a pipette. As a negative control, the same procedure was performed with 5μM IANBD-BME and a mixture of 5μM IANBD-BME and 30μM Rnr1 monomer. Each measurement was repeated three times. The fluorescent intensity at each titration point was manually recorded and analyzed by GraphPad Prism software (GraphPad, San

Diego, CA). First, the average intensity of the negative controls at each titration point was calculated. Second, the average of the negative control was subtracted from the intensities given by the corresponding titration points in IANBD-C14S/S60C Sml1. (This value will be called delta F in the following description). Third, the total concentration of Rnr1 as the independent variable was plotted against delta F as the dependent variable. Fourth, by non-linear least square regression, the data were fitted to the equation below that expresses one-to-one binding model of the Sml1-Rnr1 interactions:

$$DeltaF = DeltaF\max\frac{[Sml1]+[Rnr1]+K_d-\sqrt{([Sml1]+[Rnr1]+k_d)^2-4[Sml1][Rnr1]}}{2[Sml1]}$$

(Delta F_{max} : Delta F where Sml1 is saturated with Rnr1. K_d: dissociation constant of the Sml1-Rnr1 interactions. [Sml1]: total concentration of Sml1. [Rnr1]: total concentration of Rnr1).

The model equation is based on these assumptions:

1. The stoichiometry of Sml1 and Rnr1 in the complex is one to one
2. The change in the fluorescence intensity is proportional to the fraction of Sml1 bound to Rnr1.

Estimation of size of proteins by gel filtration chromatography

Fractions corresponding to 24 - 30 kDa (corresponding to intact Sml1) were pooled and concentrated to 2 mg/ml by stirred ultrafiltration with a membrane of MWCO 10,000 (Millipore, Bedford, MA). The protein concentration was determined by A_{595} of the

protein sample reacted with Commassie® Plus Protein Assay Reagent (Rockford, IL) using bovine serum albumin (BSA) as a standard. Based on image optical density (IOD) of an SDS-PAGE gel, we estimated that more than 85% of protein in the sample was intact Sml1. A total of 500µl (3mg) of the protein sample was applied to a gel filtration column, Superose 12 (HR10-30; Amersham Biosciences, Uppsala, Sweden) on a Pharmacia fast protein liquid chromatograph. Elution was accomplished with 50mM Tris-HCl pH 8.0, 100mM NaCl, 1mM EDTA and 5mM DTT (or in the absence of DTT) at a flow rate of 1ml/min. To estimate the apparent molecular mass of Sml1, BSA (66 kDa), Ovalbumin (43 kDa), Carbonic anhydrase (29 kDa) and Ribonuclease A (13.7 kDa) were used as molecular weight standards, and blue dextran was used to determine void volume. Based on elution volume of the standards, K_{av} [= (elution volume –void volume) / (bed volume – void volume)] was determined. A standard curve was created by plotting $\log_{10}$[Molecular weight] as the independent variable on the x-axis against K_{av} as dependent variable on the y axis and performing a least square linear regression. The chromatography run was repeated three times for both the molecular weight standards and Sml1. The standard deviation of the retention times for both Sml1 and the protein standards were less than one second.

Chapter 3. Characterization of Sml1-histag protein by ESI-FTICR-MS (Uchiki, Gupta et al. 2002)

Introduction

The major goal of study presented in this chapter was to investigate integrity and homogeneity of Sml1-histag protein expressed and purified for crystallization. A recombinant Sml1 which is conjugated with 6xHis tag at its C-terminus (Sml1-histag) was characterized by mass spectrometry. Although several trials of crystallization were not successful, the mass spectrometric analysis of Sml1-histag revealed some interesting properties of this protein. The methods described in this chapter were mainly taken from the Results and Discussion section of the article we published in 2002 (Uchiki, Gupta et al. 2002), which describes the disulfide-mediated dimer and Na^+/K^+ adducts of Sml1-histag. (Experimental procedures used in this study are described in Chapter 2.)

Results and discussion

Determination of the molecular mass for the Sml1-histag monomer

The predicted amino acid sequence of the recombinant Sml1-histag protein, based on the DNA construct used in the *E. coli* expression vector, is illustrated in Figure 13. This protein would have an average molecular mass of 13,376.87 Da. This protein is identical to the wild type Sml1, except that this histag version has 4 extra amino acids at the N-terminus and 10 extra amino acids at the C-terminus (as noted by the bold, underlined regions in Figure 13) that come from the bacterial expression vector. SDS-PAGE was conducted under reducing conditions (i.e. beta-mercaptoethanol present) on the purified

1 10 20 30
MGGSMQNSQDYFYAQNRCQQQQAPSTLRTV
31 40 50 60
TMAEFRRVPLPPMAEVPMLSTQNSMGSSAS
61 70 80 90
ASASSLEMWEKDLEERLNSIDHDMNNNKFG
91 100 110 118
SGELKSMFNQGKVEEMDF**GSRSHHHHHH**

Mr = 13376.87 Da

Figure 13. Predicted sequence of recombinant Sml1-histag

Sml1-histag protein, and verified that it had the expected molecular mass of about 14kDa, as shown in lane 2 of Figure 14. The histag purification procedure was successful in isolating this protein to a purity of about 80% (based on the IOD calculation described in Chapter 2). Although other faint bands were noted on the gel, no attempt was made to determine the identities of these minor species in the sample.

The positive ion ESI-FTICR mass spectrum of the Sml1-histag protein sample (i.e. the same sample that was used to generate the data in Lane 2 of Figure 14) is shown in Figure 15A, and reveals multiply-charged molecular ions ranging from *nominal* m/z 889 (15^+ ions) to *nominal* m/z 1333 (10^+ ions). Because the mass measurement resolution (FWHM) in Figure 15A is about 50,000, it is possible to resolve not only different protein species, but also the natural isotopic distributions within individual species. By measuring the spacing between adjacent isotopic peaks, it is possible to determine directly the charge state of any given ion. Figure 15A reveals the presence of several protein species in this sample, as indicated by the multitude of ions present at each charge state. This is illustrated in the inset of Figure 15A for the 13^+ molecular ions in the expanded m/z 1015-1040 region.

Deconvolution of the *multiply-charged* electrospray mass spectra of Figure 15A can be used to present the data in a more easily interpreted *molecular* mass spectrum, as shown in Figure 15B. Evident in this figure are several protein species with average molecular masses ranging from 13,266 to 13,391 Da (the presence of naturally-occurring isotopes such as ^{13}C provides an envelope for each molecular species). The appearance of multiple peaks differing by a few tens of Da around the molecular mass region

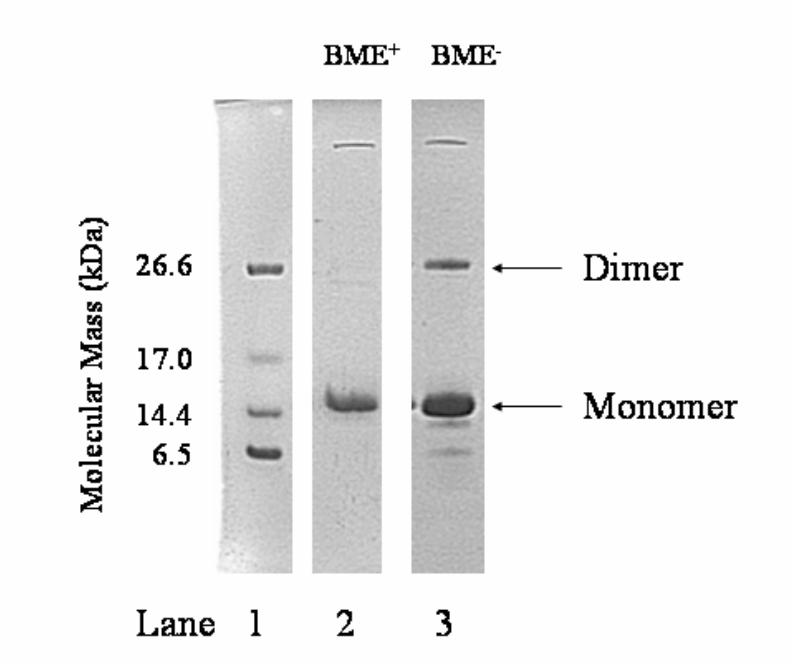

Figure 14. SDS-PAGE of recombinant Sml1-histag protein after purification, as analyzed with reducing and non-reducing 15% polyacylamide gels.

15 μg of protein was loaded onto lanes 2 and 3. Lane 1- Molecular weight marker (kDa), Lane 2 –Ni column eluate prepared with 35mM β-meracaptoethanol, Lane 3- Ni column eluate prepared without β-meracaptoethanol.

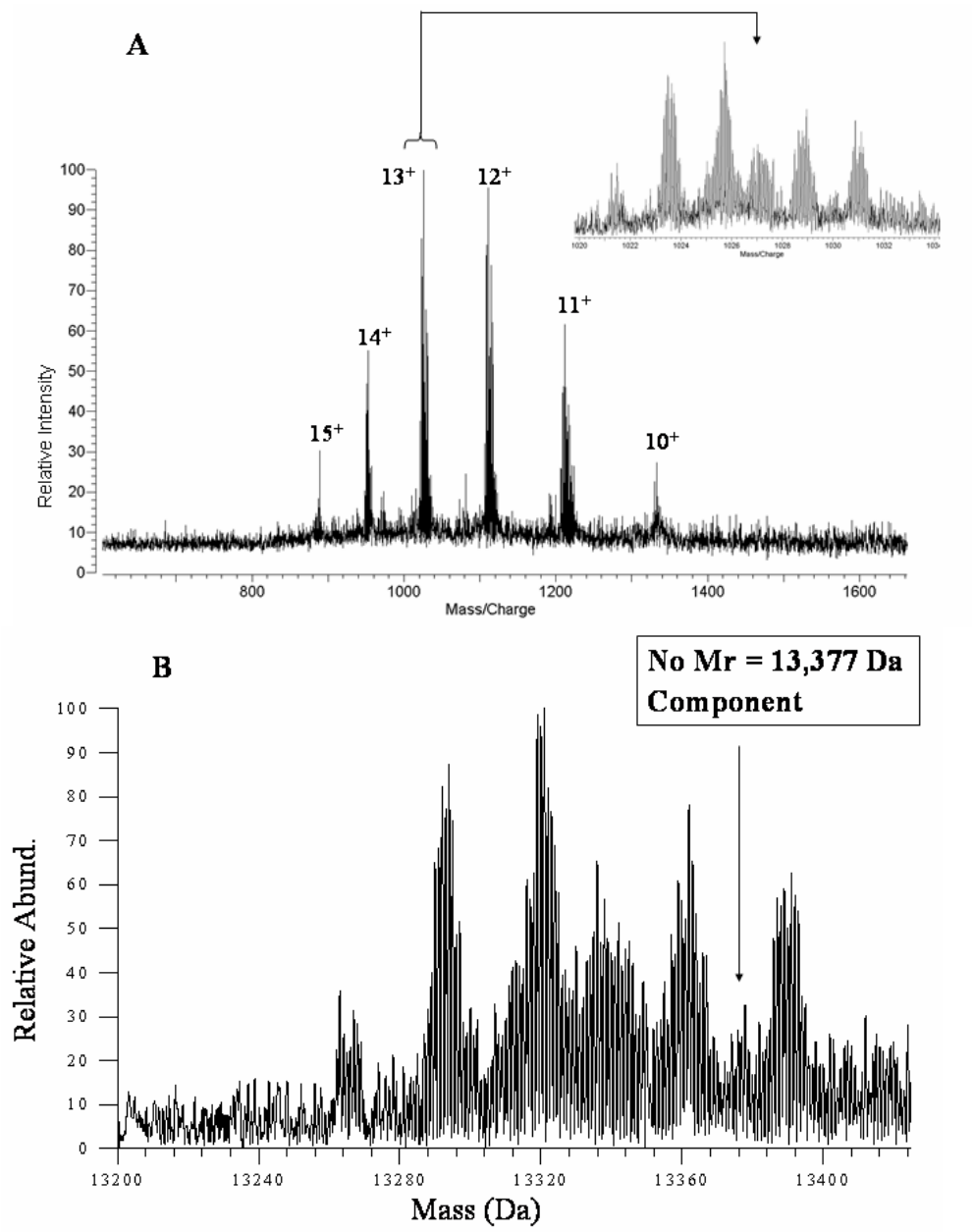

Figure 15. Positive ion ESI-FTICR mass spectrum of purified Sml1-histag.

(A) electrospray mass spectrum revealing multiply-charged molecular ions, with inset showing the multiple components at a single charge state. **(B)** Deconvoluted molecular mass spectrum illustrating the presence of several distinct species in this sample.

suggests the possibility of salt contamination (i.e. sodium and/or potassium adduction) in the protein sample; however, note that most of these species are *lower* in mass than what would be expected for alkali metal ion attachment to the protein expected from Figure 13. There is virtually no ion signal detected at the expected molecular mass of 13,377 Da. The two most abundant species in Figure 15B were measured at average Mr = 13292 Da and Mr = 13321 Da. The appearance of multiple components with similar molecular masses in this "purified" sample suggests that the recombinant Sml1-histag protein may have sequence variants, chemical modifications (such as alkali metal adduction), or post-translational modifications, although the latter is less likely as this protein was expressed in the cytoplasm of a bacterial system. These multiple species are too close in mass to be resolved on the SDS-PAGE shown in Figure 14, which shows primarily one band at approximately the expected molecular mass. This illustrates the power of mass spectrometry to provide high-resolution mass measurement and differentiation of proteins and their variants in mixtures. The nature of the molecular mass discrepancy measured in Figure 15B relative to the expected value from Figure 13 was further investigated by proteolytic digestions combined with mass spectrometry as well as collisional dissociation experiments to resolve the identity(s) of the expressed Sml1-histag protein.

Verification of the sequence for the Sml1-histag monomer

Proteolytic digestions produce a variety of characteristic peptides for a protein and can be used to determine its amino acid sequence. To further investigate the purified

recombinant Sml1-histag, the same sample used to generate Figure 15 was subjected to digestion employing either trypsin or Glu-C proteases. Trypsin digestion of the Sml1-histag protein sample yielded a variety of peptides with molecular masses in the range of 700-4500 Da, as summarized in Table 2. The masses of these tryptic peptides can be used to search protein databases of sequenced organisms to verify protein identity with no further structural information. For example, the ten peptides listed in Table 2 whose masses were measured with ESI-FTICR-MS were used to search the NCBI database for proteins of eukaryotes, bacteria and virus[19]. For this search, the peptide monoisotopic masses (with a tolerance of 0.05 Da) and a maximum of two missed cuts by the trypsin protease were the specified inputs. The most probable protein "identified" with a rank=1, a probability of 1.0, and sequence coverage of 71% was the 11.8 kDa Sml1 protein from *saccharomyces cerevisiae* (yeast). For this search, 7 of the 10 peptides corresponded to within 0.05 Da to peptides expected from the trypsin digest of wild-type Sml1. The other three peptides (Mr = 1865, 1068, and 1861 Da) did not match any expected tryptic peptides from the wild-type Sml1, as might be expected since these three peptides are from the N-terminal and C-terminal regions of the recombinant protein that are known to be different from the wild type species. Note that the protein database search correctly (and with high confidence) indicated that the recombinant protein was closely related to the wild-type Sml1 species. The second protein "hit" from the database search was lectin – *furze* with a molecular mass of 3.72 kDa, which had a probability well below 0.01 and

[19] Typtic peptides were used to search the NCBi database with the ProFound program available at http://prowl1.rockefeller.edu. NCBI database records proteins from total 23 species representing from eukaryotes, bacteria and virus.

Table 2. Peptides generated by trypsin digest of Sml1-histag protein

Measured mass (Da)	Residues	Calculated mass** (Da)	Mass Diff. (Da)	Sequence Tag***
1864.780	1-16	1864.779	0.001	NSQDYFY
1260 (by MALDI-MS)	17-27	1258.608	*1.392*	
953.463	28-35	953.464	0.001	
3705.765	36-70	3705.749	0.016	
4347.999	36-75	4348.046	0.047	STQNSMG
1413.646	76-87	1413.630	0.016	
736.383	88-94	736.375	0.008	ELK
2132.007	76-94	2131.995	0.012	MNNN
810.373	95-101	810.369	0.004	
1068.468	102-110	1068.454	0.014	
1860.809	95-110	1860.813	0.004	VEEMDFGSR
(927) – not observed	111-117	927.396		

* Monoisotopic masses.
** Expected tryptic peptides based on Sml1-histag sequence shown in Figure 13. (without the N-terminus methionine)
*** Sequence tags identified from MS/MS of the selected peptides.

matched only 3 of the 10 peptides. By including only the six most abundant tryptic peptides from Table 2 (nominal m/z 953, 3706, 1414, 2132, 810, and 1068) to the database search, the highest rank protein still found was Sml1 (rank = 1, probability of 1.0). In this case, 5 of the 6 peptides matched exactly.

Collisional dissociation methods were employed to examine the fragmentation of many of the digest peptides from Table 2, and provided amino acid sequence tags in several cases that could be used to confirm the identity of the original protein, as well as to orient the peptide fragment into the correct position from the original protein. The use of sequence tags to identify proteins was pioneered by (Mann and Wilm 1994) as an alternative to using only the tryptic peptide masses, and can be used for peptides as well as intact proteins (Mortz, O'Connor et al. 1996). The collisional dissociation experiments yield detailed sequence information, which is particularly important for the unmatched peptides that are not assignable directly from the database. For example, the "undefined" Sml1 peptide at nominal Mr = 1865 Da (i.e. m/z 933^{2+} in the electrospray mass spectrum) was fragmented selectively with collisional dissociation. This experiment, which was conducted entirely inside the analyzer cell of the FTICR instrument, generated abundant *y*-type fragment ions (cleavage of amide bond between two amino acids, with the charge retained on C-terminus (Roepstorff and Fohlman 1984; Biemann 1988), and less abundant *b*-type fragment ions (charge retained on N-terminus) from this peptide. Inspection of the mass differences between adjacent *y*-type ions yielded an amino acid sequence tag of ***NSQDYFY,*** as illustrated in Figure 16**.** Because these fragment ions could be measured so precisely, it was possible to identify the amino acids of the

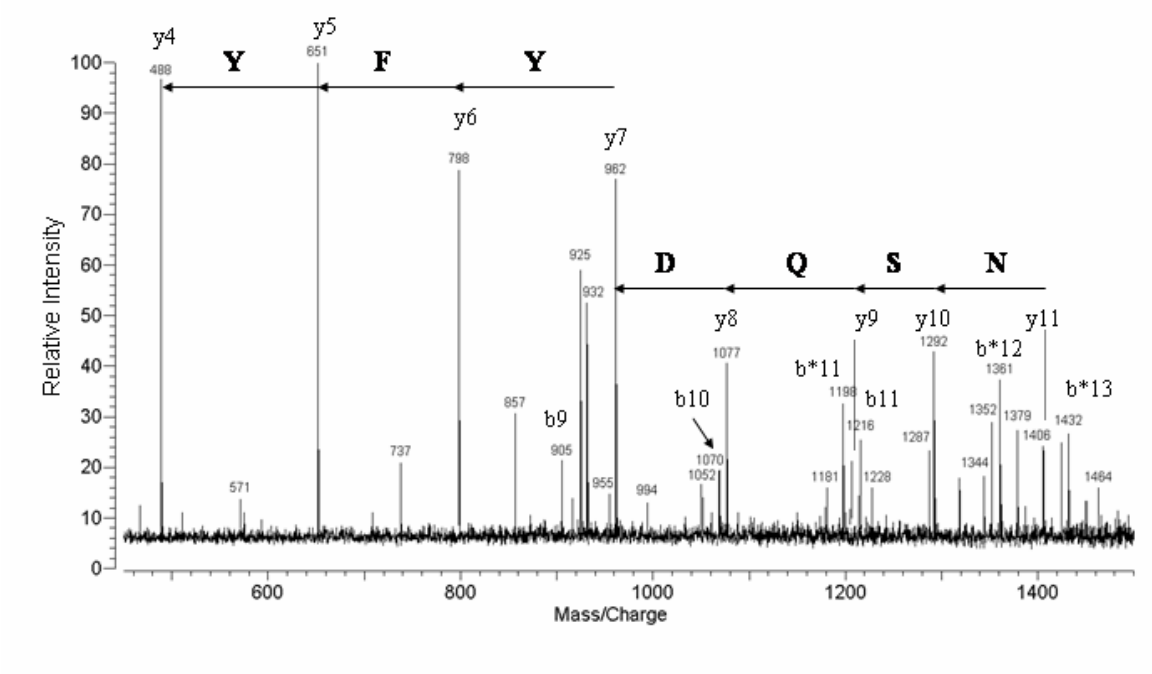

Figure 16. Positive ion collisional dissociation (SORI-CAD) ESI-FTICR mass spectrum of the Mr = 1865 Da (i.e. m/z 933^{2+}) tryptic peptide.

The abundant *y*-type fragment ions reveal an amino acid sequence tag of **NSQDYFY** in this peptide.

sequence tag to high resolution as well (i.e. less than 2 millimass units), as illustrated in Table 3. This allowed unambiguous identification and differentiation of not only asparagine (N) from aspartic acid (D), but also resolution of glutamine (Q) from lysine (K), which differ by only 0.0364 Da. Thus, the high-resolution ESI-FTICR-MS measurement unambiguously revealed that the third amino acid in this sequence tag was glutamine (Q), and not lysine (K).

Searching the NCBI protein database against all taxa with this ***NSQDYFY*** sequence tag revealed only one known possibility for this sequence; the 11.8 kDa Sml1 protein from *Saccharomyces cerevisiae*. This sequence tag occurs in the N-terminal region (residues 3-9) of the wild type protein. The collisional dissociation data and the sequence tag identification from the protein database indicated that this 1865 Da peptide was located at the N-terminus of the Sml1-histag protein, and would correspond to residues 1-16. Accurate mass measurement and collisional dissociation data verified that the expected N-terminal methionine residue was absent from the Sml1-histag protein. Likewise, collisional dissociation of the nominal Mr = 4348 Da peptide (i.e. m/z 1088^{4+}) revealed a sequence tag of ***STQNSMG***. Searching the protein database with this sequence tag verified that this peptide also corresponded only to the yeast Sml1 protein, and was located in the interior region (residues 36-70). By identifying the peptide fragments and orienting them in the correct order, more than 94% of the sequence of the Sml1-histag protein could be verified. The small C-terminal fragment peptide (residues 111-117) was not observed, and may have been lost in the sample clean-up step. The peptide fragment corresponding to residues 17-27 was also absent from the ESI-FTICR

Table 3. Fragment ions from SORI-CAD of m/z 933^{2+} (1865 Da) peptide from trypsin digest

Measured Mass (Da)	Fragment Identity	Calculated Mass (Da)	Mass Diff. (Da)	Sequence Identity	(Δ m/z between y ions)
488.2505	y_4	488.2501	0.0004	(AQNR)[a]	
651.3134	y_5	651.3134	0.000[b]	m/z 488 + **Y**	(163.0629 Da meas.) (163.0633 Da calc.)
798.3816	y_6	798.3817	0.0001	m/z 651 + **F**	(147.0682 Da meas.) (147.0684 Da calc.)
961.4440	y_7	961.4449	0.0009	m/z 798 + **Y**	(163.0624 Da meas.) (163.0633 Da calc.)
1076.4715	y_8	1076.4718	0.0003	m/z 961 + **D**	(115.0275 Da meas.) (115.0269 Da calc.)
1204.5279	y_9	1204.5304	0.0025	m/z 1076 + **Q**	(128.0564 Da meas.) (128.0586 Da calc.)
1291.5624	y_{10}	1291.5624	0.000[b]	m/z 1205 + **S**	(87.0345 Da meas.) (87.0320 Da calc.)
1405.6034	y_{11}	1405.6052	0.0018	m/z 1292 + **N**	(114.0410 Da meas.) (114.0429 Da calc.)
			Sequence Tag Identified	NSQDYFY	

[a] Only total mass rather than the sequence is determined for this peptide.
[b] These ions were used as internal calibrants for this mass spectrum.

mass spectrum, but was identified in a MALDI-TOF experiment on the same digest sample. Table 2 reveals that the masses of the tryptic digest peptides measured in the mass spectra correspond within a few millimass units of the expected calculated values (as determined from the proposed Sml1-histag sequence of Figure 13, minus the N-terminal methionine). Inspection of this table reveals that the digestion was not carried out to completion, as evidenced by the appearance of the larger peptides corresponding to one or two missed trypsin cuts. By employing both accurate mass measurements and collisional dissociation experiments, it was possible to obtain unambiguous information about protein sequence, including several informative sequence tags, even with incomplete digestion. More importantly, this data provided unambiguous information about the regions in the Sml1-histag protein that had been modified relative to the wild type species.

The Sml1-histag protein was digested with Glu-C to attempt to resolve the ambiguity with the C-terminal end from the trypsin data. This Glu-C digest also revealed a variety of peptide fragments, as listed in Table 4, whose measured masses agreed very well with the calculated values. In this case the N-terminal peptide was not observed, but the C-terminal fragment was present. Thus, the typsin and Glu-C data are complementary, and unambiguously resolve the complete sequence of the Sml1-histag protein, which corresponds to the sequence proposed in Figure 13 without the N-terminal methionine. The absence of the N-terminal methionine residue was not unexpected, even though the DNA construct should have generated the sml1-histag protein with this residue included. Many proteins expressed in bacteria have their N-terminal methionines

Table 4. Peptides generated by glu-C digest of Sml1-histag protein

Measured Mass*(Da)	Residues	Calculated Mass*(Da)	Mass Diff.
Not observed in ESI-MS	1-33	3737.660	
3434.608	34-66	3434.662	0.054
949.417	67-73	949.421	0.004
1078.468	67-74	1078.464	0.004
3107.374	67-92	3107.370	0.004
4498.064	67-104	4498.061	0.003
2661.239	70-92	2661.208	0.031
2175.941	74-92	2175.959	0.018
2046.936	75-92	2046.917	0.019
1408.711	93-104	1408.701	0.010
1620.708	105-117	1620.686	0.022

* Monoisotopic masses
** Expected Glu-C peptides based on Sml1-histag sequence shown in Figure 13. (without the N-terminus methionine)

removed, as this is the most common post-translational modification in *E. coli* (Gonzales and Robert-Baudouy 1996). The sequence tags obtained from the collisional dissociation experiments verify that this is a protein related closely to the wild-type yeast Sml1.

There are seventeen basic amino acids (K,R,H) in this protein. The electrospray mass spectra (Figure 15B) reveal that 10–15 protons are added to this protein, indicating that almost all of the basic residues (and the N-terminus amino acid glycine) are protonated. The wild type sml1-p protein is known to have two alpha-helical regions, which would correspond to residues 8-18 and 65-84 in the Sml1-histag protein sequence shown in Figure 13. In general, NMR results suggest that the three-dimensional structure of Sml1-WT is best characterized as a loosely-folded tertiary structure in which the two main alpha helices are oriented in an antiparallel fashion (Zhao, Muller et al. 1998). This loosely-folded, relatively open structure may be reflected in the observation of protonation of almost every basic amino acid residue of the protein in the ESI-MS experiment.

Identification of multiple molecular species for the Sml1-histag monomer

The Sml1-histag protein sequence determined from the digest data yielded an average molecular mass of 13245.7 Da. This value is lower than any of the molecular values measured in Figure 15B. The digestion information in Tables 2 and 4 revealed the detailed sequence of the Sml1-histag protein, and indicated that, with the exception of the missing N-terminal methionine, no other amino acid variants relative to Figure 13 were observed. Collisional dissociation experiments provide a means of directly analyzing the

intact protein, and will be discussed below to resolve the discrepancy of the Sml1-histag protein mass suggested by the digestion experiments and the intact molecular mass measured by the ESI-MS experiment in Figure 15.

The gas phase fragmentation of the Sml1-histag protein was examined with collisional dissociation in the hexapole ion guide. This experiment was conducted by trapping all of the protein ions in the hexapole ion guide for 3100 msec. to induce fragmentation. The ESI-mass spectra generated under these conditions revealed substantial fragmentation, as evidenced by the appearance of additional ions with lower charge states than the intact molecular ions observed in Figure 15A. Deconvolution of the multiply-charged fragment ions, shown in Figure 17, revealed that these species correspond to characteristic *y*-type ions, some of which are identified on the scheme shown in the inset of this figure. Although the fragmentation of this intact protein does not reveal the complete amino acid sequence under these experimental conditions, all of the most abundant fragment ions observed are completely consistent with the protein sequence suggested by the digestion experiments (shown in Figure 17, inset). In fact, the average masses of these fragment ions corresponded to within 0.075 Da of the calculated fragment ion masses expected from the proposed sequence. Note that this collisional dissociation technique did *not* provide for isolation and fragmentation of a single protein species, but rather fragmented all of the protein forms present in Figure 15 simultaneously. It is interesting that even though a variety of molecular species are observed in the Mr = 13,200 – 13,500 Da region of Figure 15B, all the fragment ions observed in Figure 17 are consistent with the

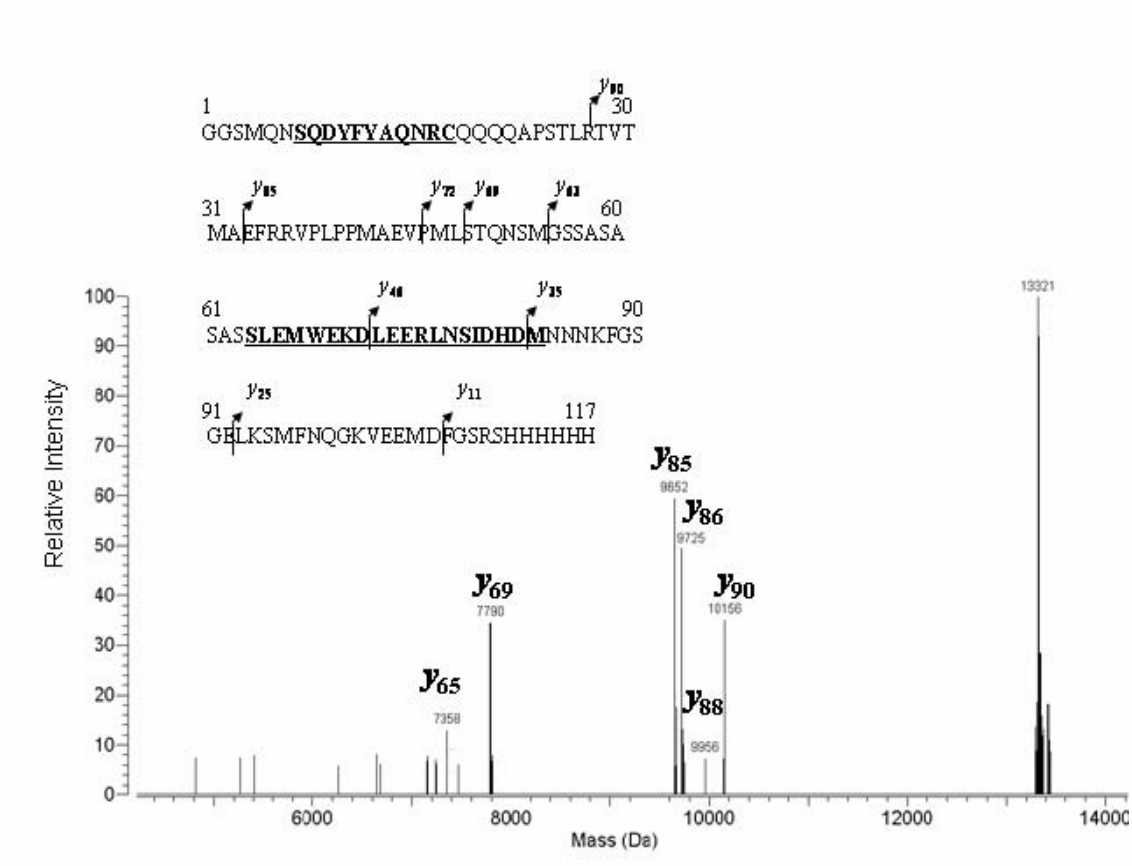

Figure 17. Positive ion collisional dissociation (extended hexapole accumulation) ESI-FTICR mass spectrum of Sml1-histag protein sample.

Even though all the distinct molecular species observed in Figure 15(B) were dissociated simultaneously in this experiment, the deconvoluted mass spectrum reveals fragment ions that are all consistent with the sequence determined by the proteolytic digestion experiments (inset reveals the identities of representative *y*-type fragment ions, based on the sequence obtained from the digestion). The bold, underlined regions in the sequence highlight the two alpha-helix regions of Sml1 identified by NMR (Zhao, Georgieva et al. 2000).

sequence obtained from the proteolytic digest data. This lends further support to the digestion data that the various protein species observed in Figure 15B are not sequence variants, but rather represent closely-related protein species that are slightly altered in total molecular mass.

Collisional dissociation experiments on the intact Sml1-histag proteins were also conducted by isolating specific molecular ions inside the FTICR-MS analyzer cell, and then inducing collisions with argon as a target gas. This provided a means of interrogating the dissociation of selective parent molecular ions *one at a time*, rather than the *all-at-once* hexapole CAD experiment. In general, the fragment ions observed for the dissociation of the two most abundant (13^+) ions at m/z 1024 (Mr = 13,292 Da) and 1027 (Mr = 13,321 Da) were similar to those observed for the hexapole dissociation shown in Figure 17. Even though the fragmentation in this case also was not very extensive, every fragment ion observed correlated with a *y*-type ion identified in the inset of Figure 17. This provided additional support for the overall proposed sequence.

The information obtained from the digestion and CAD experiments, along with high-resolution mass measurement of the intact molecular species in Figure 15B, suggest that the Sml1-histag protein species all consist of the amino acid sequence illustrated in the inset of Figure 17, but contain different levels of alkali metal (sodium, potassium) adducts. For example, even though some variation in the relative abundances of the molecular species were observed depending on purification and sample handling procedures, in general the most abundant molecular species was observed at average Mr=13,321 Da, corresponding to (Sml1-histag + 2K – 2H). The identity of the species at

average Mr= 13292 Da is slightly more difficult to determine unambiguously. This ion may correspond primarily to a fragment ion of the Mr = 13,321 Da species (loss of CO), although is it not possible to rule out some contribution from sodium-bound sml1-histag (Sml1-histag + 2Na – 2H). The other ions observed in Figure 15B would correspond to the molecular Sml1-histag protein with different combinations of sodium and/or potassium attached. The presence of salt in this histag protein is not unexpected, as the protein purification step involved the use of 300 mM sodium chloride. Likewise, the potassium may originate in the bacterial expression system, from contaminants in the sodium sample, or may be due to sample handling conditions. It appears that the Sml1-histag protein binds alkali metals quite strongly. For example, attempts to desalt this protein sample by dialysis against pure water were unsuccessful (i.e. the sodiated and potassiated versions were still observed), verifying the strong alkali metal-ion binding affinity of this species. None of the fragment ions, including the abundant y_{90} fragment ion at 10156 Da (which still contains the histidine "tail") observed in Figure 17, contain the alkali metals, indicating that the metal ions are eliminated once the protein fragments.[20] It is possible that the folded, intact Sml1-histag protein monomer has one or more strong metal-binding sites that disappear upon fragmentation of the protein into smaller peptides.

[20] Mass/charge range monitored in all the mass spectrometric experiments was from 300 to 2500, which was well above possible mass/charge of sodium or potassium ion. Therefore, these metal ions could not be observed.

Evidence for a disulfide-bonded Sml1-histag dimer

When the expressed Sml1-histag protein sample was analyzed with non-reducing SDS-PAGE, a band corresponding to about 27 kDa was observed (see lane 3 of Figure 14). Because this is approximately twice the molecular mass for the Sml1-histag monomer, it was suspected that this protein forms a covalent dimer under non-reducing condition, possibly via an intermolecular disulfide bond. ESI-MS was used to determine the identity of this dimeric species. For this particular experiment, Sml1-histag was purified in the absence of reducing agent.

The ESI-FTICR-MS of this dimeric species is shown in Figure 18, revealing abundant ions ranging from *nominal* m/z 1105 (24^+) to *nominal* m/z 914 (29^+). Even though these ions appear in the same region of the mass spectrum as the ions of the monomer (Figure 15A), the high-resolution capability of the ESI-FTICR-MS is sufficient to verify the higher charge states of the ions from the dimeric species. The amount of protonation observed in Figure 18 is approximately twice that observed for the monomer, suggesting that this dimer probably does not have a tightly bound higher order structure, but rather provides relatively open accessibility for protonation of the basic amino acid residues. The deconvoluted molecular mass spectrum, shown in the inset of Figure 18, reveals the presence of an abundant protein at average Mr=26,488 Da, which would correspond to the dimer of the expected monomer (Mr = 13245.6 Da) with concomitant loss of H_2. The appearance of an abundant Sml1-histag dimer without the ubiquitous metal ion adduction observed for the monomer may suggest that the metal ion binding

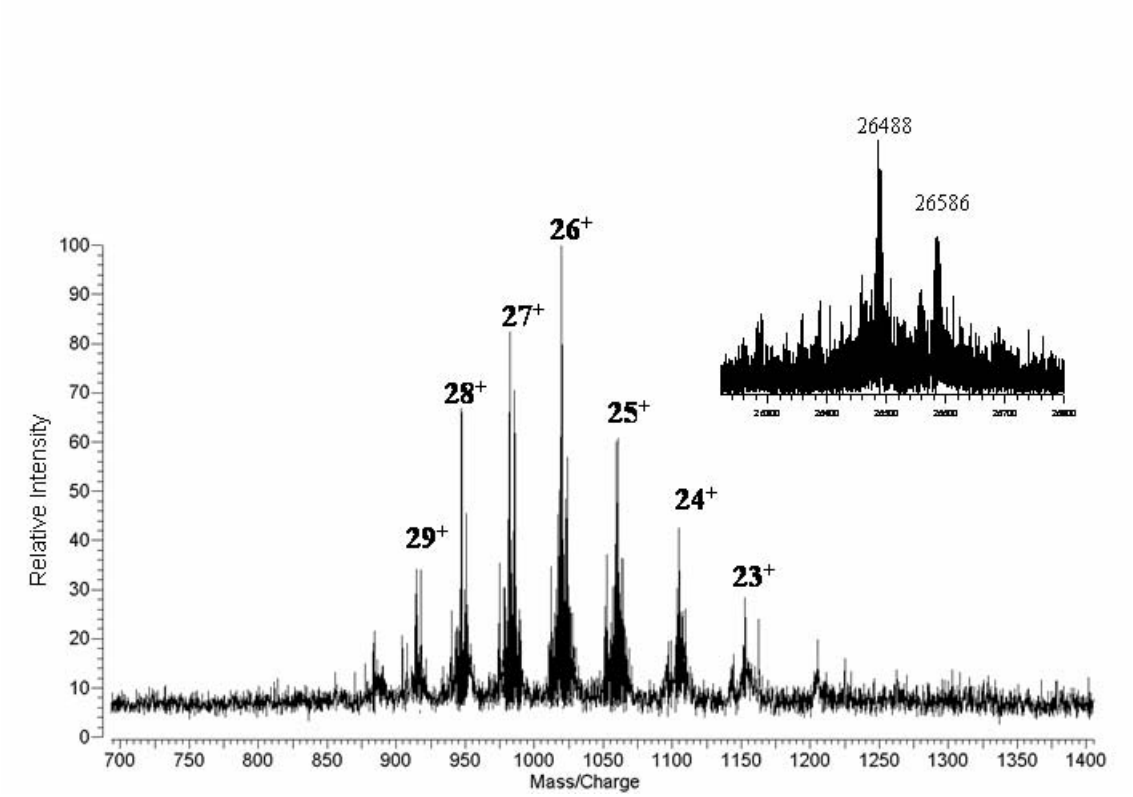

Figure 18. Positive ion ESI-FTICR mass spectrum of recombinant Sml1-histag protein purified under non-reducing conditions.

Multiply-charged molecular ions corresponding to dimeric Sml1-histag are shown. Inset shows deconvoluted molecular mass spectrum illustrating the presence of abundant dimeric species in this sample.

sites prevalent in the monomer are not retained in the dimer. No attempt was made to comprehensively examine this issue, as the monomer and dimer were purified under different experimental conditions. Based on the metal ion binding in the monomeric Sml1-histag, the species at Mr=26,586 Da in Figure 18 inset most likely corresponds to addition of multiple alkali metals (i.e. dimer + 2K + Na – 3H). These results suggest the presence of a single disulfide bond in the Sml1-histag dimer. Since the Sml1-histag protein monomer contains one cysteine residue at position 17 (see Figure 17, inset), it is likely that a disulfide bond might have formed between the cysteines of two monomers and produced a stable dimeric species. The application of more energetic electrospray ionization conditions (i.e. higher voltages in the ES source) and the addition of a higher amount of organic solvent (75% methanol: 25% water) did not diminish the amount of dimeric protein species observed, verifying that the dimers were strongly bound (as would be the case for a covalently-linked disulfide-bonded species).

Collisional dissociation of this dimer was undertaken to probe the fragmentation products in order to further verify the suspected structure of this species. This experiment was accomplished by trapping the dimer molecular ions in the hexapole accumulation region for times ranging from 1 sec. to 2.5 sec. (longer accumulation times result in more extensive fragmentation). Trapping the electrosprayed ions in the hexapole ion guide for times less than 700 msec. did not generate any appreciable fragmentation, but rather revealed mass spectra that were very similar to that shown in Figure 18. Increasing the hexapole accumulation time to 1 sec. revealed not only the molecular species observed in Figure 18, but also an abundant fragment ion at m/z 8130.

This is a y_{72} fragment ion, generated by cleaving the protein between residues 45 (valine) and 46 (proline), and corresponds to a peptide consisting of residues 46-117 of the original protein (see Figure 19, inset). Extensive fragmentation of the dimeric Sml1-histag protein was observed with a hexapole accumulation time of 1.9 sec. These fragment ions have a wide range of charge states, and thus require high-resolution mass measurement to identify and resolve unambiguously their identities. The deconvoluted mass spectrum (Figure 19) reveals a variety of characteristic fragment ions that range from y_{72} at m/z 8130 to y_{11} at m/z 1375 (refer to Figure 19 for identities). Closer inspection of the m/z 6400 – 7000 region reveals fragment ions that identify an amino acid sequence tag of **ASASAS**, as illustrated in the inset of Figure 19, which corresponds to residues 58-63 from the sequence proposed in Figure 17. In general, the fragmentation of the dimeric Sml1-histag protein reveals abundant *y*-type fragment ions from the C-terminal regions of the dimer, with virtually no fragment ions originating from the N-terminal section. While the presence of the disulfide bond at position 17 clearly reduces fragmentation in this region, the absence of dissociation on the N-terminal side of this bond suggests that there may be higher order structure (such as an interaction between the two suspected alpha-helices at residues 7-17 in each monomer) in this dimer which reduces fragmentation in the N-terminal region. The two C-terminal "tails" of the dimer may have a relatively open structure, as evidenced by the ease of fragmenting this region.

Trypsin digestion under non-reducing conditions of the Sm1lp-histag dimer was undertaken to confirm the presence of the disulfide bond. ESI-MS of this digest reveals similar peptide fragments as those observed for the monomeric species. Accurate mass

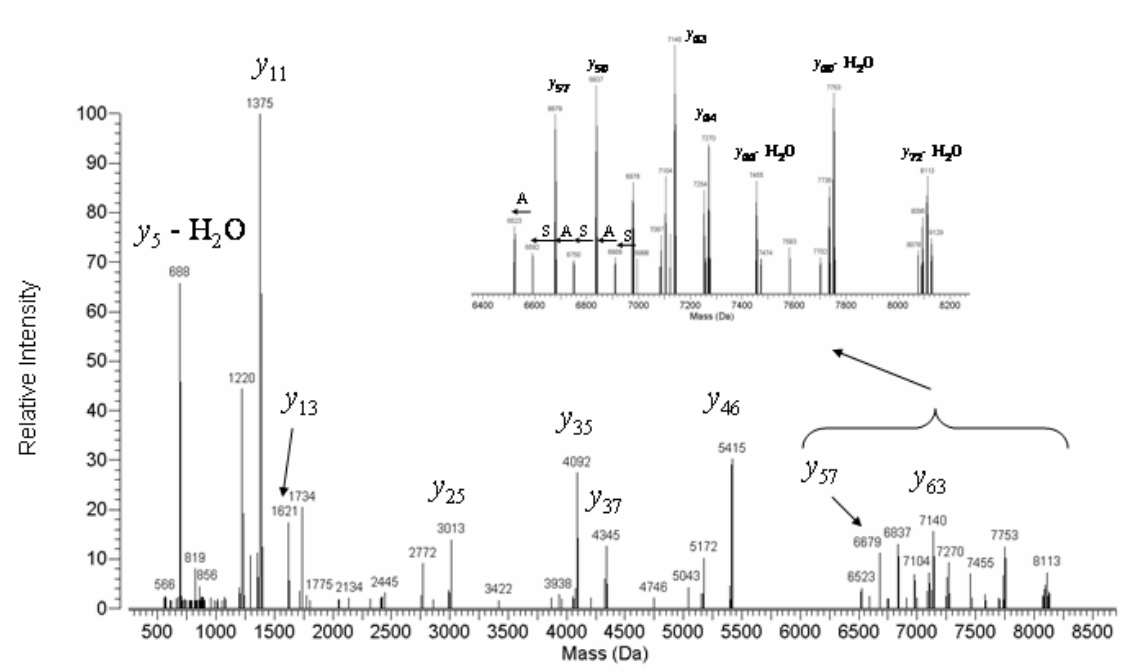

Figure 19. Positive ion collisional dissociation (extended hexapole accumulation) ESI-FTICR mass spectrum of Sml1-histag protein sample from Figure 18.

This decovoluted mass spectrum reveals fragment ions which are consistent with the protein sequence shown in Figure 17. Expansion of the m/z 6400-8400 region illustrates an amino acid sequence tag of **ASASAS** obtained from these y-type fragment ions.

measurement of the peptide fragments from the dimeric Sml1-histag yielded values that were each within 5 millimass units of the calculated values. One particular peptide fragment, m/z 839.4^{3+} (Mr = 2515.207 Da), was present in the digest of this Sml1-histag dimer that was absent in the digest of the monomer, and was of special interest, because it would correspond to the fragment containing the disulfide bond, as shown in Figure 20 (calculated Mr = 2515.216 Da). Collisional dissociation of this species revealed fragmentation supporting a partial amino acid sequence of **QQQAPSTLR**, corresponding to the expected region of residues 19-27 of the original Sml1-histag protein proposed in Figure 17, and verified that this was a disulfide-bonded dimer of tryptic fragment 17-27. Reduction of the Sml1-histag dimer protein with DTT prior to tryptic digestion, generated identical fragment peptides with the exception that the Mr = 2515 Da was absent in this case (data not shown).

The molecular mass measurements, along with the collisional dissociation and digestion data, verify that this Sml1-histag dimer consists of two Sml1-histag monomers joined by a single disulfide involving the cysteine residues present at position 17 in each monomer. The ease of forming this dimeric species suggests that it is important to keep the protein in a reducing environment when purifying the Sml1-histag proteins to avoid undesirable formation of the dimeric species.

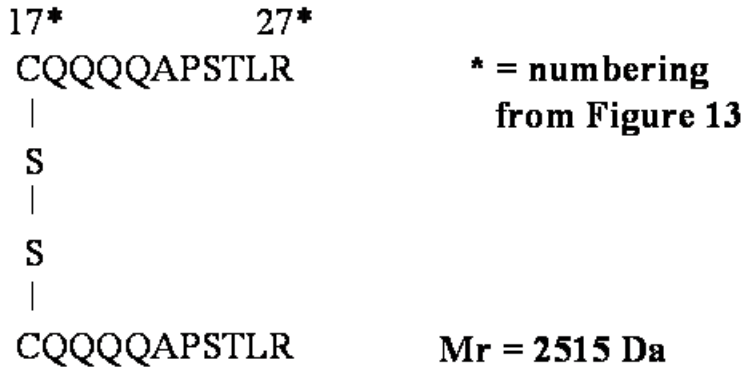

Figure 20. Sml1 peptide containing disulfide bond

Summary

This study demonstrates that recombinant proteins can be characterized at the molecular level with mass spectrometry, in particular ESI-FTICR-MS. The accurate mass measurement and ion interrogation capabilities of this technique were used to verify the sequence of a recombinant Sml1 protein and verified that the N-terminal methionine had been removed from the expressed protein. Strong alkali metal binding affinity for the Sml1 protein was determined from the collisional dissociation data and the molecular mass measurements. The ease of forming disulfide-linked dimeric Sml1 protein species under non-reducing conditions also was verified by ESI-FTICR-MS, confirming the C-14 as the site of disulfide bond attachment between two monomers. MS results were compared with gel electrophoresis data, and revealed the ability of MS to provide high-resolution identification and characterization of proteins and their variants.

Chapter 4: Characterization of Sml1's oligomerization state

Introduction

Based on SDS-PAGE and mass spectrometry, we have previously observed that Sml1-histag readily forms a homodimer via a disulfide bond between the lone cysteine groups of each monomer (Uchiki, Hettich et al. 2002). To further investigate the effect of the disulfide linkage on oligomerization of Sml1, a site-directed mutagenesis was performed to express a mutant Sml1 protein (C14S Sml1) in which the lone cysteine involved in intermolecular disulfide bond (Cys 14) was replaced into serine.

Previous studies have demonstrated that the site directed mutagenesis of residues involved in intermolecular interaction can affect both the oligomeric state and the activity of proteins (Cutler, Pielak et al. 1987; Amatayakul-Chantler, Qian et al. 1994; Olsen, Ludvigsen et al. 1996; Nickerson and Wong 1997). For example, in cytochrome $P450_{cam}$, substitution of cysteine 334, which forms an intermolecular disulfide bond, to alanine eliminates the presence of dimeric species and increases the overall activity of the protein (Nickerson and Wong 1997). Similarly site directed mutagenesis of the four glutamic acid involved in monomer-monomer interaction in insulin abolishes the dimer and hexamer species, resulting the mutant protein suitable for determination of 3D structure by NMR.

A critical component for site directed mutagenesis is the confirmation of the mutation, which is usually achieved by DNA sequencing. However, even a recombinant protein expressed in a bacterial host such as *E. coli* can be different from the one predicted from the DNA sequence due to N-terminus methione cleavage (Tchelet, Vogel

et al. 1997; Ge, Lawhorn et al. 2002; Uchiki, Hettich et al. 2002), formylation and acetylation (Bradshaw 1998) and other posttranslational modifications (Krishna and Wold 1993; Yan, Caldwell et al. 1999; Yan, Caldwell et al. 1999). In addition, due to sequencing errors and automated prediction of open reading frames, one needs to be cautious when analyzing protein sequences derived from the DNA sequence (Mann and Pandey 2001; Ge, Lawhorn et al. 2002). Therefore, examination of the *protein* rather than *DNA sequence* is more informative. We analyzed intact molecular masses of wild type and C14S Sml1, as well as masses of their tryptic peptides, by ESI-FTICR to verify their protein sequence. The high resolution achieved by ESI-FTICR unequivocally verified Cys to Ser mutation and also presence of disulfide linked dimer in wild type Sml1. Furthermore, we identified a truncated C14S Sml1 obtained during purification of full length C14S Sml1.

The goal of this study was to investigate the types of interactions and regions of Sml1 that are responsible for oligomerization. Measuring the size of a protein complex is a common approach to determine the stoichiometry of proteins in the complex. For this purpose, gel filtration chromatography can be used to determine the size of a protein complex with relative ease (Stevens 1989; Raffen and Stevens 1999). In this study, we analyzed oligomeric states of wild type and C14S Sml1 in solution by gel filtration chromatography. Our results showed that Sml1 forms an oligomer through non-covalent interaction. (Experimental procedures used in this study are described in Chapter 2.)

Results

Verification of Sml1 wild type and variant protein identities

When analyzed by SDS-PAGE under reducing conditions (i.e. DTT present), both wild type and C14S Sml1 showed apparent molecular masses of about 14 kDa. However, when examined under non-reducing SDS-PAGE conditions, wild type Sml1 migrated slower, with an apparent molecular mass of about 24 kDa, as shown in Figure 21. On the other hand, C14S Sml1 did not exhibit an altered shift in migration. These results suggest that wild type Sml1 forms a covalently-linked dimer under non-reducing conditions (most likely due to a disulfide bond), while C14S Sml1 exists as monomer in both reducing and non-reducing conditions.

Mass spectrometry was used to measure the molecular masses of these proteins with high-resolution to verify their molecular identities. When purified in the presence of DTT, wild type Sml1 showed only a monomeric species (Figure 22A). On the other hand, in the absence of DTT, primarily dimeric species of wild type Sml1 was observed (Figure 22B). As expected, C14S Sml1 consistently showed only monomeric species regardless of the presence or absence of DTT (Figure 22C). Not surprisingly, mass of both wild type and C14S Sml1 indicated that N-terminal methionine of these proteins was removed. It has been reported that N-terminal methionine elimination is the most common form of bacterial post-translational processing (Gonzales and Robert-Baudouy 1996). Table 5 illustrates that the measured molecular masses of the wild-type and C14S Sml1 correspond to within ~10 ppm of these adjusted molecular masses (i.e. with methionine truncation).

Table 5. Intact molecular masses of wild type and variant Sml1

	Observed mass*	Expected mass*	Accuracy (ppm)
Wild type Sml1 Monomer	11977.653	11977.496	13
Wild type Sml1 Disulfide dimer	23954.036	23953.979	2
C14S Sml1	11961.625	11961.519	9

* Most abundant isotopic mass of protein (including loss N-terminus methionine)

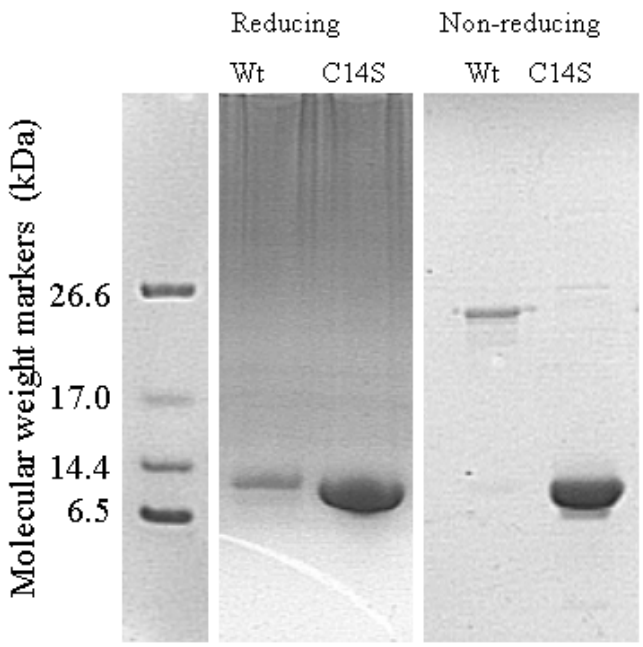

Figure 21. SDS-PAGE of Sml1 proteins conducted in reducing and non-reducing conditions.

Covalent dimer was observed in wild type Sml1 in non-reducing condition, while only monomer was observed in C14S Sml1.

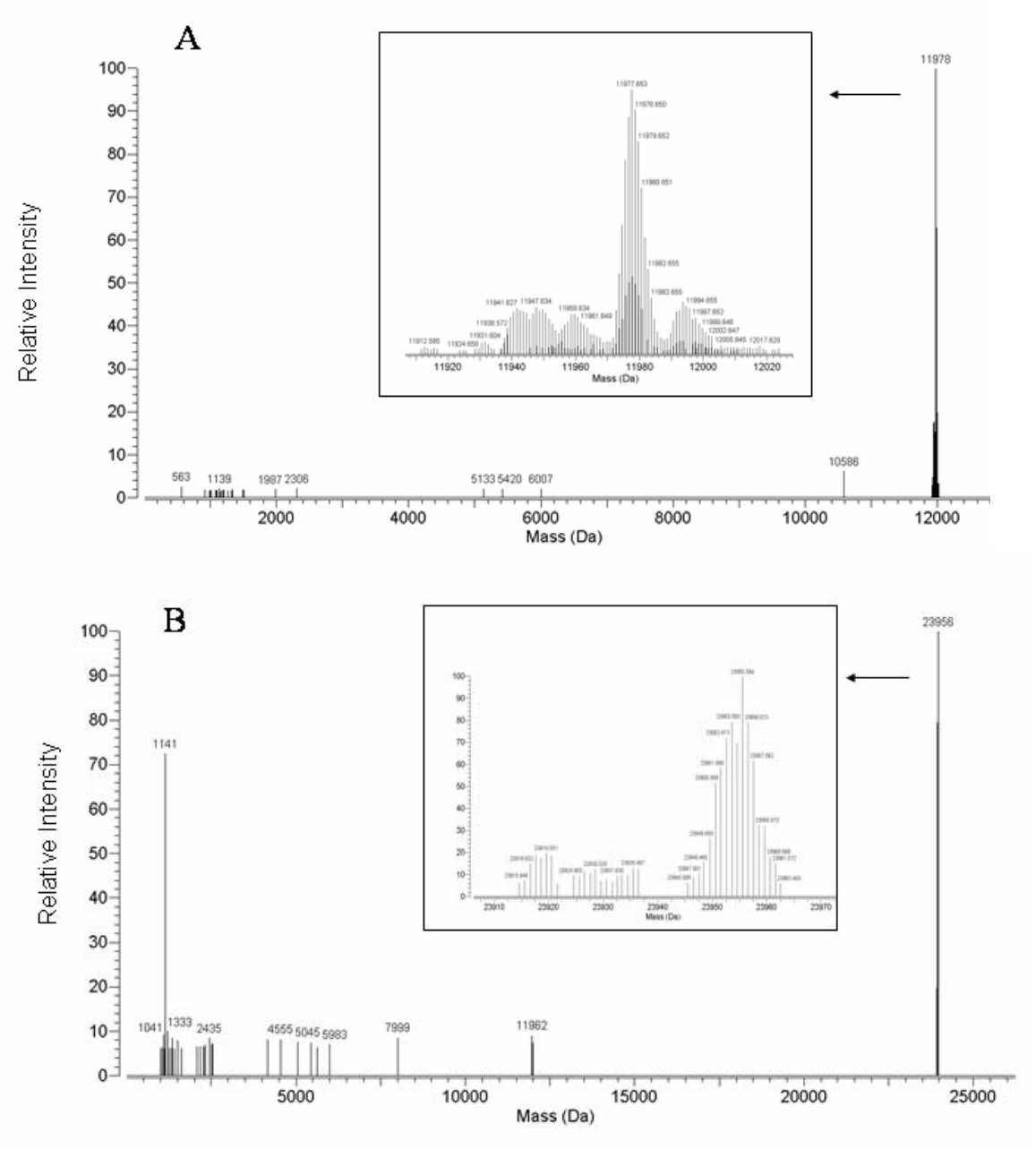

Figure 22. Deconvoluted mass spectrum of wild type and C14S Sml1.

(A) Wild type Sml1 purified in presence of DTT. Mainly monomer was observed in this sample. High resolution spectrum (inset) showed five to six envelopes of isomers, which corresponds to Na^+/ K^+ adduct and loss of CO_2 / H_2O. **(B)** Wild type Sml1 purified in absence of DTT. Mainly disulfide-linked dimer was observed in this sample. Like the monomer, five to six envelopes of isomers were observed in high-resolution spectrum (inset).

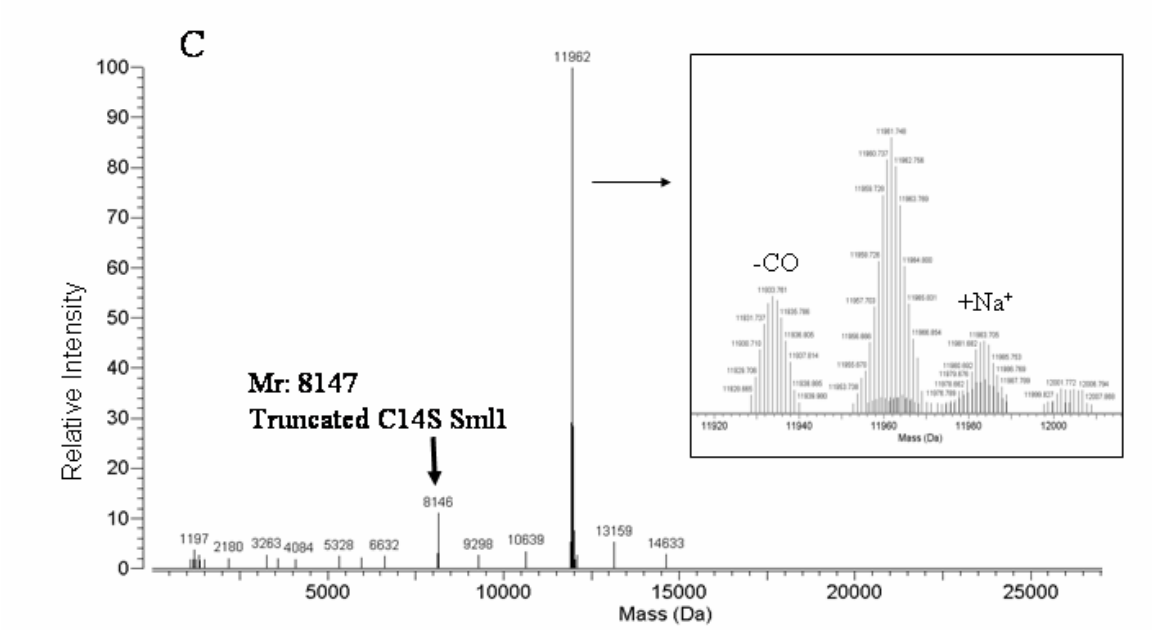

Figure 22 Continued

(C) C14S Sml1. Only monomer was observed in the samples purified in presence and absence (data not shown) of DTT. As was seen in wild type, isomers were also observed in C14S Sml1 (inset). A low abundant peak of 8147 is truncation product at C terminus end of residue 71E.

Closer inspection of mass spectra of intact proteins (Figure 22) reveals that, in addition to the species corresponding to Sml1 sequences, there are clusters of other species with close molecular masses. Most of these species correspond to addition of sodium, potassium, or both to the intact proteins. Although we could not assign some of these clusters, they are likely isomeric species of Sml1 that arose from addition or loss of chemical group during sample preparation.

SORI-CAD of the wild type and C14S Sml1 proteins generated y- and b-type fragment ions characteristic to these proteins (Figure 23). In both proteins, there was extensive fragmentation in the region from residue 42 to 55[21] under a wide range of collision energies (and different charge states). In wild type Sml1, the N-terminus end of the protein also fragmented extensively and several y or b type ions were observed. On the other hand, only two y type ions were identified from N-terminus end of C14S Sml1. For both proteins, the fragmentation was entirely consistent with the expected sequences. Unlike intact Sml1, no fragments corresponding to Na^+ or K^+ adduct were observed. The disulfide linked wild type Sml1 dimer was also examined by SORI-CAD. Over 20 fragment ions ranging from 6000 to 9000 Da were observed. Two abundant fragment ions at 6541 and 6920 Da were detected in all experiments conducted on the dimer. However, we could not identify these fragments. Overall, we could assign only two fragments of 5928 and 8722Da could as y53 and y78 (Data not shown).

---

[21] From Chpter 4 to Appendix, residues of Sml1 are numbered based on Sml1 ORF. Three residues, Met-Gly-Ser, are added to the N-terminus of the protein by the expression plasmid. These three residues (Met, Gly and Ser) are numbered –3, -2 and –1 respectively.

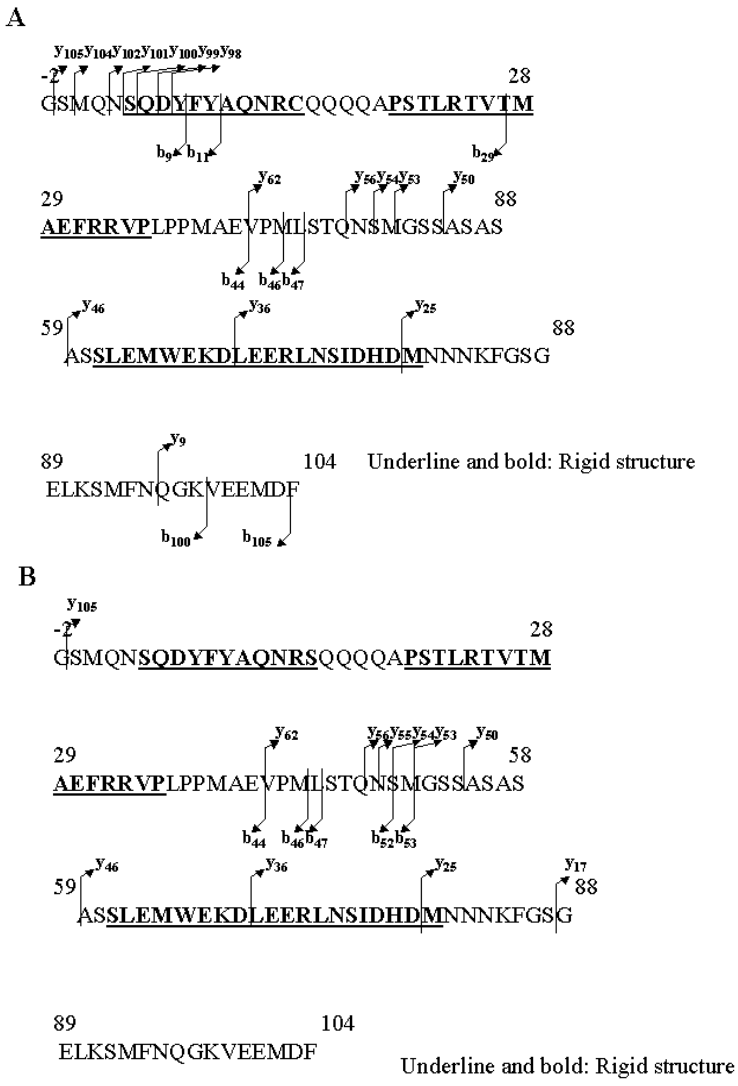

Figure 23. SORI-CAD fragmentation patterns of intact Sml1

(A) Wild type Sml1. **(B)** C14S Sml1. Fragment was assigned based on calculated mass of b and y type ions. The line across the sequence represents sites of fragmentation. Underline represents regions reported to have a high degree of backbone order (Zhao, Georgieva et al. 2000). Residues are numbered based on Sml1 ORF. Two residues added to the N-terminus (Gly-Ser) are numbered –2 and –1 respectively

To further verify the protein species, the wild type and C14S Sml1 were digested with trypsin and the tryptic peptides were analyzed by ESI-FTICR-MS. In both wild type and C14S Sml1, we observed tryptic fragments that covered the entire sequence of these proteins, as summarized in Table 6. In the analysis of both wild type and C14S Sml1, mass accuracies for the most of peptides were less than 3 ppm.

Inspection of tryptic peptide sample reveals that all the peptides for the wild-type and C14S Sml1 are identical (Figure 24), with the lone exception of the fragment 2 (residues 14-24). This is expected, since this is the only peptide that should have the amino acid substitution, and can be used for confirmation of the mutation. Residue 14 of the wild type Sml1 is cysteine, while that of C14S Sml1 is serine, and they differ in mass by 15.994 Da. We conducted CAD experiments of this particular peptide to further localize the site of the mutation. In both C14S and wild type Sml1, doubly charged ions of this peptide were subjected to SORI-CAD. The tryptic fragment of wild type and C14S Sml1 generated similar y type or b type ions. From y5, y6, y7 and y8 ions, a short sequence QQA was identified (Figure 25 A and B).

A species with monoisotopic mass 2515Da was observed in the wild type Sml1 digest, but completely absent from the C14S Sml1 digest. With a mass accuracy of 9.54 ppm, this species was identified as disulfide liked dimer of tryptic peptide from residue 14 to 24 (Table 6 bottom row). This disulfide linked peptide was also subjected to SORI-CAD. We chose ion of 840^{3+} as parent ion. 11 fragments were observed and five of them were identified as b or y type ion (Figure 25C). No fragmentation occurred at residues close to N-terminus end of the peptide, where disulfide bond is formed.

Table 6. Peptides generated by tryptic digestion of wild type Sml1 and C14S Sml1.

Residues*	Wild-type Sml1 measured mass (Da)	C14S Sml1 measured mass (Da)	Expected mass (Da)
85-91	736.377**	736.380**	736.375**
99-104	768.302	768.302	768.300
92-98	810.370	810.372	810.369
25-32	953.466	953.462	953.464
14-24	--	1242.631	1242.631
14-24	1258.609	--	1258.608
73-84	1413.631	1413.641	1413.630
92-104	1560.663	1560.664	1560.658
-2-13	1807.759	1807.773	1807.758
73-92	2132.000	2132.006	2131.990
34-67	3549.652	3549.656	3549.648
33-67	3705.744	3705.781	3705.749
34-72	4193.946**	4193.938**	4193.952**
33-72	4350.089	4350.041	4350.053
33-92	6466.043	6466.045	6466.042
14-24 S-S	2515.177	--	2515.201

* Residues are numbered based on ORF of Sml1. Two residues added to the N-terminus (Gly-Ser) are numbered –2 and –1 respectively

**For peptides of mass below 4000 Da, monoisotopic masses are listed. For peptides of mass above 4000 Da, the masses of the most abundant isotopic species are listed.

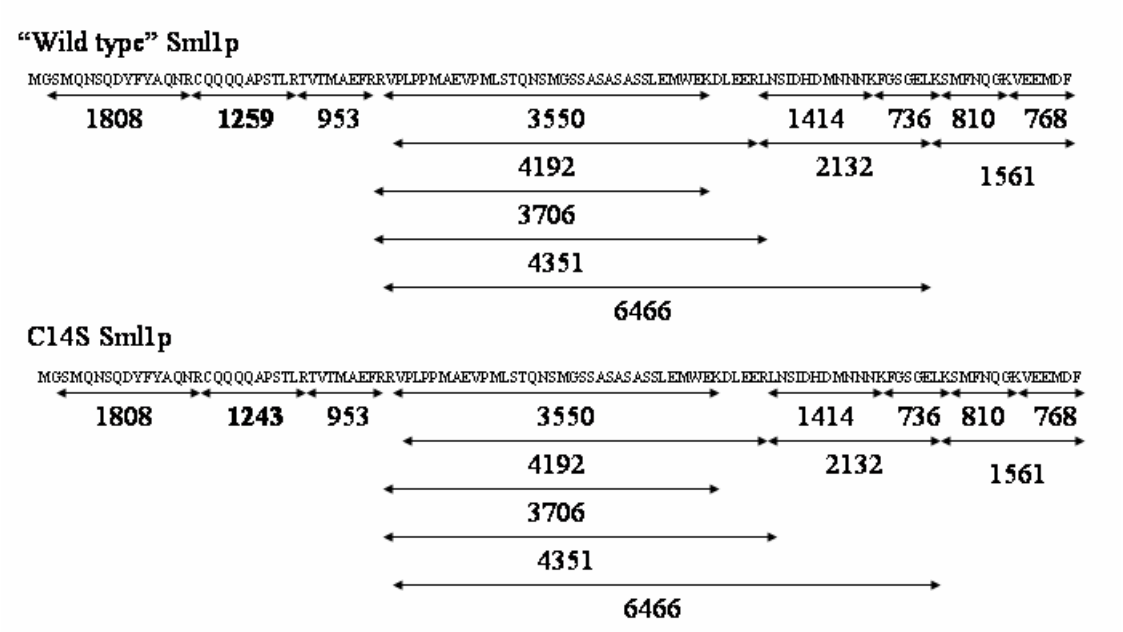

Figure 24. Tryptic digest fragments of wild type and C14S Sml1 observed by ESI-FTICR.

Double headed arrows below the amino acid sequence represents the region covered by each tryptic fragment. The numbers below the arrow are the nominal mass of the fragments.

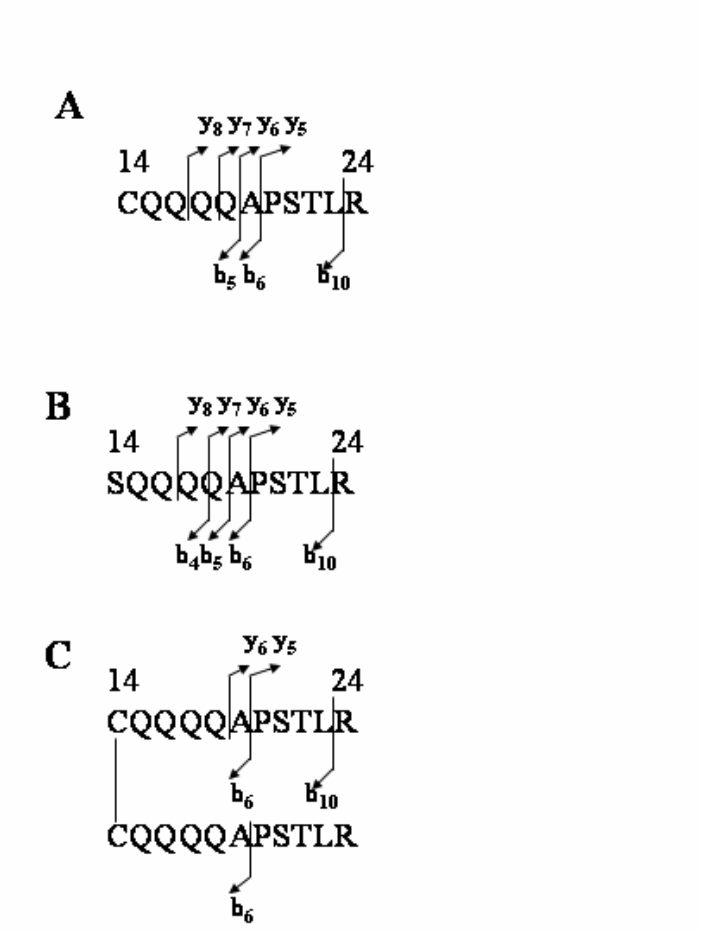

Figure 25. SORI-CAD fragmentation patterns of tryptic fragment spanning from residue 14 to 24.

(A) Wild type Sml1. **(B)** C14S Sml1. **(C)** wild type Sml1 in non-reducing condition. Fragment was assigned based on calculated mass of b and y type ions. The line across the sequence represents sites of fragmentation.

Identification of truncation products of C14S Sml1

In addition to intact protein species, we observed some minor species with mass ranging from 500-11,000 Da, and suspected that some of these were truncated products of Sml1. Particularly in C14S Sml1, a species of 8147 Da was consistently detected in several samples prepared independently. During purification of C14S Sml1 by gel filtration chromatography (Superdex 75), we noticed presence of a shoulder on the right side of the peak corresponding to C14S Sml1 (Figure 26A). SDS-PAGE analysis showed that this shoulder contains a protein species slightly smaller than C14S Sml1, as well as intact C14S Sml1. This shoulder was also analyzed by ESI-FTICR and, in this case the relative abundance of the species of 8147Da was higher (Figure 26B). Like intact C14S Sml1, this species also showed multiple envelopes of isomers corresponding to loss of CO or addition of alkali metals. We suspected this species to be a proteolytic fragment of C14S Sml1 generated by cleavage at C-terminus of residue 71 Glu. The calculated mass of this truncated species is 8146.781 Da (the most abundant isotopic mass), which agreed with the observed mass at 8146.923 Da (17 ppm error).

To further investigate the identity of this species, we conducted a SORI-CAD experiment. The parent molecular ion at m/z 1359 $(M+6H)^{6+}$ was isolated and fragmented with a range of collisional energies. A series of fragments corresponding to masses from 7600 to 8120 Da were generated. Most of these fragments were a, a*, b, b* or y" type ion derived from fragmentation at C-terminus of the C14S Sml1 truncation product. Through the CAD experiment, a sequence from the C-terminus of this species, EKDLE was identified (Figure 26C).

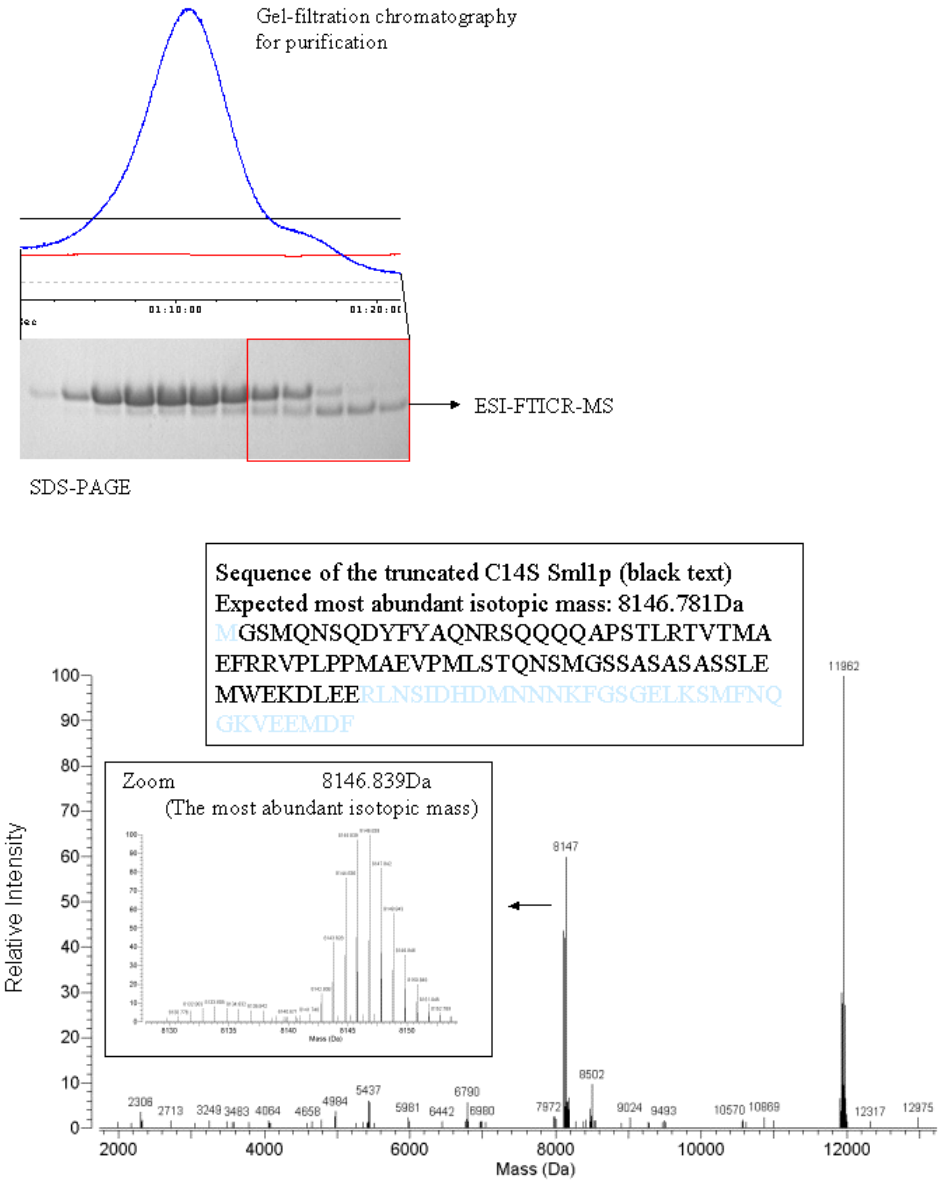

Figure 26. Fragment of C14S by cleavage at C-terminus of E71

(A) Chromatogram of C14S Sml1 purification. On SDS-PAGE gel, two bands were observed from the fractions that contained C14S Sml1. Fast migrating band was abundant in early fractions, while slowly migrating bands was abundant in later fractions
(B) Deconvoluted mass spectrum obtained from late fractions of Superdex 75 eluent (last five lanes on SDS-PAGE gel of (A)). Species with 8147Da showed higher abundance than intact C14S Sml1. The molecular mass, 8147Da, corresponds to cleavage at C-terminus end of E71. This species also showed multiple envelopes corresponding to chemical isomers (inset).

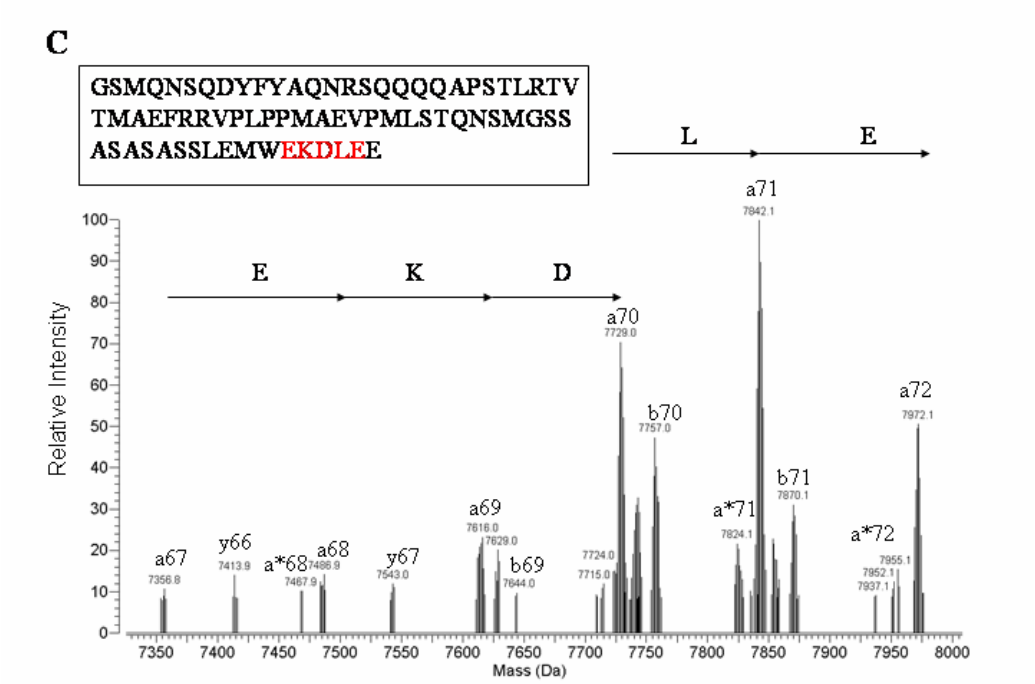

Figure 26 continued

(C) Deconvoluted mass spectrum of CAD fragments generated from the species of 8147Da (m/z 1359^{6+}). Extensive fragmentation at N terminus of the truncation product produced, a, a*, b, b* and y" ions. A sequence tag EKDLE was identified.

Oligomeric state of wild type and C14S Sml1

To analyze the oligomeric states of the proteins in solution, the wild type and C14S Sml1 proteins were subjected to gel filtration chromatography (using Superose 12 HR10-30). In the presence of DTT, both wild type and C14S Sml1 showed peaks at the elution volume corresponding to 32kDa (Table 7). In non-reducing conditions, the major peaks of wild type and C14S Sml1 also showed the same retention time corresponding to 32kDa. However, in non-reducing condition a shoulder corresponding to ~50kDa was also observed in wild type Sml1 (Figure 27), while peak profile of C14S Sml1 was essentially identical in non-reducing and reducing conditions.

Discussion

Integrities of recombinant Sml1 wild type and variant proteins

The point mutation that replaces residue Cys14 of wild type Sml1 with serine was confirmed based on intact proteins and tryptic fragments determined by ESI-FTICR-MS. The observed mass of wild type and C14S Sml1 proteins differs by 16.028 Da, closely corresponding to the calculated mass difference between cysteine and serine, which is 15.977 Da. Similarly, when the tryptic fragments were analyzed, the mass difference between the peptides of wild type and C14S Sml1 spanning residues 14 to 24 was 15.978 Da, demonstrating excellent agreement with the calculated mass difference. A SORI-CAD experiment of the tryptic peptides of residue 14 to 24 also confirmed the mutation (data not shown).

Table 7. Gel filtration chromatography elution profiles of wild type and C14S Sml1

	Retention time (min)	Estimated molecular weight (kDa)
Reducing condition		
Wild type Sml1	13.7±0.0	32.5±0.2
C14S Sml1	13.7±0.0	32.7±0.3
Non-reducing condition		
Wild type Sml1	14.1±0.1	32.0±1.1
Wild type Sml1 shoulder	12.6±0.1	49.9±0.9
C14S Sml1	14.1±0.1	32.0±0.1

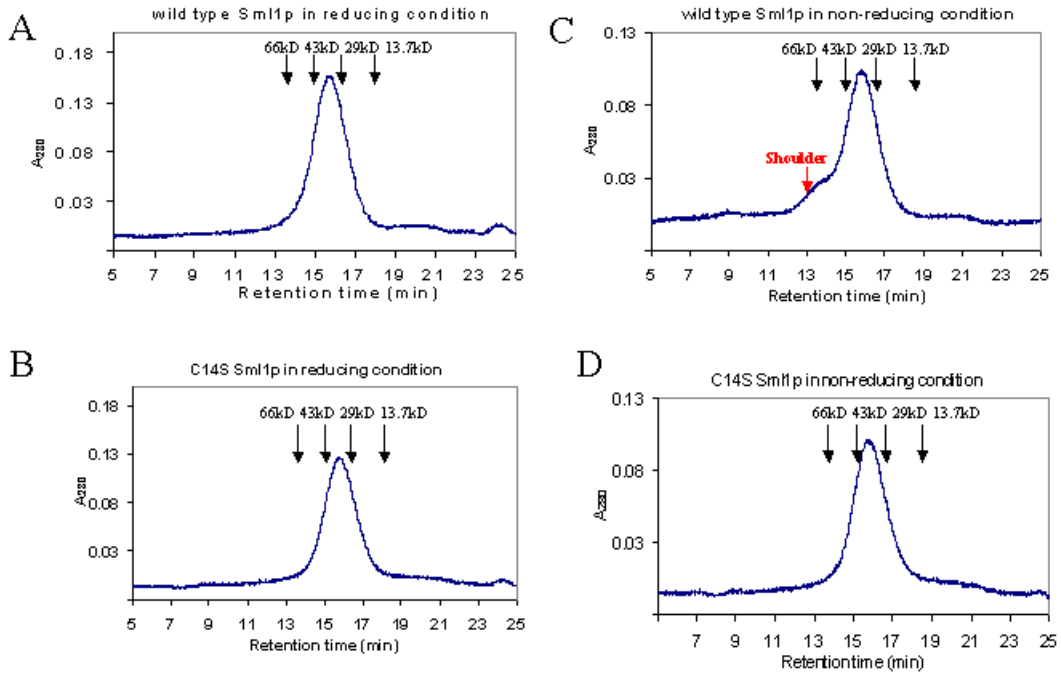

Figure 27. Gel filtration chromatography (Superose 12 molecular sieving column) of WT-Sml1 and C14S Sml1.

Samples were prepared at ~2 mg/ml in pure water. Four experimental conditions were examined. **(A)** WT-Sml1 under reducing conditions (DTT present) **(B)** WT-Sml1 under non-reducing conditions (DTT absent) **(C)** C14S Sml1 under reducing conditions (DTT present), and **(D)** C14S-Sml1 under non-reducing conditions (DTT absent). Note that the appearance of C14S-Smlp is unaltered in either condition, whereas WT-Sml1 reveals a higher molecular mass "shoulder" under non-reducing conditions, implying the formation of a larger oligomeric species.

The methionine at N-terminus of both wild type and C14S Sml1 was absent in the purified proteins. The masses of the intact wild type and C14S Sml1 were 131Da (130.885 for wild type and 130.936 for C14S) less than calculated mass based on their gene sequence. This mass difference corresponds to the loss of N-terminus methionine (calculated loss is 131.42Da). In the tryptic digest, a peptide of 1807.7Da in both wild type and C14S Sml1 sample corresponds to the mass of peptide spanning residues 1 to 14 (Table 6), verifying the loss of methionine at the N-terminus.

SORI-CAD fragmentation patterns of the intact wild type and C14S Sml1 can possibly reflect their higher order structure. In both proteins, the region from residue 42 to 55 fragmented extensively, and almost the same b type and y type ions were identified from the two proteins (Figure 23). This extensive fragmentation is probably due to the higher order structure of Sml1. Based on the report from Rothstein and co-workers (Zhao, Georgieva et al. 2000), this region does not have rigid structure. It has been reported that SORI-CAD fragmentation pattern reflects higher order structure of protein (Wu, Van Orden et al. 1995).

The molecular mass measurement and SORI-CAD fragmentation data indicated that the species of 8147 Da was a truncation product generated by cleavage at the C-terminus end of residue 71 glutamic acid. This species possibly corresponds to a minor species migrating slightly faster than C14S Sml1 on SDS-PAGE gel (data not shown). The same minor species was observed when whole cell lysate was subjected to an SDS-PAGE analysis (Data not shown). Heterogeneity of recombinant protein due to proteolytic cleavage has been reported (Frimpong, Darnay et al. 1993; Hejazi, Piotukh et

al. 2002). This minor species were likely to be produced by proteolytic cleavage of C14S Sml1. We also observed similar minor protein species in the sample of wild type Sml1, although we could not determine whether this species was fragment of wild type Sml1.

Anomalous species of wild type and C14S Sml1

In mass spectrometric analysis of intact proteins, wild type and C14S Sml1 are associated minor species possibly due to alkali metal adducts. The most intense peaks of wild type and C14S Sml1 in mass spectrometry matched expected molecular mass of each protein. However, high resolution spectrum of these proteins showed five to six envelopes of minor species whose molecular mass differ from the expected molecular mass in 15 to 100Da (Figure 22A and B, inset). Similar envelopes were also observed in a disulfide-linked dimer of wild type Sml1 (Figure 22C, inset). Some of these species were consistently found in different samples, while some were observed in only particular samples. Although we resolved isotopic peaks of these species, determining the average mass of these species were difficult because envelopes of these species are closely spaced and the sizes of the envelopes were relatively small. However, based on previous observation from Sml1 histag (Uchiki, Hettich et al. 2002), we believe that these species were possibly due to combination of Na^+ / K^+ adduct and the loss of some groups such as water, ammonia, and carbon monoxide. In addition, once proteins were fragmented by tryptic digestion or SORI-CAD, most of these minor species disappeared, indicating that these minor species are due to non-covalently bound adducts.

It should be noted such anomalous peaks were not observed in both tryptic fragments and fragments generated by CAD experiments. It is possible that a binding pocket of the alkali metal exists in the intact proteins. Once proteins are digested into peptide, they might lose the property to create such a binding pocket. We also tried to dissociate only the metal ion from the protein without breaking the peptide bond by SORI-CAD and in source decay in the hexapole ion guide. However, our attempt with lower energy in SORI-CAD or longer accumulation time in the hexapole ion guide of the ion inlet system did not yield intact Sml1 without the anomalous species.

Covalent dimer formation in wild type Sml1

A disulfide linked dimer of intact wild type Sml1 was observed by mass spectrometry. When purified in the absence of reducing agents, the mass of major species in wild type Sml1 was 23954Da, corresponding to its disulfide linked dimer (11978 Da x 2- 2H). On the other hand, only monomer was observed in C14S Sml1, whether it was prepared in the absence or presence of reducing agents. The same observation was made in SDS-PAGE under non-reducing condition. A similar observation was made when DTT was removed from wild type Sml1 sample originally purified with DTT and left in 4°C over a few days. Since the intracellular milieu is a reducing environment, the disulfide linkage between Sml1 monomer was unlikely to be formed when the protein is in the cell. Nevertheless, the formation of the disulfide linkage in a relatively short period of time suggests that Cys 14 of two Sml1 molecules are located in a proximity in solution. This

was further supported by our recent study (Gupta, Peterson et al. 2004) which shows that Sml1 dimerization occurs within a region spanning residues 8 to 20.

Disulfide linked peptides spanning residues 14 to 24 were also observed. When a tryptic peptide sample of wild type Sml1 was left for a few days in 4°C, a species with nominal mass of 2515Da appeared in mass spectrum (Table 6 bottom row). On the other hand, this species was not observed in tryptic peptides of C14S Sml1. This species was possibly due to a disulfide-linked dimer of peptide corresponding to residue 14 to 24. In this case, the disulfide bond was probably formed after digestion because purification and digestion was carried out in presence of 5mM DTT. It is possible that DTT was oxidized after digestion, and the non-reducing conditions allowed formation of the disulfide linkage. Similarly, the same disulfide linked species appeared in the peptides desalted by Zip Tip and kept in 50% acetonitrile in 4°C for a few days.

Tandem mass spectrometry (SORI-CAD) was also conducted with this disulfide linked peptide, and it showed fragmentation unique to the disulfide linked peptide. Among eleven fragments, five of them were identified as y type or b type ions of the disulfide linked peptides. For example, the fragment of 2341Da corresponds to b type ion generated by fragmentation at residue 23 on one peptide tail (Figure 25). Similarly, the fragment of 1943Da corresponds to b-type ion generated by fragmentation at residue 19 on one peptide tail. The fragment of 1371Da corresponded to b-type generated by fragmentation at residue 19 two peptide tails. These observations showed that fragmentation of disulfide-linker peptides occurs on either one or both of the peptide tail.

Non- covalent oligomer formation of wild type and variant Sml1

Analysis of purified wild type and C14S Sml1 by gel filtration chromatography (using Superose 12 HR10-30) showed oligomers with apparent size of 32kDa. These species were possibly dimers (12 kDa x 2 = 24 kDa) or trimers (12 kDa x 3 = 36 kDa) of Sml1. In addition, the symmetric peak of these proteins observed in reducing conditions indicates the presence of a single oligomeric species. In many cases, the symmetric peak of a protein complex in gel filtration chromatography indicates that the rate for the complex to dissociate into smaller subunits is slow and the dissociation does not occur during the chromatographic run (Raffen and Stevens 1999). Furthermore, the retention volumes of wild type and C14S Sml1 were essentially the same in non-reducing conditions. On the other hand, wild type Sml1 appears as disulfide-linked dimer in mass spectrometry, while C14S Sml1 cannot form disulfide linkage. These data indicate that C14S Sml1 can form an oligomer in solution through non-covalent interactions. It is unlikely that the monomer (12kDa) of C14S Sml1 and disulfide-linked dimer of wild type Sml1 (24kDa) elute at the same retention volume. These results are also consistent with our recent studies by sedimentation equilibrium (Gupta, Peterson et al. 2004) which showed that Sml1 can form non-covalent dimers. On the other hand, the shoulder of wild type Sml1 form a larger oligomer. Possibly, disulfide bonding of wild type Sml1 contributes to the formation of the larger oligomer.

Further information regarding Sml1 oligomerization was obtained by a combination of mass spectrometric identification of the truncated C14S Sml1 and

retention of the truncated C14S Sml1 on gel filtration chromatography. In addition to analysis of the purified proteins by Superose 12 column, we estimated the molecular mass of proteins eluted from Superdex G75 column during purification of C14S Sml1. As described above, a minor species of 8147Da was identified as a truncated C14S Sml1 spanning residues 1 to 71 by mass spectrometry. On the chromatogram of C14S Sml1 on Superdex G75 column, this truncated C14S Sml1 appeared as a small peak overlapping with the major peak of intact C14S Sml1 (Figure 26A). Based on its retention volume on Superdex G75 column, the size of the truncated C14S Sml1 in solution was estimated to be 14 to 16kDa, corresponding to its dimeric form (its expected size is 16294Da). In addition, no protein was detected from the fraction corresponding to 8 to 9kD. This indicates that the region responsible for oligomerization of C14S Sml1 may reside within residues 1 and 71, and C terminus end may not be involved in the dimerization. This is consistent with our recent studies with sedimentation equilibrium (Gupta, Peterson et al. 2004) that demonstrated that a region lying between residues 8 to 20 is responsible for dimerization of Sml1.

Summary

This study demonstrates the characterization of wild type and C14S Sml1 at molecular level by mass spectrometry at high-resolution and high mass accuracy (ESI-FTICR-MS). First, analysis of intact protein and tryptic peptide confirmed that the purified proteins were wild type and C14S Sml1. Mass spectrometric analysis of the intact protein also showed anomalous Sml1 variants that is possibly generated by alkali metal adduct on the

protein. By a combination of mass spectrometry and gel filtration chromatography, we showed that Sml1 can form dimers through both disulfide linkage and non-covalent interaction. Furthermore, a C14S Sml1 fragment truncated at residue 71 was identified. The retention profile of the truncated C14S Sml1 in gel filtration chromatography indicated that the non-covalent interaction occur in a region between residue 1 to 71.

Chapter 5. Identification of phosphorylation sites on Sml1 (Uchiki, Dice et al. 2004)

Introduction

As discussed in Chapter 1, the knowledge of specific phosphorylation sites on Sml1 provides clues for specificity of the Dun1 kinase, and such information is also useful when studying other substrates of Dun1. Ideally, phosphorylation of Sml1 *in vivo*[22] is biologically more relevant than phosphorylation *in vitro*. However, our initial trial of Sml1 purification indicated that it would be difficult to prepare a sufficient amount of Sml1 expressed in *S.cerevisiae* for identifying the phosphorylation sites. Additionally, because Sml1 is degraded *in vivo* after its phosphorylation, it was suspected that it would be even more challenging to obtain Sml1 phosphorylated *in vivo*. Therefore, we focused on an *in vitro* analysis. Descriptions in this chapter was mainly taken from Results and Discussion section of the article we published in March 2004 (Uchiki, Dice et al. 2004). Some new data and additional discussion have been included. (Experimental procedures used in this study are described in Chapter 2.)

Results

Results from the *in vitro* phosphorylation of Sml1 by Dun1 where ^{32}P is incorporated are shown in Figure 28. Sml1 is specifically phosphorylated by Dun1 with ^{32}P-incorporation

---

[22] *In vivo* Sml1 phosphorylation sites has not been identified to date. White and co-worker Ficarro, S. B., M. L. McCleland, et al. (2002). "Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae." Nat Biotechnol **20**(3): 301-5. identified 216 phospho-peptide sequence of *S.cerevisiae* proteins expressed in logalismic growth phase. However, no phospho-peptide of Sml1 was identified.

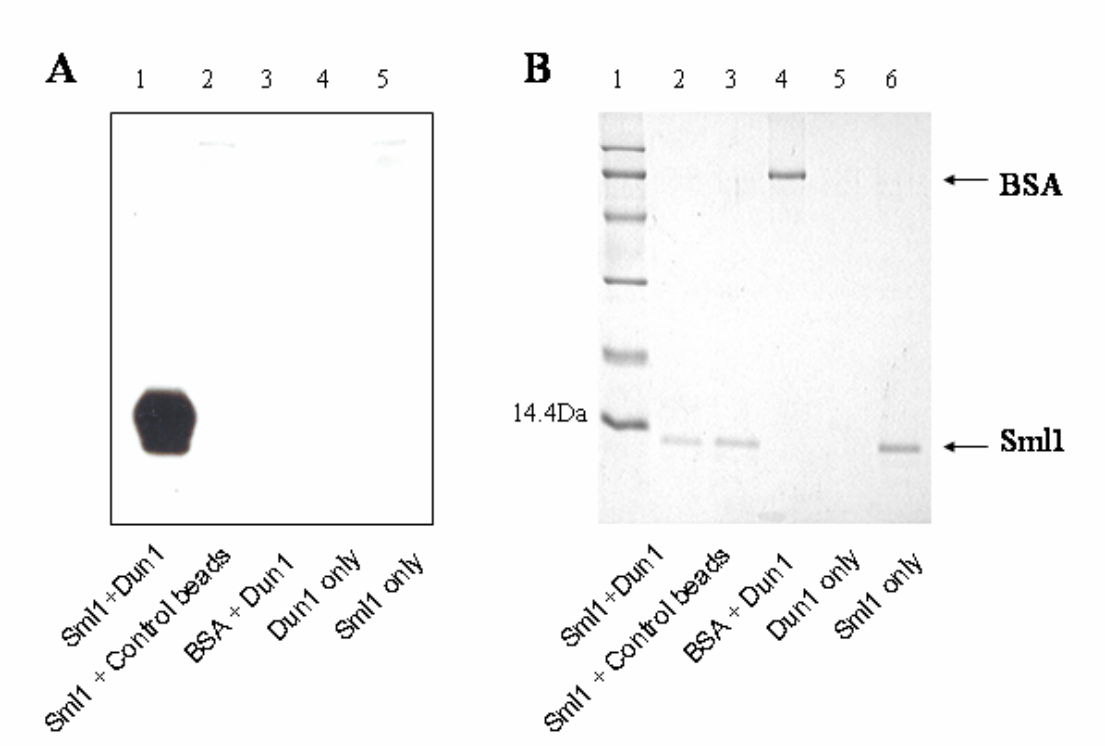

Figure 28. *In vitro* phosphorylation of Sml1 in the presence of γ[^{32}P]ATP.

(A) Autoradiography: *Lane 1: E. coli* expressed Sml1 in the reaction buffer containing 250μM ATP (cold) and 1μCi/μl γ[^{32}P]ATP was incubated for two hours with the glutathione beads pre-incubated with yeast lysate containing GST-Dun1 (see Experimental procedures). 10μl of supernatant was subjected to 15% (w/v) acrylamide SDS-PAGE. *Lane 2*: Sml1 was incubated with glutathione beads pre-incubated with the lysate of untransformed yeast cells (U952-B). *Lane 3*: BSA was incubated with the GST-Dun1 bound glutathione beads. *Lane 4*: The GST-Dun1 bound glutathione beads alone were incubated in the reaction buffer. *Lane 5*: Only Sml1 was incubated in the reaction buffer. **(B)** Coomassie Blue staining of the gel presented on panel A. *Lane 1*: Molecular weight standard of 97.4 Da, 66.2 Da, 45.0 Da, 31.0 Da, 21.5Da and 14.4Da (from top to bottom), Lanes 2-6: the same samples in lane 1 to 5 on panel A

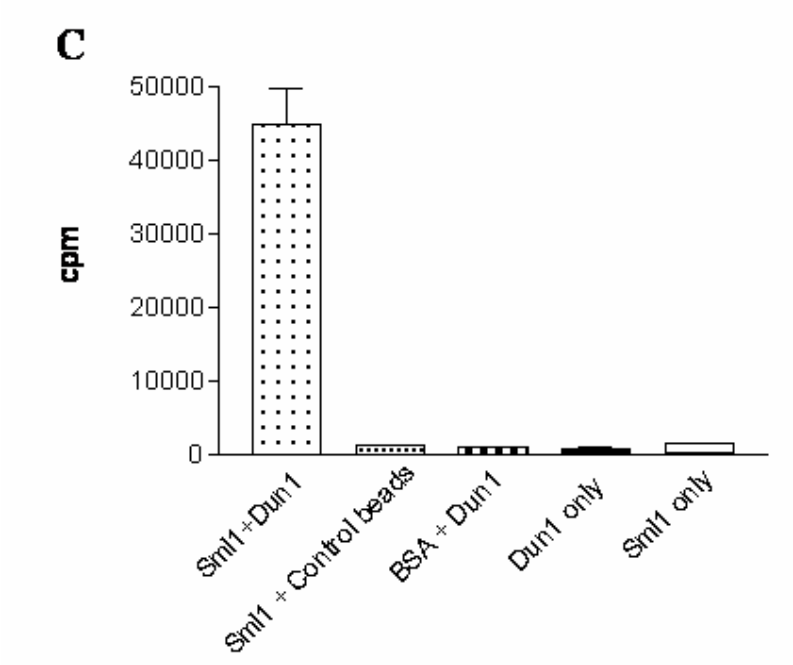

Figure 28 continued

(C) Radioactivity from 20µl of the same samples as in (A) measured by a liquid scintillation counter. Subtraction of average radioactivity of all the control samples from that of the experimental sample (Sml1 + Dun1 beads) was corresponds to phosphate incorporation (moles phosphate / moles protein) of 0.31. (Number of replicates measured was four. The error bars represent standard error of the mean.)

at 40 to 50 times greater than that of the negative control samples (Figure 28). These results show that *in vitro* phosphorylation of Sml1 observed in this experiment is solely due to GST-Dun1 and that the *in vitro* kinase activity of GST-Dun1 is specific to Sml1. The number of phosphoryl groups attached to Sml1 was determined by a combination of MS and site-directed mutagenesis. In order to estimate the minimum number of phosphorylation sites, intact molecules of Sml1 were first analyzed by ESI-FTICR. The attachment of a phosphate to a hydroxyl group of serine, threonine or tyrosine results in mass increases of the proteins or peptides by 79.966 Da in monoisotopic mass, and it was possible to estimate the number of phosphates on a protein by analyzing the mass shift. We observed monophosphorylated, di-phosphorylated, tri-phosphorylated, and unphosphorylated Sml1. These observations indicated that at least three sites of Sml1 could be phosphorylated (Figure 29).

In order to find the region of Sml1 containing the phosphorylation sites, the protein was digested by trypsin or CNBr, and then analyzed by ESI-FTICR. In the positive ion analysis of the tryptic digest, we detected peptides that covered 75% of the Sml1 sequence spanning residues 1 to 67 and 73 to 84. In the negative ion analysis, we detected peptides constituting 89% of the Sml1 sequence, which almost covered the entire Sml1 sequence with the exception of residues 16 to 26 (Figure 30A). In both positive and negative ion analysis, the spectrum for the samples taken from the *in vitro* Dun1 phosphorylation assays showed that peptides spanning residues 33-67 (M=3549.720 Da) and 34-67 (M=3705.857 Da) are associated with their singly phosphorylated forms of M=3629.695 Da and M=3785.829 Da, respectively (Figure 31).

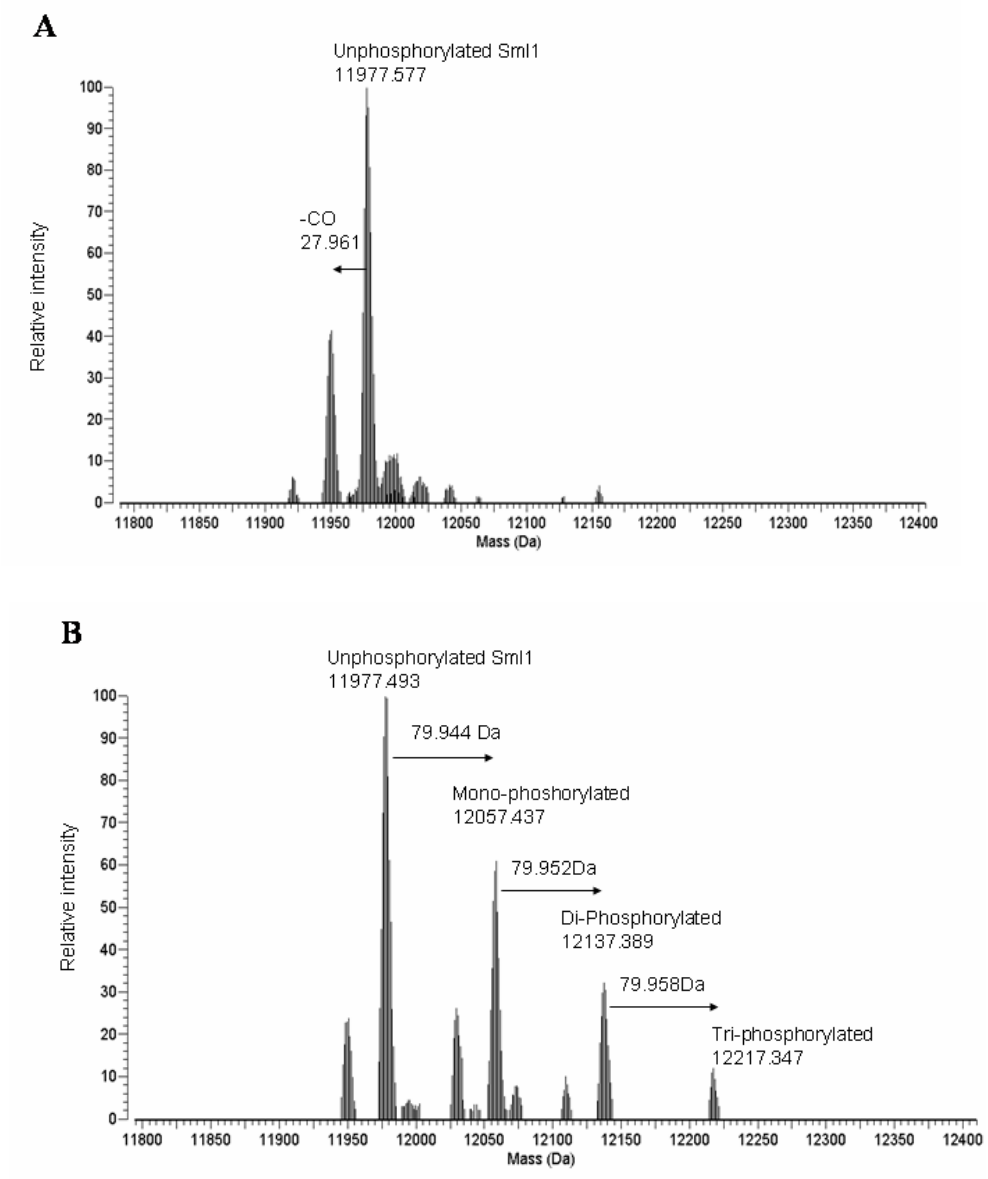

Figure 29. Positive ion analysis of intact phosphorylated Sml1.

(A) Deconvoluted mass spectrum of the negative control in which Sml1 was incubated in the presence of 250μM ATP without GST-Dun1 for two hours. The spectrum shows only unphosphorylated species. No peaks corresponding to the mass of phosphorylated Sml1 was observed. **(B)** Deconvoluted mass spectrum of Sml1 incubated with GST-Dun1 and 250μM ATP for two hours. In this particular spectrum, an internal calibration was accomplished with the known peaks of unphosphorylated Sml1. Values above each peak are the observed mass of the most abundant isotopic species. Note that the observed mass differences between different phosphorylation states closely correspond to the calculated monoisotopic mass of HPO_3 which is 79.966 Da.

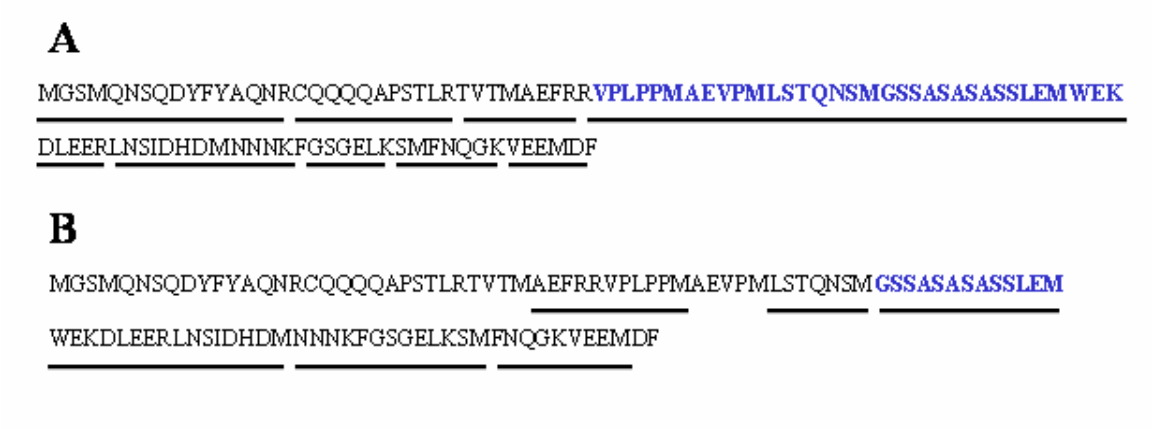

Figure 30. Sequence coverage of Sml1 obtained from (A) tryptic digest and (B) CNBr digest.

The peptides observed in ESI-FTICR-MS are underlined. Peptide sequences that contain phosphorylation sites are shown in blue.

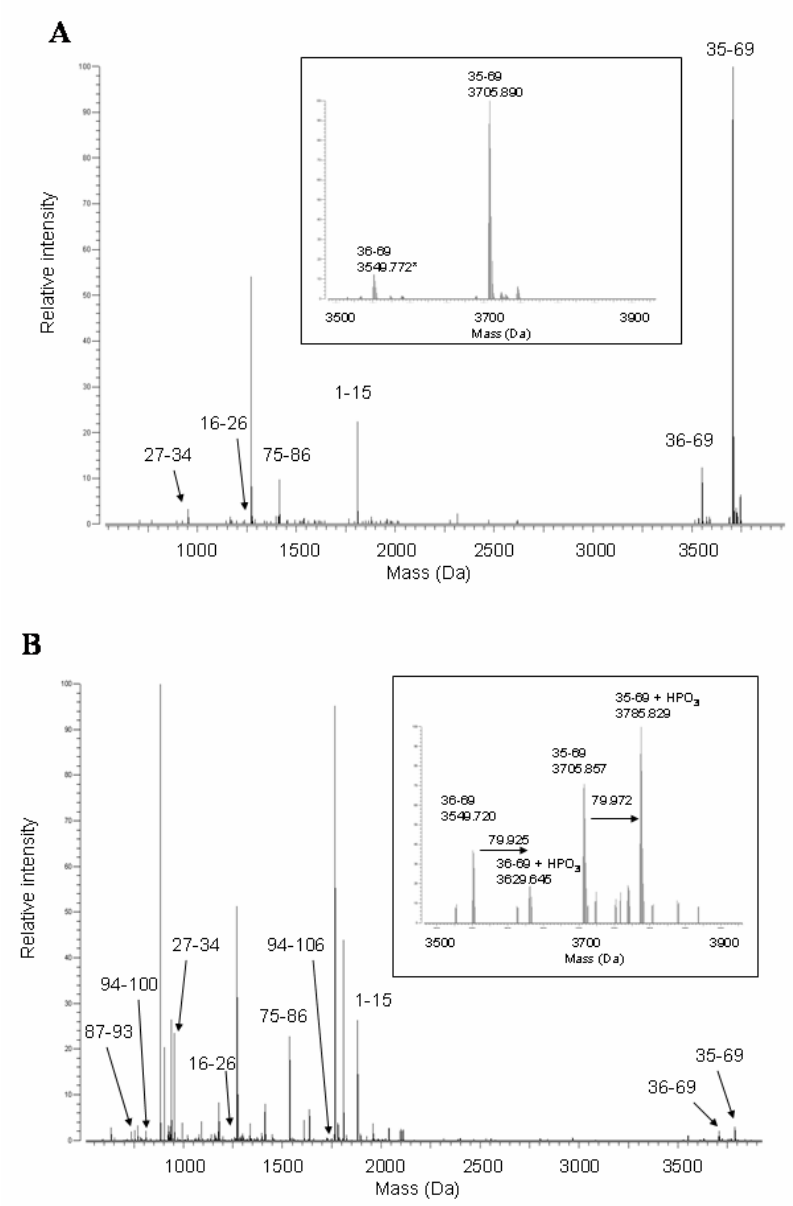

Figure 31. Positive ion analysis of Sml1 Tryptic digest.

(A) Tryptic digest of the negative control in which Sml1 was incubated in the presence of 250μM ATP without GST-Dun1. As expected, no phosphorylated species were identified (see inset). **(B)** Tryptic digest of the experimental sample in which Sml1 was incubated with GST-Dun1 in the presence of 250μM ATP. Monophosphorylated peptides spanning residues 35-69 and 36-69 were detected.

* The value above each peak is the observed monoisotopic mass based on the external calibration.

In the positive ion analysis of CNBr digests, we detected four peptides that cover 34% of the Sml1 sequence spanning residues 29 to 39 and 81 to 104. However, none of these peptides were phosphorylated. In the negative ion analysis of CNBr digests, 7 to 8 peptides that cover 71% of Sml1 sequence (residue 31-106) were observed (Figure 30B). These results showed that the peptide spanning residues 52-64 consisting of GSSASASASSLEM (M=1153.539 Da) was associated with its singly phosphorylated form (Figure 32A & B). The phosphopeptides of Sml1 identified by both CNBr and tryptic digest are entirely consistent. The combined data generated from the CNBr and tryptic digest gave us 100% sequence coverage of Sml1.

To enrich phosphopeptides, CNBr digest samples were subjected to Ga(III) IMAC chromatography and analyzed in the negative ion mode. Singly, doubly, and triply phosphorylated forms of residues 52-64 were observed (Figure 32C and Table 8), which were absent in the negative control which contained CNBr peptides generated from only unphosphorylated Sml1. In most of the samples, peak intensities of doubly or triply phosphorylated species were stronger than the singly phosphorylated species. Singly phosphorylated species could only be observed prior to IMAC chromatography. To confirm that the observed mass shift was due to phosphorylation, an ion of doubly phosphorylated species (M-H=1294.452 Da) was subjected to collisional activated dissociation (CAD). In this experiment, the species of interest was isolated in the analyzer region of the mass spectrometer, accelerated with a RF voltage and bombarded into nitrogen gas. Depending on RF voltage used to accelerate the ion, we observed fragment ions of 1196.489 and 1098.466 that corresponded to the loss of one or two

Table 8. Monoisotopic mass of peptides generated by CNBr digestion corresponding to residue 52-64

	Observed mass(Da)	Calculated mass(Da)	Mass difference(Da)
Unphosphorylated	1135.525	1135.400	0.125
Unphosphorylated C-hSer*	1153.557	1153.411	0.146
Singly phosphorylated	1215.444	1215.366	0.078
Singly phosphorylated C-hSer	1233.477	1233.377	0.100
Doubly phosphorylated	1295.428	1295.332	0.096
Doubly phosphorylated C-hSer	1313.454	1313.343	0.111
Triply phosphorylated	1375.361	1375.299	0.062
Triply phosphorylated C-hSer	1393.405	1393.309	0.096

*C-hSer: Species in which C terminus homolactone is converted to homoserine.

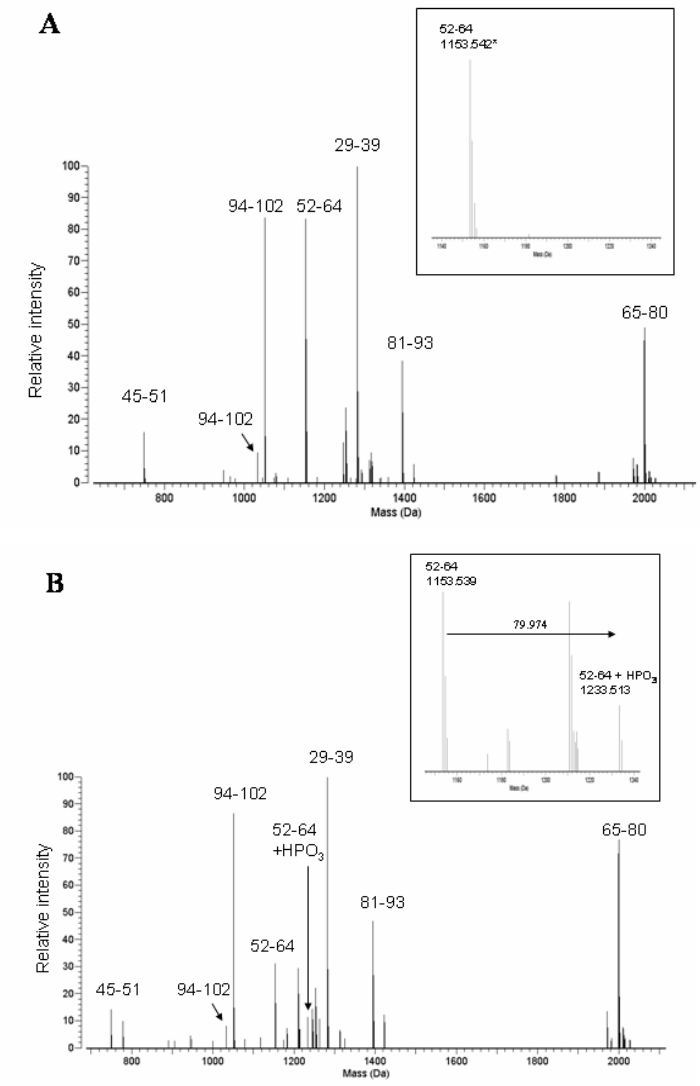

Figure 32. Negative ion analysis of Sml1 CNBr digest.

(A) The deconvoluted mass spectrum of the negative control of the kinase reaction subjected to CNBr digestion in which Sml1 was incubated in the presence of 250μM ATP without GST-Dun1. As expected, no phosphorylated species were identified (see inset). In general, residues at the C-terminus of peptides generated by CNBr digestion are homolactones, which can be further converted to homoserine by intra-molecular hydrolysis. In this spectrum, residues at the C-terminus of most of the peptides were homoserine. **(B)** The deconvoluted mass spectrum of phosphorylated Sml1 subjected to CNBr digestion in which Sml1 was incubated with GST-Dun1 in the presence of 250μM ATP. The peak (M=1233.535) corresponds to a singly phosphorylated peptide consisting of residues 52-64 (see inset). In this spectrum, residues at the C-terminus of the majority of the peptides were homoserine.

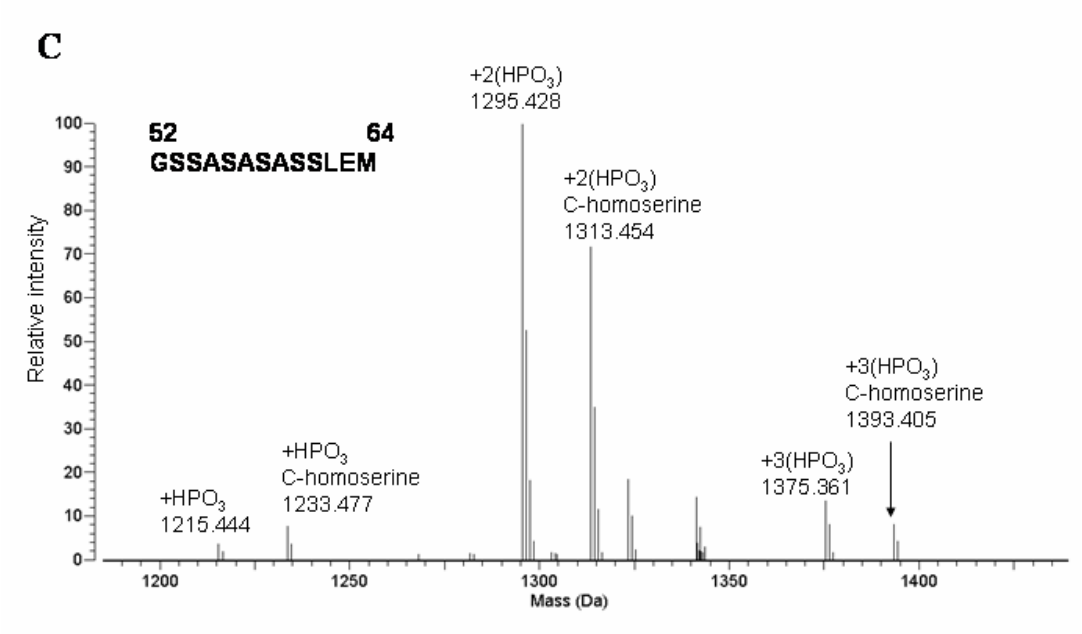

Figure 32 continued

(C) Phospho-peptides enriched by Ga(III) IMAC shows singly, doubly, and triply phosphorylated species. In this spectrum, residues at the C-terminus of these peptides were either homolactone or homoserine (denoted as C-term-homoserine).

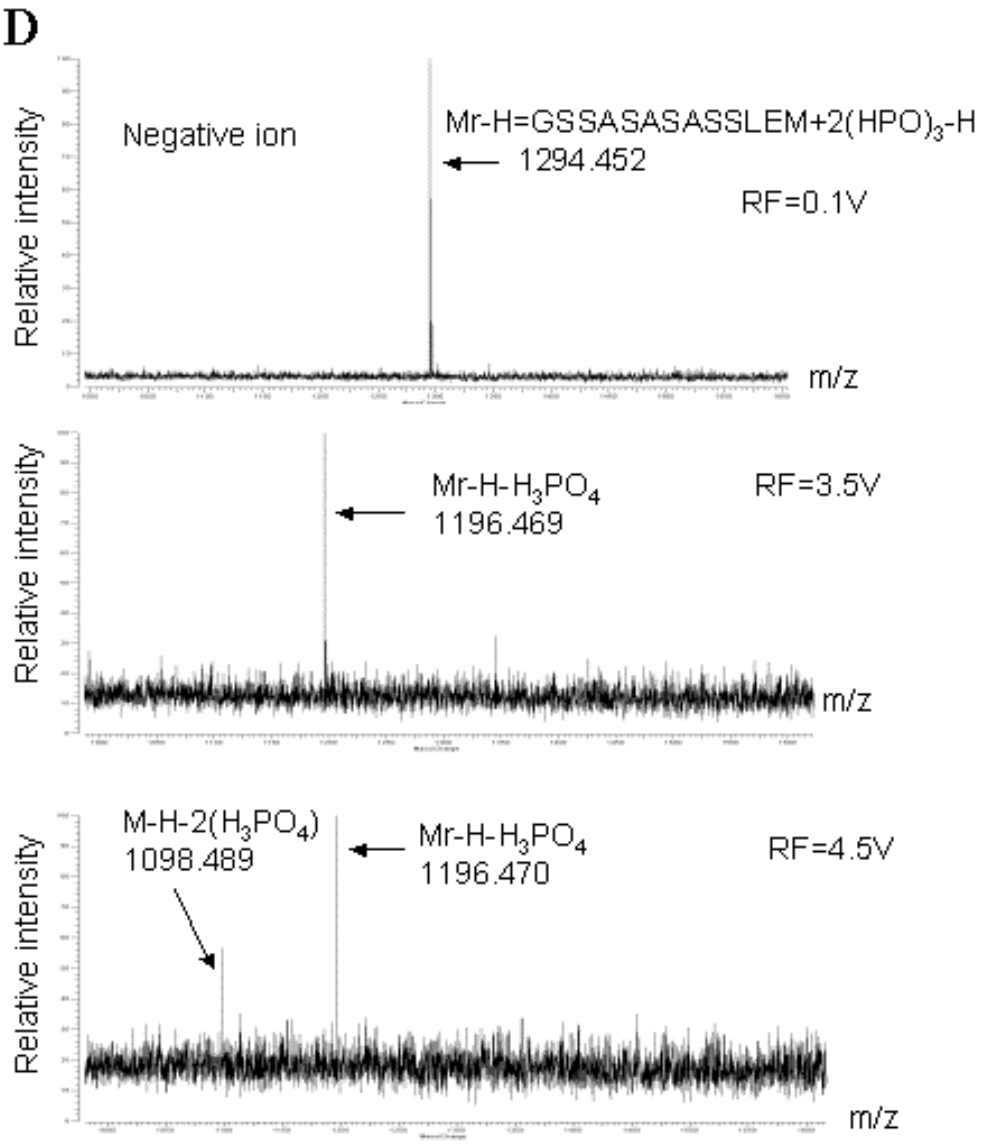

Figure 32 continued

(D) SORI-CAD fragmentation of a doubly phosphorylated peptide consisting of residues 52-56. Ions corresponding to the doubly phosphorylated peptide were isolated in the analyzer cell by an arbitrary waveform *(top)* and subjected to SORI-CAD fragmentation with MS burst voltages of 3.5V*(middle)* and 4.5V*(bottom).* The mass of two fragment ions corresponds to neutral loss of one phosphoric acid (-H3PO4= -97.977) and two phosphoric acids (-H_3PO_4x2=-195.954).

*The value above or beside each peak is the observed monoisotopic mass based on the external calibration.

phosphoric acids (Figure 32D). We could not obtain a reasonable signal in the CAD experiment for singly and triply phosphorylated species, which was possibly due to the low abundance of the parent ions. To further confirm our findings, we repeated the same experiment with the C14S mutant form of Sml1 and its proteolytic fragment. During purification of C14S Sml1 by gel filtration chromatography, the intact molecule was separated from a smaller degraded C14S Sml1 fragment corresponding to a mass of 8147.073Da. Mass spectrometric analysis of the latter revealed that the degraded peptide was an N-terminal fragment of C14S Sml1 consisting of residue 1 to 71 (Uchiki, Gupta et al. 2002), which we denoted as FRAG71. We conducted the kinase reaction using both the intact and FRAG71 of the C14S mutant. Both intact C14S Sml1 and FRAG 71 were singly, doubly, and triply phosphorylated (Figure 33A & B). Although this experiment was conducted with a mutant form of Sml1, it strongly suggests that three sites of Sml1 can be phosphorylated and that all three sites reside within the N-terminal 71 residues. In addition, triple phosphorylation of FRAG71 C14S strongly suggested that the C-terminal 33 residues, which is the minimal Rnr1 binding domain of Sml1(Zhao, Georgieva et al. 2000), were not required for its phosphorylation.

The data above indicated that a peptide of Sml1 consisting of residues 52-64 (GSSASASASSLEM) produced by CNBr digestion includes all the three phosphorylation sites. It should be noted that the above sequence does not contain any threonine residues, suggesting that only serine residues in Sml1 are phosphorylated by Dun1. However, as reported in other studies (Zhang, Herring et al. 1998; Stensballe, Jensen et al. 2000), we too were unable to obtain sequence specific fragmentation of the

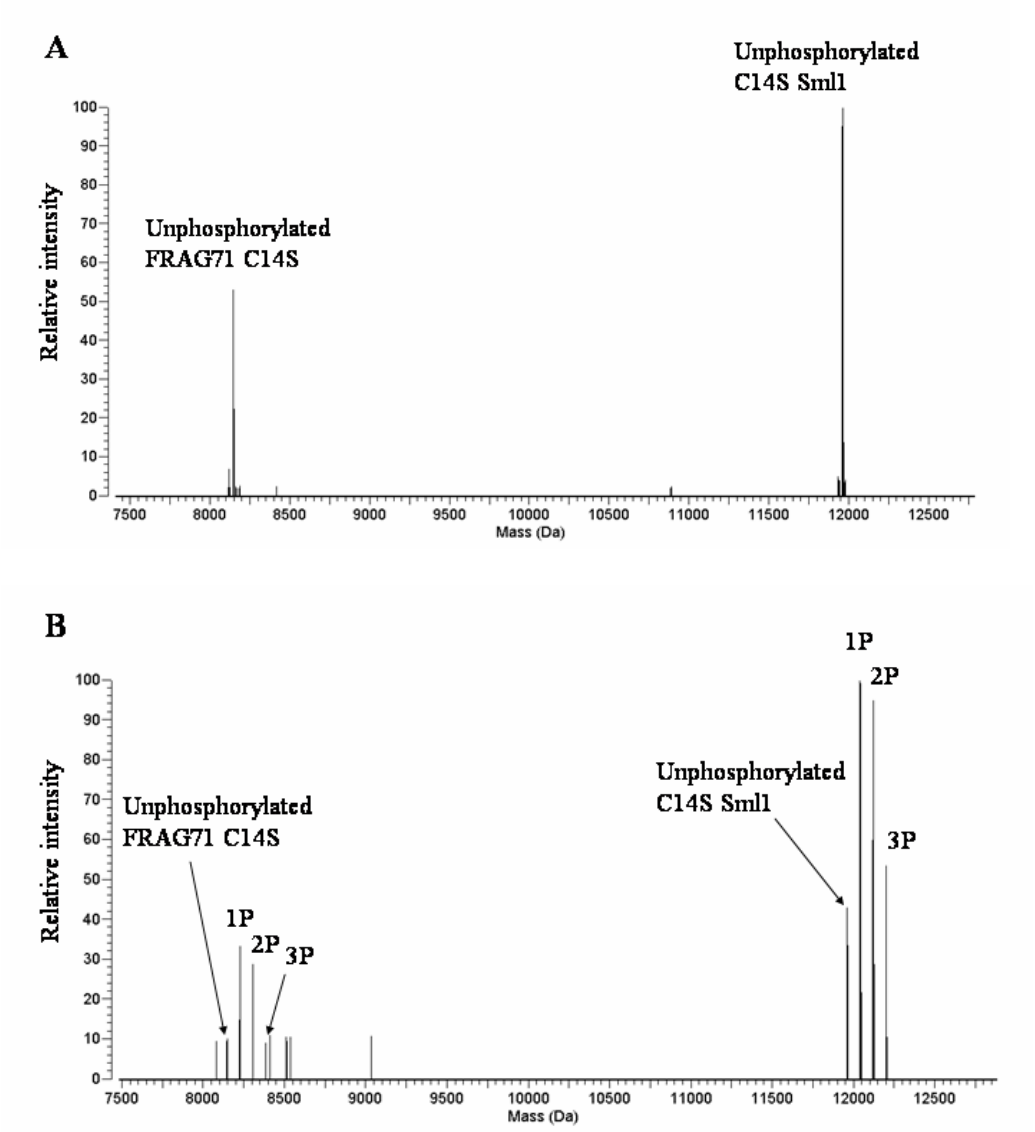

Figure 33. Positive ion analysis of intact phosphorylated C14S Sml1 and FRAG71C14S.

(A) The negative control in which Sml1 was incubated in the presence of 250μM ATP without GST-Dun1 for two hours showed only unphosphorylated species. **(B)** The deconvoluted mass spectrum of the phosphorylated C14S Sml1 and FRAG71 C14S shows singly, doubly, and triply phosphorylated species (denoted as 1P, 2P and 3P respectively).

doubly phosphorylated peptide by CAD as the predominant fragmentation was the neutral loss of phosphoric acid. In addition to this, the identification of phosphorylated residues from a tandem cluster of serine residues (GSSASASASSLEM) by MS is challenging. In order to simplify the identification of phosphoserine residues by CAD, we converted the phosphoserine residues to S- ethylcysteine by β–elimination in the presence of ethanethiol. This reaction specifically replaces phosphoryl group on serine or threonine with ethanethionyl giving a 44.008 Da mass shift from normal serine or threonine (Meyer, Hoffmann-Posorske et al. 1991; Lapko, Jiang et al. 1997; Adamczyk, Gebler et al. 2001; Oda, Nagasu et al. 2001). Unlike phosphoryl groups, ethanethionyl groups were more resistant to fragmentation by CAD so that fragmentation of the peptide mainly takes place at the peptide backbone (Meyer, Hoffmann-Posorske et al. 1991; Lapko, Jiang et al. 1997). In positive ion mode, we have observed peaks which correspond to modification of doubly and triply phosphorylated species, respectively (Figure 34A). First we conducted CAD with doubly modified species that was fragmented into a series of y and b type ions, from which two sites of phosphorylation were unambiguously determined (Figure 34B and Table 9). b*8, b*9, and b*10 ions can be produced only when phosphoryl groups are attached to Ser58 and Ser60. Similarly, only two possible combinations, Ser58/Ser60 or Ser58/Ser61, can produce observed values for y6, y7, and y8. Overall, more than 74% (20/27) of the peaks in the MS/MS spectrum could be assigned as fragment ions of a peptide in which ethanethionyl groups (that replaces phosphoryl groups) are attached to Ser58 and Ser60. The unassigned peaks do not match the possible fragments of peptides with other combinations of ethanethionyl

Table 9. Ions observed in CAD fragmentation of species corresponding to the doubly phosphorylated peptide of Sml1

Ion type	*Observed m/z*	*Calculated m/z*	*m/z difference*
Parent	1242.551	1242.428	0.123
b9	794.327	794.320	0.007
b11	994.454	994.436	0.018
b12	1123.506	1123.479	0.027
b13	1224.560	1224.480	0.080
b*8	645.270	645.272	0.002
b*9	776.316	776.313	0.003
b*10	863.354	863.345	0.009
b*11	976.442	976.429	0.013
b*12	1105.492	1105.472	0.020
b*13	1206.550	1206.420	0.130
b**8	627.260	627.261	0.001
b**9	758.304	758.302	0.002
b**11	958.430	958.418	0.012
b**12	1087.481	1087.161	0.320
y10	1011.467	1011.346	0.121
y9	940.431	940.305	0.126
y8	853.391	853.273	0.118
y7	782.352	782.236	0.116
y6	651.307	651.196	0.111
y*7	764.342	764.229	0.113
y*6	633.296	633.185	0.111

This peptide consists of residues 52-64 in which phosphoserine residues are derivatized by ethanethiol.
C terminus of the parent ion is homoserine.
b*: ions produced by neutral loss of one H_2O from b ions.
b**: ions produced by neutral loss of two H_2O from b ions
y*: ions produced by neutral loss of H_2O from y ions

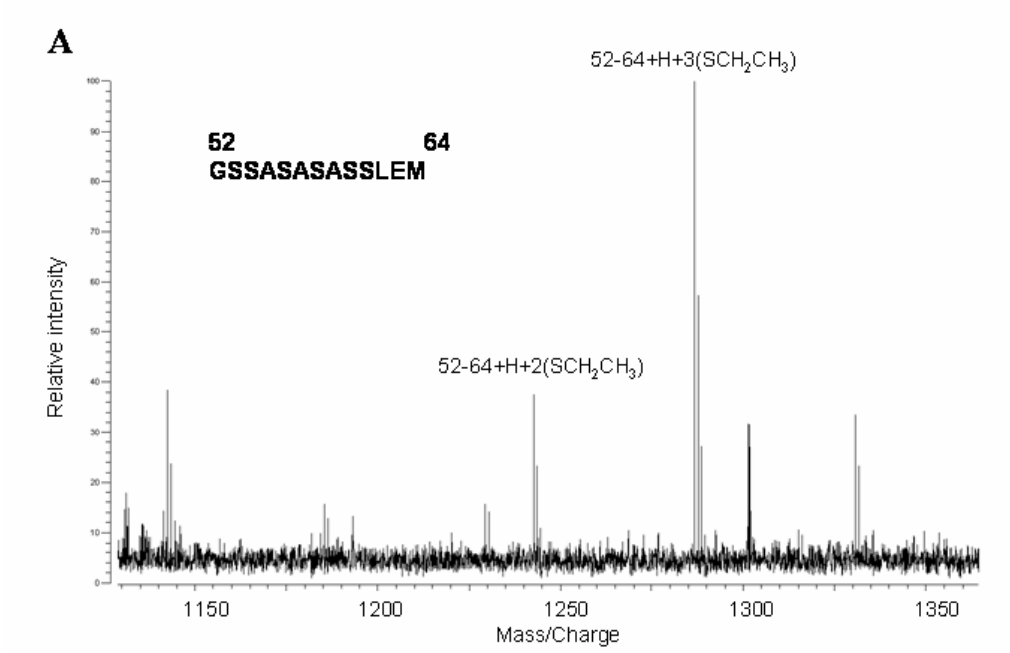

Figure 34. Positive ion SORI-CAD fragmentation of species corresponding to the peptide of Sml1 consisting of residues 52-64 in which phosphoserine residues are derivatized by ethanethiol. Ser56, Ser58, and Ser60 are identified as residues with phosphoryl attachment.

(A) The ions of derivatized peptides corresponding the doubly and triply phosphorylated. These ions are isolated as parent ions and fragmented as in (B) and (C).

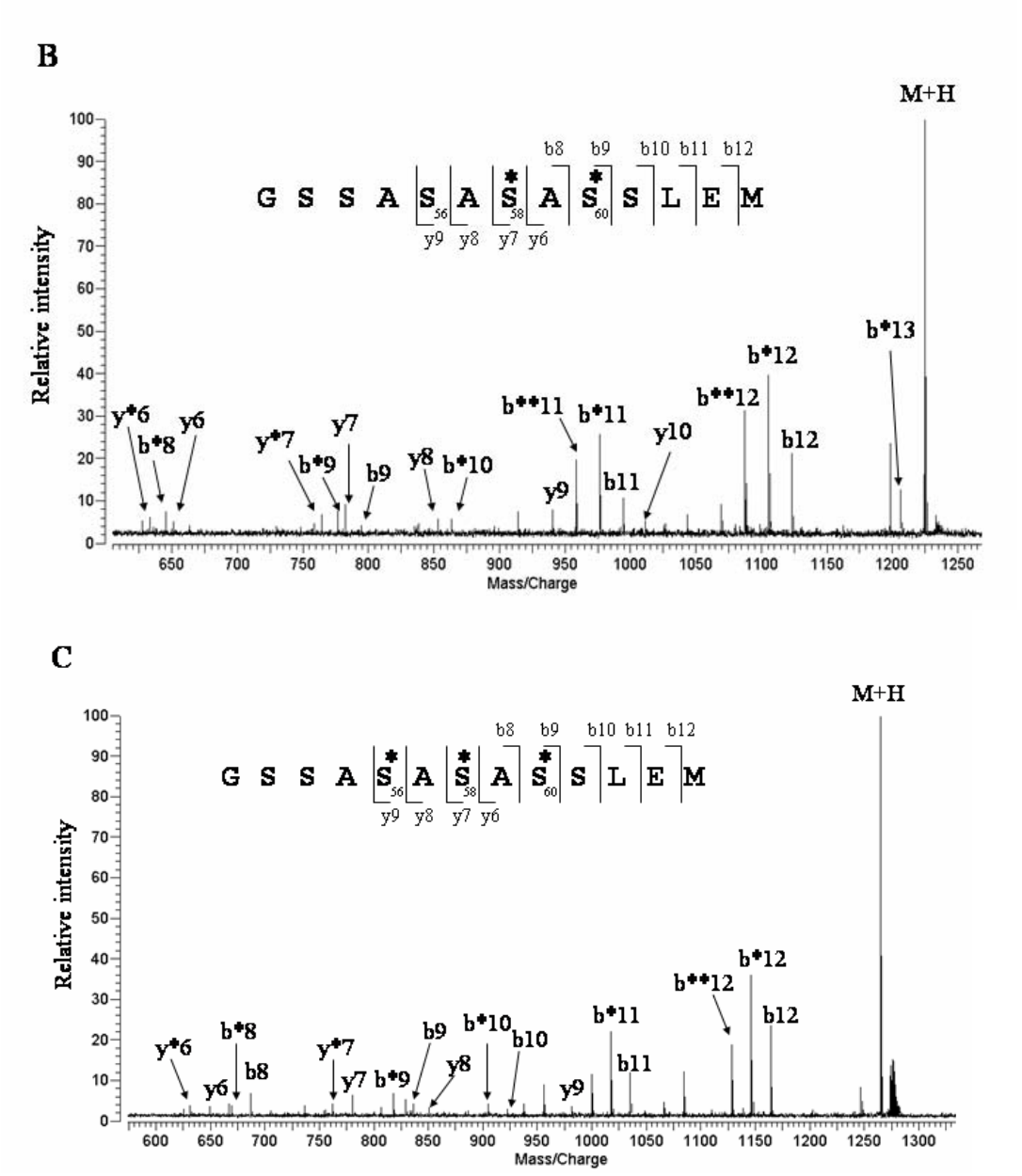

Figure 34 continued

(B) CAD fragmentation of ions corresponding to the doubly phosphorylated peptide. Over 70% of the observed peaks could be assigned when Ser58 and Ser60 were considered to be the phosphorylation sites. b* and b** ions are produced by the neutral loss of one or two H_2O from b ions. Likewise, y* ions are produced by the neutral loss of one H_2O from y ions. **(C)** CAD fragmentation of ion corresponding to the triply phosphorylated peptide. Over 70% of observed peaks could be assigned when Ser56, Ser58, and Ser60 were considered to be the phosphorylation sites.

attachment. This indicated that residues other than Ser58 and Ser60 are less likely to be the phospho-acceptors on this peptide. CAD experiment with the triply modified species also provided a series of y and b type ions corresponding to ethanethionyl attachment at specific serine residues (Figure 34C and Table 10). Y9, y8, y7, and y8 can be produced only when phosphate groups are attached to Ser56, Ser58, and either Ser60 or Ser61. Likewise, b*10, b*9, and b*8 can be generated only when Ser60, Ser58 and either Ser56, Ser54, or Ser53 are phosphorylated. Combining these data together, it is most likely that Ser56, Ser58, and Ser60 are phosphorylated. Over 70% of observed peaks (19/26) could be assigned as fragment ions of a peptide in which ethanethionyl groups (that replaced phosphate groups) are attached to Ser56, Ser58, and Ser60.

In order to determine primary phosphorylation sites of Sml1, we individually replaced Ser56, Ser58, and Ser60 with Ala and compared the degree of phosphorylation by a kinase assay based on incorporation of ^{32}P (Figure 35A & B). As compared to wild type Sml1, S60A Sml1 showed more than a 90% decrease in the amount of phosphate incorporation, while S58A Sml1 showed approximately a 25% decrease. No significant difference was observed between S56A and wild type Sml1. From autoradiography studies we observed that the phosphorylated form of these mutants appeared to migrate slightly faster than the wild type (in an order of S60A>S58A>S56A>Wild type Sml1). These results are in agreement with the results from mass spectrometry analysis. Although according to the MS results Ser56, Ser58 and Ser60 can be all phosphorylated, only the mutation at Ser60 showed a dramatic decrease in phosphate incorporation.

Table 10. Ions observed in CAD fragmentation of species corresponding to the triply phosphorylated peptide of Sml1

Ion type	*Observed m/z*	*Calculated m/z*	*m/z difference*
Parent	1286.536	1286.436	0.100
b9	838.332	838.329	0.003
b10	925.364	925.361	0.003
b11	1038.451	1038.445	0.006
b12	1167.497	1167.487	0.010
b13	1268.547	1268.44	0.107
b*8	689.28	689.281	0.001
b*9	820.32	820.321	0.001
b*10	907.355	907.353	0.002
b*11	1020.44	1020.437	0.003
b*12	1149.486	1149.48	0.006
b**8	671.268	671.27	0.002
b**12	1131.474	1131.469	0.005
y9	984.429	984.314	0.115
y8	853.386	853.273	0.887
y7	782.348	782.236	0.112
y6	651.306	651.196	0.110
y*7	764.339	764.226	0.113
y*6	633.296	633.185	0.111

This peptide consists of residues 52-64 in which phosphoserine residues are derivatized by ethanethiol

C terminus of the parent ion is homoserine

b*: ions produced by neutral loss of one H_2O from b ions

b**: ions produced by neutral loss of two H_2O from b ions

y*: ions produced by neutral loss of one H_2O from y ions

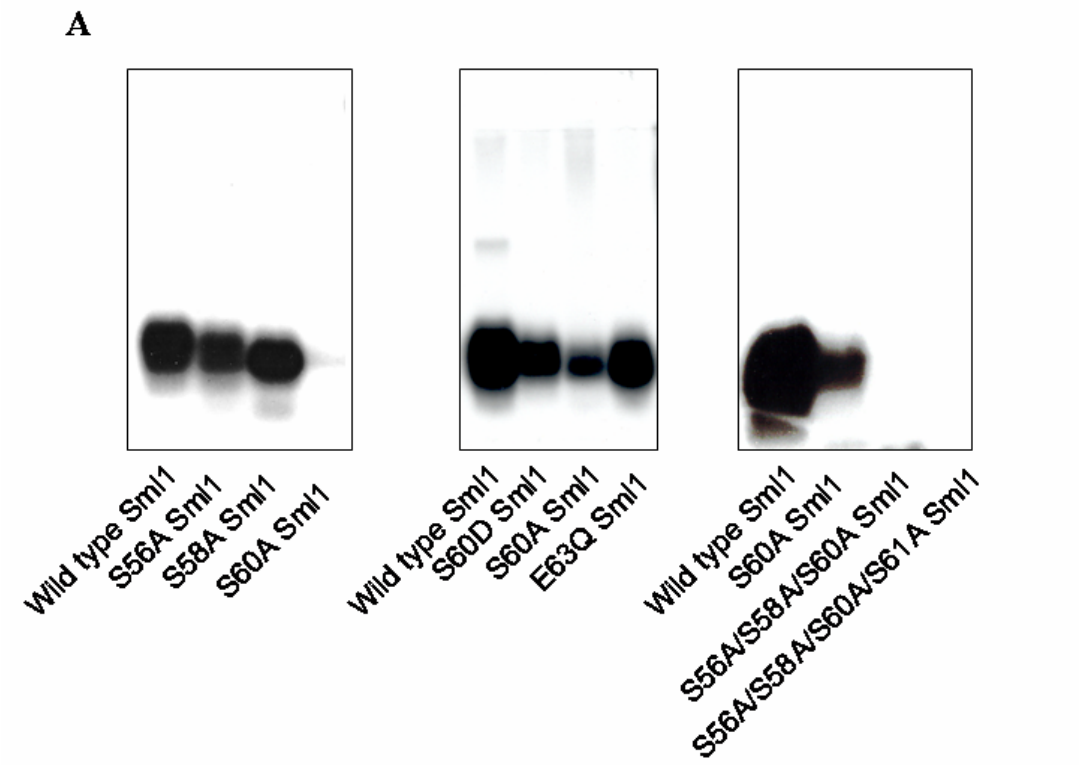

Figure 35. Quantitative analysis of phosphate incorporation in wild type, S56A, S58A, and S60A Sml1.

(A) Autoradiography. (*Left)* 30µl of 0.1µg/µl of *E. coli* expressed wild type (Lane 1), S56A (Lane 2), S58A (Lane 3) or S60A Sml1 (Lane 4) was incubated with GST-Dun1 bound glutathione beads in the presence of 0.06µCi/µl γ[^{32}P]ATP and 60µM ATP (cold) for two hours. After the reaction, 10µl of each sample was subjected to 15% (w/v) acrylamide SDS-PAGE, and the SDS-PAGE gel was exposed to X-ray film. (*Middle)* In the same manner the Dun1 kinase assay was conducted with the single mutants S60D (Lane2) and E63Q Sml1 (Lane 4) along with wild type (Lane 1) and S60A Sml1 (Lane 3). *(Right)* In the same manner the Dun1 kinase assay was conducted with the S56A/S58A/S60A triple mutant (Lane 3) and S56A/S58A/S60A/S61A quadruple mutant (Lane 4) along with wild type (Lane 1) and S60A Sml1 (Lane 2).

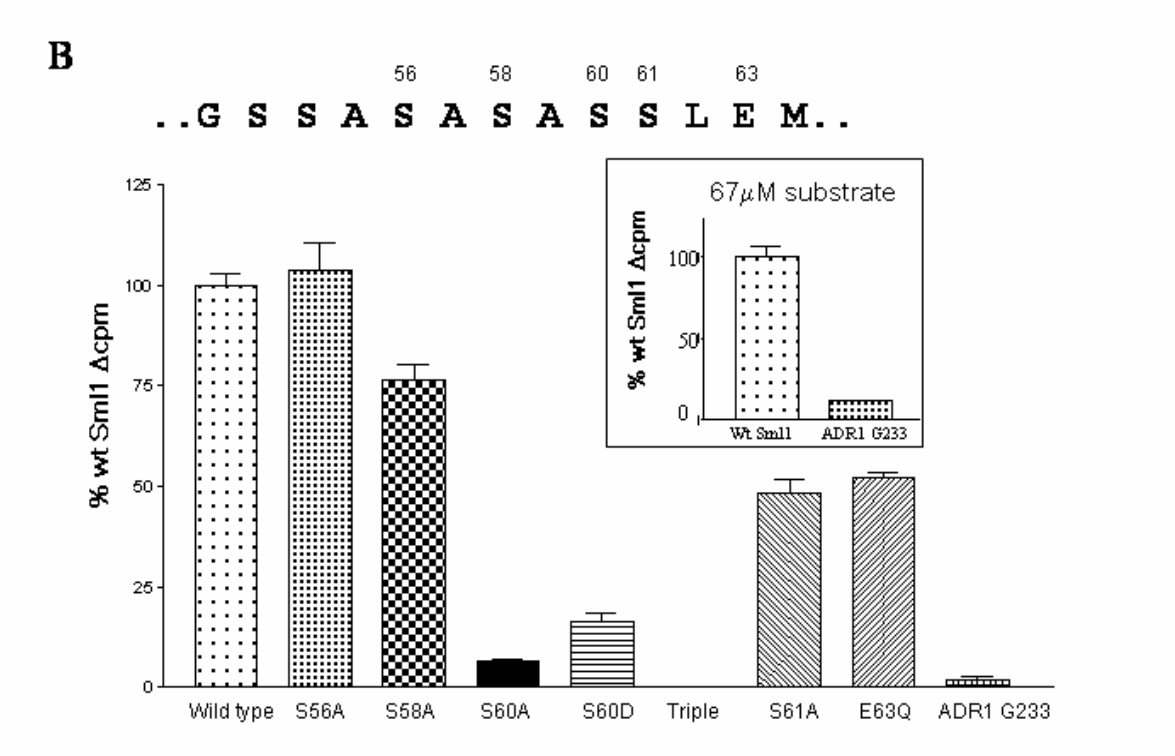

Figure 35 continued

(B) Phospho-cellulose filter assay. 20µl of the supernatant from the same samples as in (A) was taken, and their radioactivity (cpm) was measured. In a similar manner, the assay was performed with 8.3µM of ADR1 G233 peptide (LKKLTRRASFSGQ) for 30 minutes along with the same concentration of wild type Sml1 as a positive control. The inlet shows noticeable phosphorylation of the same peptide as compared to wild type Sml1 at a concentration of 67µM. Radioactivity given by incorporation of phosphate was determined by subtracting the values from samples in which only substrate proteins were incubated in the presence of 0.06µCi/µl $\gamma[^{32}P]$ATP and 60µM ATP (cold). The data was represented as a fraction of the radioactivity obtained for wild type Sml1. (Number of replicates was three. The error bar represents standard error of the mean.)

Moreover, due to its close proximity to Ser56, Ser58 and Ser60, we constructed the S61A mutant and tested it for phosphate incorporation. The S61A mutant had 48% phosphate incorporation (p=0.0001) as compared to wild type Sml1 (see Figure 35B). However, Ser61 was not identified as a potential phosphorylation site by MS analysis. To further test if Ser61 was a potential phosphorylation site we tested the triple mutant S56A/S58A/S60A for phosphate incorporation. Our results show that the triple mutant did not show any phosphate incorporation (see Figure 35A & B) strongly suggesting that it is Ser56, Ser58 and Ser60 that are the sole phosphorylation sites of Sml1 *in vitro*. Moreover, to test for cooperativity between the phosphorylation sites, we replaced Ser60 with an aspartic acid (S60D Sml1) which mimics the negative charge resulting from phosphorylation. On average the percentage incorporation of phosphate in S60D Sml1 was 17% of wild type Sml1 (Figure 35B). Nonetheless, this is a 2.5-fold increase in phosphate incorporation as compared to S60A Sml1 (p=0.01) (Figure 35B). To test if an acidic group (Glu63) located at the +3 position which is C-terminal of the three phosphorylation site acts as a recognition site for the Dun1 kinase we tested the E63Q Sml1 mutant for phosphate incorporation. Phosphate incorporation in E63Q Sml1 was 52% that of wild type Sml1 showing a significant decrease (p=0.00006). We also compared the Dun1 kinase activity for wild type Sml1 with respect to a synthetic peptide (named ADR1 G233) previously reported to be phosphorylated by Dun1 (Sanchez, Zhou et al. 1997). Using the substrate concentration of 8.3μM. no significant kinase activity above background was detected for ADR1 G233 peptide. In contrast, a significant amount phosphorylation was detected with Sml1 as the substrate. Next, to be consistent

with the experiment previously reported (Sanchez, Zhou et al. 1997), the kinase assay was conducted at a substrate concentration of 67μM. In this case, the average phosphate incorporation of ADR1 G233 was 11% of that in wild type Sml1 (Figure 35B, inset), suggesting that Sml1 is a far superior substrate for Dun1 as compared to the peptide.

We also exposed the yeast cells to the DNA alkylating reagent methyl methane sulfonate (MMS) during expression of GST-Dun1 for 2 hours. After addition of MMS, we noticed that cells stopped growing, while cells in media without MMS grew two fold during a 2 hours time period. The apparent kinase activity of GST-Dun1 purified from the cells exposed to MMS was lower than that of GST-Dun1 from the regularly grown cells (Data not shown). However, we could not conclude whether or not specific activity of these two sets of the Dun1 kinase are significantly different since the amount of Dun1 bound to the glutathione resin was not quantified. On the other hand, when the phosphorylated Sml1 was resolved by SDS-PAGE, we observed a slight difference in migration of Sml1 phosphorylated by GST-Dun1 from the cells exposed to MMS and the regularly grown cells; the former migrated lower than the latter (Figure 36A). A repeated experiment also showed similar migration patterns. The same experiment was conducted by using the S56A/S58A/S60A triple mutant Sml1 as substrates. Like regularly expressed GST-Dun1 kinase, we did not observe any level of phosphorylation on S56A/S58A/S60A Sml1 by GST-Dun1 from the cells exposed to MMS (Figure 36B).

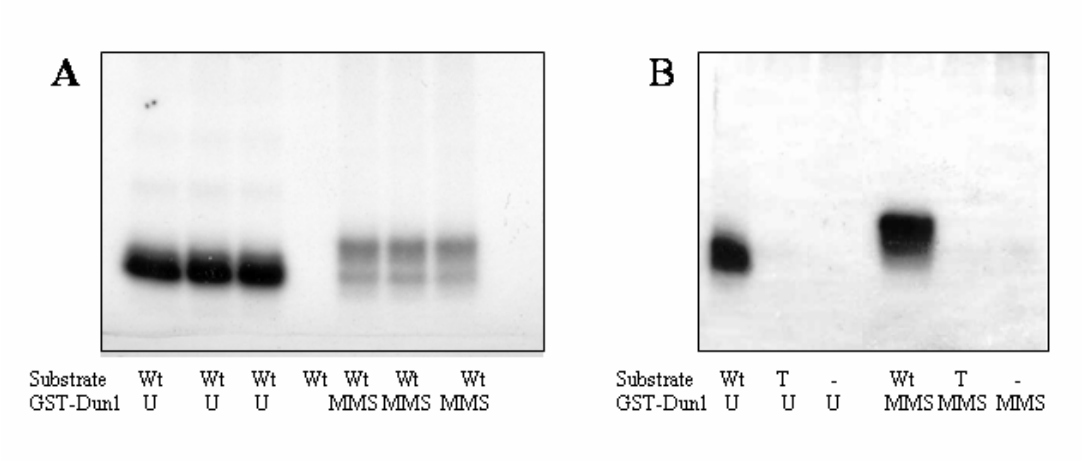

Figure 36. Phosphorylation of Sml1 by GST-Dun1 purified from *S.cerevisiae* cells grown in the presence of in the absence of methyl methane sulfonate (MMS).

(A). Phosphorylation of wild type Sml1. *Lanes 1-3.* 8.3μM of wild type Sml1 was incubated with 60μM ATP and 0.06μCi/μl γ[^{32}P]ATP and glutathione resin bound to GST-Dun1 purified from cells grown *in the absence of MMS*. Lane 4. Negative control in which 60μM wild type Sml1, 60μM ATP, and 0.06μCi/μl γ[^{32}P]ATP were incubated. *Lanes 5-7.* 60μM wild type Sml1, 60μM ATP, and 0.06μCi/μl γ[^{32}P]ATP and glutathione resin bound to GST-Dun1 purified from cells *grown in the presence of MMS*. **(B)** Phosphorylation of wild type, S56A / S58A / S60A triple mutant (denoted as T). *Lanes 1-4.* 16.6μM of wild type (lane 1), S56A / S58A / S60A triple mutant (lane 2), Sml1 and buffer without Sml1 (lane 3) was incubated with 60μM ATP and 0.1μCi/μl γ[^{32}P]ATP and glutathione resin bound to GST-Dun1 purified from cells grown *in the absence of MMS*. *Lanes 5-8.* 16.6μM of wild type (lane 5), S56A / S58A / S60A triple mutant (lane 6), and buffer without Sml1 (lane 7) was incubated with 60μM ATP and 0.1μCi/μl γ[^{32}P]ATP and glutathione resin bound to GST-Dun1 purified from cells grown in the presence of MMS.

Discussion

In this study, we have identified the phosphorylation sites of Sml1, the only biological substrate of Dun1 identified to date. First, we analyzed the intact phosphorylated Sml1 by ESI-FTMS to estimate the number of phosphate groups attached to Sml1. We observed a mixture of singly, doubly, and triply phosphorylated Sml1 as well as unphosphorylated species. This indicated that at least three sites of Sml1 could be phosphorylated. Next, in order to determine regions of phosphorylation, we digested phosphorylated Sml1 with trypsin or CNBr and analyzed the proteolytic fragments. In both the trypsin and CNBr digests, we observed singly phosphorylated peptides spanning residues 34 to 67 or 52 to 64. Repeated analysis of CNBr fragments enriched by Ga(III) IMAC column showed two to three phosphate attachment within residues 52-64. These data are consistent with the molecular weight of the intact phosphoSml1 with three phosphate attachments. Moreover, in order to determine the specific residues of phosphoryl attachment, we conducted further experiments on the fragment consisting of residues 52-64. Although initial CAD experiments showed fragmentation of phosphate groups, we could not observe fragmentation at the peptide backbone. Furthermore, residues 52-64 (GSSASASASSLEM) contain a cluster of six serine residues making it difficult to identify the specific phosphoserines. Therefore, the phosphoserine residues were converted into S-ethylcysteine by a β-elimination reaction in the presence of ethanethiol, which made the peptide amenable to fragment at its backbone. CAD fragmentation of the ions corresponding to doubly and triply phosphorylated peptides showed that Ser56, Ser58, and Ser60 were phosphorylated.

Under our experimental conditions, we did not observed a significant increase of Dun1 activity when the yeast cells were exposed to the DNA alkylating agent MMS. This was different from a previous report by Elledge and co-workers (Sanchez, Zhou et al. 1997) showing that the Dun1 kinase activity toward a synthetic peptide, ADR G233, increased 20 fold when the yeast cells were exposed to MMS. Although we did not quantify the specific activity of the Dun1 kinase, it is unlikely that the activity of our GST-Dun1 toward Sml1 increases 20 fold because the apparent band intensity (although it was very weak) on an SDS-PAGE gel stained with Coomassie Blue was not significantly different. This is similar to another report by Heyer and coworkers (Bashkirov, Bashkirova et al. 2003) that there is no significant change in autophosphorylation activity of GST-Dun1 when cells were exposed to MMS. However, we cannot exclude the possibility that specific activity of the Dun1 kinase toward Sml1 increases to a certain extent during DNA damage response.

On the other hand, we have observed a mobility shift on SDS-PAGE of Sml1 phosphorylated by GST-Dun1 prepared from the cells exposed to MMS. Although the difference in the migration is relatively small, similar mobility shift was repeatedly observed. A possible explanation for the different migration rate of Sml1 is that the number of phosphorylation sites on these two sets of Sml1 is different. For example, hyper-phosphorylated form of the Dun1 kinase migrate slower than its hypophosphorylated form (Zhou and Elledge 1993). However, like regularly expressed GST-Dun1, GST-Dun1 expressed in the presence of MMS did not phosphorylate S56A/S58A/S60A Sml1. This indicates that residues other than Ser56, Ser58 and Ser60

cannot be phosphorylated by GST-Dun1 regardless of the presence or the absence of DNA damage induced by MMS. Given the fact that the majority of Sml1 phosphorylated by regularly expressed GST-Dun1 appears to be a singly phosphorylated form on mass spectrometry, it is possible that GST-Dun1 expressed in the presence of MMS produce more doubly and triply phosphorylated Sml1. This hypothesis should be further tested in the future, possibly, by a systematic mass spectrometric analysis.

Our studies indicate that Ser56 might be phosphorylated after Ser58 and Ser60 due to the following: we observed phosphorylation of all three residues in the triply phosphorylated peptide, while only the phosphorylation of Ser58 and Ser60 were observed in the doubly phosphorylated peptides. Some serine/threonine kinases such as casein kinase II (Flotow and Roach 1989) and glycogen synthase kinase 3 (Fiol, Wang et al. 1990) phosphorylate multiple sites of their substrates by a "hierarchal phosphorylation mechanism" (Flotow and Roach 1989) in which prior phosphorylation of one site is a prerequisite for phosphorylation of other sites. We constructed several site-directed mutants to investigate if Sml1 phosphorylation by Dun1 follows hierarchal phosphorylation and also to identify the primary phosphorylation sites. Mutation of Ser60 to Ala eliminates phosphorylation by greater than 90%. On the other hand, mutating Ser58 to Ala only had a small reduction in phosphorylation, and no significant difference between S56A and wild type Sml1 was observed. These results suggest that Ser60 is the first primary site of phosphorylation, while Ser56 and Ser58 are minor sites of phosphorylation. Moreover, while the mutation of Ser60 to Asp also showed a significantly lower level of phosphate incorporation than wild type Sml1, it was

significantly higher than S60A Sml1. This shows that a negative charge at residue 60 enhances phosphate incorporation at Ser 56 and Ser58 suggesting that there is positive cooperativity between the primary phosphorylation site Ser60 and the minor sites Ser56 and Ser58. Although Ser61 was not identified as a potential phosphorylation site by MS analysis, the S61A mutant showed 48% phosphate incorporation as compared to wild type Sml1, suggesting that Ser61 could possibly be phosphorylated or it is necessary for recognition by the Dun1 kinase. However, the S56A/S58A/S60A triple mutant showed no phosphate incorporation strongly corroborating the MS result and unambiguously identifying Ser56, Ser58 and Ser60 to be the sole Sml1 sites of phosphorylation by the Dun1 kinase. Based on these results, we interpret the role of Ser61 to be involved in substrate recognition by Dun1.

Although our data support a hierachical mechanism of Sml1 phosphorylation, intramolecular phosphoryl transfer from Ser60 to Ser58 and from Ser58 to Ser56 cannot be completely excluded. The side chains of these residues may be close enough to each other on the polypeptide chain that the phosphate groups can be transferred from one to the other without enzymatic catalysis. In addition, S56A and S58A mutations have a little effect on the degree of Sml1 phosophorylation indicating that Ser56 and Ser58 are not the major phosphorylation sites. However, Ser56 and Ser58 are likely to be phosphorylated by Dun1 kinase because S60A and S60D mutant Sml1 are still phosphorylated. Moreover, if Ser56 and Ser58 are phosphorylated solely by intramolecular phosphoryl transfer, rather than the enzymatic catalysis, there must be a doubly phosphorylated species in which phosphates are attached on Ser56 and Ser58. CAD analysis of the

peptide corresponding to doubly phosphorylated species showed that Ser58 and Ser60 are mainly phosphorylated on the doubly phosphorylated species. Nevertheless, we cannot exclude this possibility.

The amino acid sequence of the Sml1 phosphorylation sites (GSSASASASSLEM) identified in our study is completely different from that identified by the synthetic peptide substrates by Sanchez et al. (Sanchez, Zhou et al. 1997). This study demonstrated that synthetic peptides containing a consensus cAPK recognition sequence (RRXS/TY; X, small residues; Y, residues having large hydrophobic group) can be phosphorylated by Dun1, and replacement of the –3 Arg to Ala abolishes phosphorylation. Although Sml1 does not contain the cAPK consensus sequence, flanking regions of a few serine residues have some similarity to it. For example, Ser75 has an Arg at the –3 position and an Ile, which is a relatively large hydrophobic group, at +1 position (RLNSI). Similarly, Ser87 has a Lys (as a basic residue) at the –3 position and a Gly (as a small residue) at the –1 position (KFGS). However, analysis of the wild type Sml1 by CNBr digest as well as the tryptic digest and FRAG71 of the C14S mutant indicated that Ser75 and Ser87 are unlikely to be phosphorylated. On the contrary, the flanking regions of the phosphorylation sites identified by our study did not show any similarity to the cAPK consensus sequence. Furthermore, we demonstrated that a synthetic peptide used by Sanchez et al. (1997) can be phosphorylated by Dun1 only at a high concentration, and the relative level of its phosphorylation is only about 11% of Sml1. These findings show that the Dun1 recognition motif is different from the cAPK

recognition motif as Sml1 is a better substrate of Dun1 as compared to the cAPK-based peptide.

We investigated if the Sml1 phosphopeptide (GSSASASASSLEM) is similar to phosphorylation motifs found in other proteins. A Phi-BLAST search as well as simple BLAST searches in the NCBI protein sequence database showed that very few proteins have similar sequences to the Sml1 phosphorylation sites, and that none of these proteins seems to be involved in cell cycle checkpoint pathways. However, we cannot rule out the possibility of other unidentified substrates of Dun1 that may have different phosphorylation motifs as compared to that of Sml1. To further compare the specificity of phosphorylation by Dun1, we have conducted a database search for possible phosphorylation sites of Sml1 by several unknown serine/threonine kinase using the web-based server PhosBase (http://www.cbs.dtu.dk/databases/PhosphoBase/). In this search, each serine and threonine residue of Sml1 was scored based on sequence comparisons of Sml1 and substrates for several known serine/threonine kinases. Ser61, Ser58 and Ser56 were the first, second and third most probable phosphorylation sites with scores of 0.991, 0.939 and 0.845, respectively. The recognition motif of 10 known serine/threonine kinases were also searched, and Ser60 was predicted to be a potential phosphorylation site for only Casein kinase II. The consensus specificity motif for Casein kinase II is XS*/T*XXE/DX (Hanks and Quinn 1991). If Ser60 is considered to be the site of phosphorylation then the XS*/T*XXE/DX consensus is satisfied by the ASSLEM stretch of the Sml1 phosphopeptide sequence where Glu63 is the +3 site. Our results show that mutation of Glu63 to glutamine decreases phosphate incorporation down to 52% of wild

type Sml1, and therefore it is involved in recognition by Dun1. Nevertheless, the phosphorylation sites Ser56 and Ser58 of Sml1 cannot be accommodated within the Casein kinase II recognition motif XS*/T*XXE/DX, suggesting that the Dun1 kinase has a unique phospho-recognition motif.

A prerequisite for phosphorylation requires that the potential site to be freely accessible to the kinase of interest. Previously, in the absence of a three-dimensional structure, we constructed a molecular model of Sml1 based on *ab initio* Rosetta algorithms (Gupta, Peterson et al. 2004), which was in close agreement with CD and NMR data (Zhao, Georgieva et al. 2000). In this model, Ser56, Ser58, and Ser60 are located at a flexible region adjacent to a C-terminal alpha helix (Figure 37). Side chain solvent accessibilities of Ser56, Ser58, and Ser60 in our model were 30.6Å^2, 49.6 Å^2, and 61.6 Å^2, respectively. The accessibilities tabulated for Ser56, Ser58, and Ser60 reflect 40%, 64%, and 80% of side chain accessible area for serine in a random coil flanked by two Gly residues on either side (Fraczkiewicz and Braun 1998). These results suggest that Ser56, Ser58, and Ser60 are sites relatively accessible to the Dun1 kinase for phosphorylation.

The biological consequence of Sml1 phosphorylation is to dramatically reduce the cellular pools of Sml1 resulting in the elevation of RNR activity (Zhao and Rothstein 2002). At present the mechanism involving Sml1 down-regulation after its phosphorylation by Dun1 is not known. However, phosphorylation followed by

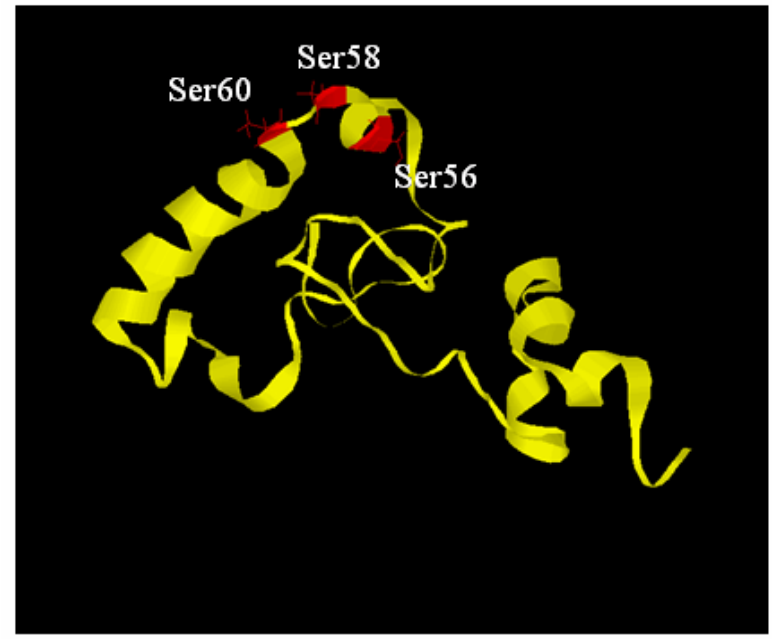

Figure 37. Phosphorylation sites mapped on a molecular model of Sml1.

A molecular model of Sml1 was constructed by a template-based program Rosetta/I-sites (Gupta, Peterson et al. 2004). The model shows that the phosphorylated residues Ser56, Ser58, and Ser60 are exposed to solvent indicating that these residues are easily accessible to Dun1.

degradation of proteinaceous inhibitors has been observed in the regulation of several biological processes. In *S. cerevisiae*, Cdc4 is a part of the SCF ubiquitin ligase complex that is responsible for recruiting several target proteins such as CDK inhibitor Sic1 (Verma, Annan et al. 1997) and $p25^{rum1}$(Benito, Martin-Castellanos et al. 1998) which are phosphorylated at the end of G_1 phase. Based on binding of Cdc4 to its natural targets and synthetic peptides, a consensus phosphopeptide motif for Cdc4 was identified, and it was designated as the Cdc4 phospho-degron (CPD) motif (L/I-L/I/P-pT-P<RK> (Nash, Tang et al. 2001). However, Sml1phosphorylation sites did not show any indication of a CPD motif. Possibly degradation of Sml1 is mediated by an unidentified ubiquitin ligase that recognizes multiple phosphorylated residues. Future work will involve the identification of the pathway responsible for Sml1 degradation.

Summary

In this study, we characterized the phosphorylation of Sml1 at the molecular level. Initial mass spectrometric analysis of intact phosphorylated Sml1 showed that Sml1 can be singly, doubly and triply phosphorylated. Further mass spectrometric analysis of Sml1 phospho-peptides generated by trypsin and CNBr digestion identified phospho-acceptors of Sml1 to be Ser56, Ser58 and Ser60. To further investigate Sml1 phosphorylation, we constructed the single mutants S56A, S58A, S60A and the triple mutant S56A/S58A/S60A and compared their degrees of phosphorylation to that of wild type Sml1 based on ^{32}P incorporation. There was no observed phosphate incorporation in the triple mutant, supporting that Ser56, Ser58 and Ser60 in Sml1 are the sole

phosphorylation sites. We observed a drastic decrease in the relative degree of phosphorylation of S60A as compared to that of wild type, while S56A and S58A mutants showed only a little or no decrease. This result indicates that Ser60 is the primary phosphorylation site on Sml1. We also compared the degree of phosphorylation in S60A Sml1 and a phospho-mimic mutant, S60D Sml1. S60D Sml1 showed a 2.5 fold higher level of phosphorylation than S60A Sml1. Combining results in the mass spectrometric mutagenesis studies, we hypothesize that phosphorylation of Sml1 follows a hierarchical phosphorylation mechanism in which Ser60 is the primary phosphorylation site and phosphorylation at Ser60 enhances phosphorylation of Ser56 and Ser58. Site directed mutagenesis to replace Ser61 and Glu63 with alanine also indicates that these residues are important for recognition of Sml1 by the Dun1 kinase. Taken together, this study revealed the *in vitro* phosphorylation sites of Sml1 by the Dun1 kinase, the mechanism of Sml1 phosphorylation and substrate specificity of the Dun1 kinase.

Chapter 6 Relation between the Sml1-Rnr1 interactions and phosphorylation of Sml1 by the Dun1 kinase

Introduction

Protein phosphorylation is the most ubiquitous mode of protein regulation in Eukaryotes. Nearly 30% of all proteins in Eukaryotes are thought to be phosphorylated at any given time (Zolnierowicz and Bollen 2000). The importance of protein phosphorylation is underscored by the fact that the protein kinases are the largest family of proteins in eukaryotes, comprising 1.5 to 2.5% of their genome (Hunter 1994; Manning, Plowman et al. 2002). Furthermore, all major kinase families are found among eukaryotes indicating, that similar modes of regulation are conserved through evolution. Protein phosphorylation is an effective way to switch biochemical properties of proteins. Negative charge introduced by phosphorylation cause significant effects on target proteins, either by inducing a conformational change or by changing their ability to interact with other macromolecules. Moreover, protein phophorylation can be quickly reversed by removal of phosphate (de-phosphorylation) by protein phosphatase. By a combination of phosphorylation and de-phosphorylation, cells can switch properties of proteins from one state to another rapidly. As a consequence of this, phosphorylation can be used to quickly change the biochemical environment in the cell without synthesizing new proteins. Indeed, protein phosphorylation plays a key role in several complex cellular functions (Manning, Plowman et al. 2002) such as neuronal development, morphogenesis, inflammatory response, stress responses, and cell cycle control. As

discussed in chapter 1, DNA damage response and replication blocks are also largely regulated through protein phosphorylation.

A majority of proteins are found in multi-protein complexes within the cell (Kumar and Snyder 2002; Janin and Seraphin 2003). For example, a large scale affinity purification of 589 *S.cerevisiae* proteins followed by mass spectrometric analysis[23] of proteins co-purified with the 589 yeast proteins revealed that at least 78% of these proteins form stable complexes (Gavin, Bosche et al. 2002). These proteins are involved in almost all types of fundamental biological processes, such as energy metabolism, synthesis and turnover of macromolecules (such as protein, nucleic acid and lipid), cell cycle, maintenance of cell polarity and structure, and gene expression. In the yeast proteome consisting of over 6000 proteins, approximately 30,000 protein-protein interactions are thought to occur (Kumar and Snyder 2002). These numbers reflect the fact that biochemical activities of proteins are largely regulated by protein-protein interactions. A subset of these interactions must occur during chemical modifications of molecules that require protein-protein interactions. Not surprisingly, protein-protein interaction plays essential roles in DNA damage response and replication blocks. For example, based on co-immunoprecipitation followed by mass spectrometric analysis of the immunoprecipitated proteins, Tyer and co-workers (Ho, Gruhler et al. 2002) identified proteins interacting with the 86 *S.cerevisiae* proteins implicated in DNA

[23] In this analysis, a cassette of a tandem affinity tag comprising of calmodulin binding protein and protein A was fused to 1739 ORFs. Of the 1739 proteins, 589 of them were successfully purified. The co-purified proteins associated with the 589 proteins were separated on SDS-PAGE gel, and they were identified by tryptic digestion followed by MALDI-TOF mass spectrometry. Of the 589 proteins 22% could not be identified due to technical problems such as detection limit or stability issues. However, this result does not exclude the possibility of complex formation by these 22% of proteins.

damage response. Their result revealed a complex network of protein-protein interactions involving majority of these proteins.

It is quite common for phosphorylated proteins to interact with proteins besides the kinases or phosphatases responsible for their phosphorylation or dephosphorylation. Cdk2 is an example that has been studied intensely. Cdk2 is inhibited by inactivating Tyr-15 and Thr-14 phosphorylation by Wee 1 (Gu, Rosenblatt et al. 1992; Chow, Siu et al. 2003; Coulonval, Bockstaele et al. 2003), and by the interaction with a Cdk inhibitor, $p21^{cip1}$ (Sherr and Roberts 1999). On the other hand, Cdk2 is activated by the phosphorylation of Thr=160 by the Cdk activating kinases (Kaldis 1999; Coulonval, Bockstaele et al. 2003), the dephosphorylation of Tyr-15 and Thr-14 by Cdc25 phosphatase (Gabrielli, Lee et al. 1992; Hoffmann, Draetta et al. 1994), and physical interactions with cyclins E and A (Murray 2004). Furthermore, p21 is phosphorylated by Akt kinase (Rossig, Jadidi et al. 2001; Zhou and Hung 2002), while cyclin E is phosphorylated by glycogen synthase kinase and Cdk2 (Welcker, Singer et al. 2003). This example illustrates a complexity of protein regulation by phosphorylation and protein-protein interactions.

Protein phosphorylation reinforces protein-protein interaction. For example, F box proteins of SCF E3 ubiquitin ligases specifically bind to phosphorylated forms of their target proteins, bringing the target to a proximity of the ubiquitin ligase catalytic components (Skowyra, Craig et al. 1997). *S. pombe* Rad24 specifically binds to a phosphorylated form of Cdc25, tethering Cdc25 to the cytoplasm and blocking its access to Cdc2/cyclin B (Zeng and Piwnica-Worms 1999). As discussed in chapter 1, in DNA

damage response and replication blocks, Rad53 and Dun1 are activated upon their interaction with phosphorylated proteins (Gilbert, Green et al. 2001; Bashkirov, Bashkirova et al. 2003). It should be noted that recognition of phosphoproteins is often mediated by distinct domains or proteins, such as Src-homology domain 2, WD40 repeats[24], Folk Head Associated domain (FHA domain), and 14-3-3 proteins that specifically recognize particular phosphoproteins[25] (Yaffe and Elia 2001; Pawson 2004). On the other hand, phosphorylation can also weaken protein-protein interaction in some cases. For example, during the G1-phase of the cell cycle, cell cycle inhibitory protein $p21^{cip1}$ forms a stable complex with Cdk2 and inhibits activity of Cdk2/E-cyclin. It also interacts with proliferating cell nuclear antigen (PCNA) and inhibits activity of PCNA to recruit DNA polymerase to the origin of replication. Before mitosis, $p21^{cip1}$ is specifically phosphorylated by Akt kinase, and the phosphorylation causes dissociation of $p21^{cip1}$ from Cdk2 and PCNA (Rossig, Jadidi et al. 2001). Consequently, Cdk2 and PCNA become active after the phosphorylation of p21, leading to initiation of the S phase of the cell cycle.

A noteworthy question in biology is how protein-protein interaction (other than interaction with kinases or phosphatases) affects protein phosphorylation. Sometimes such interactions may block phosphorylation by causing the phosphor-accepting residue to be inaccessible to the kinase. For example, eIF4E is a component of 5' mRNA cap binding complex, which is involved in translation initiation of Eukaryotic mRNA. In

---

[24] A protein-binding motif containing 7 repeated regions of about 40 amino acids containing conserved Trp and Asp.
[25] Some 14-3-3 and WD40 recognize proteins that are not phosphorylated.

response to certain chemical stresses, eIF4E is phosphorylated by Mnk kinase, and the phosphorylation of eIF4E stimulates the initiation of mRNA translation. On the other hand, 100-KDa adenovirus protein, L4 100K, competitively binds to eIF4E and blocks its phosphorylation by Mnk kinase (Cuesta, Xi et al. 2004). Interestingly, Mnk kinase and L4 100K share a basic amino acid rich sequence, which is responsible for their binding to eIF4E. On the other hand protein-protein interaction may do the opposite by exposing the phosphorylation site on the proteins enhancing phosphorylation. Such an enhancement of protein phosphorylation has been artificially demonstrated. The dominant negative mutant of phage Mu immunity repressor, Vir, interacts with the wild type repressor, Rep, forming a heterodimer of Vir and Rep within bacterial cells. Nakai and co-workers (Marshall-Batty and Nakai 2003) attached cAMP-dependent kinase (PKA) motif to the C-terminus of Rep. The dimerization of Vir and Rep induces a conformational change on Rep and causes exposure of the PKA motif on Rep. Consequently, interaction between Vir and Rep increases the rate of Rep phosphorylation by PKA *in vitro*. These studies show that protein-protein interactions can affect protein phosphorylation both positively and negatively by changing accessibility of phosphorylation sites to protein kinases.

The work described in this chapter involves the investigation of the relationship between phosphorylation of Sml1 by Dun1 and interaction of Sml1 with Rnr1. Sml1 interacts with both Dun1 and Rnr1 as demonstrated by several studies (Zhao, Muller et al. 1998; Zhao, Georgieva et al. 2000; Ho, Gruhler et al. 2002). An interaction of Sml1 with Dun1 is responsible for phosphorylation of Sml1 by the Dun1 kinase, leading to the removal of Sml1. On the other hand, Sml1 inhibits RNR activity by interacting with

Rnr1. The first question to address was whether the Sml1-Rnr1 interactions influenced the degree of Sml1 phosphorylation by the Dun1 kinase. The data show not only the effect of Rnr1 on Sml1 phosphorylation but also possible structural insights of the Sml1-Rnr1 interactions concerning whether the phosphorylation sites of Sml1 are on the interface of the Sml1-Rnr1 complex. The second question involved using a phospho-mimic Sml1 mutant to address whether phosphorylation of Sml1 had a significant effect on its ability to inhibit RNR. The data suggests that Sml1 phosphorylation has a negative effect on RNR inhibition by Sml1. Due to time constraints, the question of whether phosphor-Sml1 could directly interact with Rnr1 was not addressed. In the end of this chapter, further experiments to address this issue will be proposed. (The experimental procedures employed here are described in Chapter 2.)

Results

Before performing the experiments to address the major issues, the activities of Sml1 and Rnr1 used for this study were measured to ensure that these proteins used for this study were biologically active. Figure 38 shows the activity of RNR defined as the velocity of the steady state kinetics. In this particular case, specific activity of Rnr1 was 103nmol/min/mg, which is similar to values previously reported by other groups (Nguyen, Ge et al. 1999; Chabes, Domkin et al. 2000). In repeated experiments, 40 to 110nmol/min/mg of Rnr1 specific activities were obtained.

Next, the RNR assay was performed in the presence of Sml1 ranging from 0 to 16.6μM following the procedure described by Thelander and co-workers

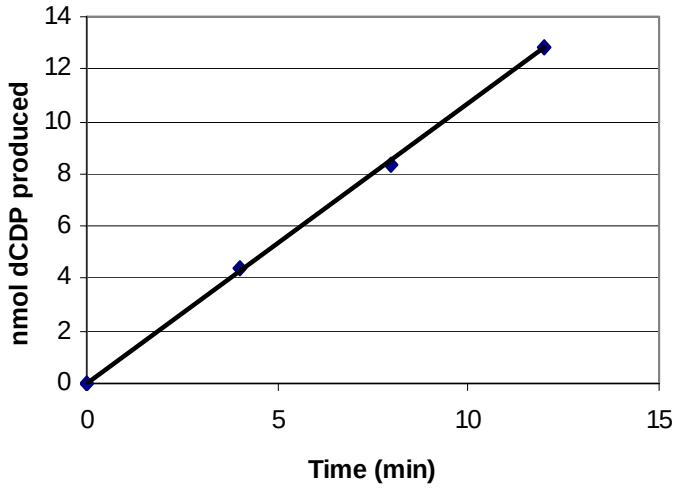

Figure 38. RNR activity in steady state kinetics.

RNR activity assay was performed with 1μM Rnr1 dimer and 5μM Rnr2/Rnr4 heterodimer. 50μl of reaction mixture was taken and boiled at 0, 4, 8, 12 minutes over time course of the reaction. Based on radioactivity given by dCDP in each aliquot, the amount of dCDP produced up to each time point was determined. The slope of the curve represents the velocity of the reaction, which is defined as activity. Based on the amount of Rnr1 in each aliquot (10μg or 100pmol), the specific activity was determined. In this case, the activity is 1.06nmol/min and specific activity is 106nmol/min. Triplication of the same experiment gave specific activity of 103±3nmol/min/mg

(Chabes, Domkin et al. 1999). A gradual decrease of RNR activity was observed as concentration of Sml1 increased in the reaction mixture (Figure 39). A molar ratio of 1:1 between Sml1 and Rnr1 monomer (2μM Sml1 and 2μM Rnr1) gave approximately 50% inhibition, similar to the extent of inhibition reported by Thelander and co-workers (Chabes, Domkin et al. 1999).

To test if the phosphorylation of Sml1 can be affected by interaction between Sml1 and Rnr1, a Dun1 kinase assay was performed with 3μM of Sml1 in the presence of Rnr1 at various concentrations. A gradual decrease of phosphorylation was observed at higher concentration of Rnr1 (Figure 40A & B). At 20 to 30μM of Rnr1 monomer, the curve appeared to plateau at lower levels of phosphate incorporation. It should be noted that even at this plateau, the radioactivity was well above the background level, indicating that a small amount of Sml1 is still phosphorylated.

As a negative control, the same experiment was conducted using the mutant E71A/R72A Sml1. Previously Rothstein and co-workers (Zhao, Georgieva et al. 2000) reported that R72G Sml1 did not show interaction with Rnr1 *in vivo*. Similarly, *in vitro* inhibition of RNR was significantly impaired by the R72A mutation (Zhao, Georgieva et al. 2000). We also observe that the ability of E71A/R72A Sml1 to inhibit RNR is significantly less than that of wild type Sml1 (Figure 39B). As expected, we did not observe significant change in phosphorylation of E71A/R72A Sml1 in the presence of Rnr1, supporting that the decrease of wild type Sml1 phosphorylation in the presence of Rnr1 is due to the Sml1-Rnr1 interactions. Although the data of E71A/R72A Sml1 had some fluctuations, statistical analysis showed that there is no significant correlation

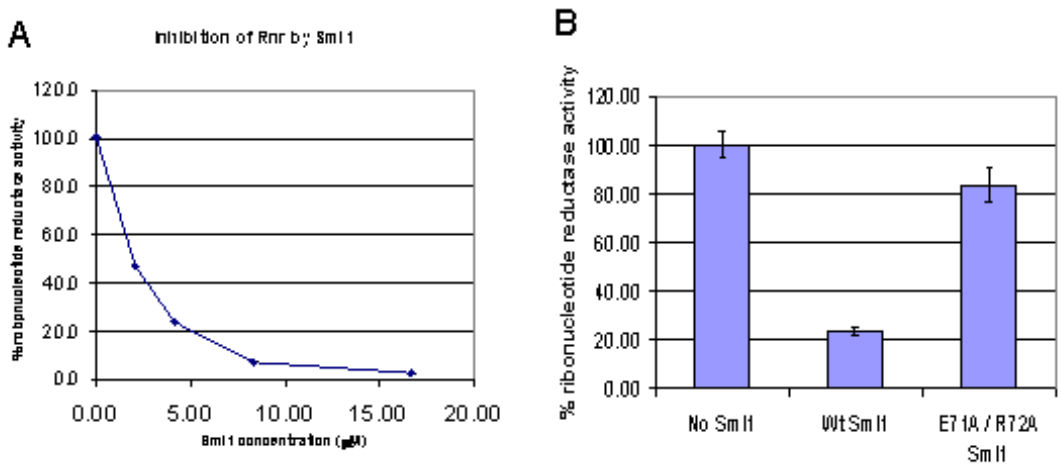

Figure 39. Inhibition of RNR by Sml1.

(A) RNR inhibition by wild type Sml1. An RNR activity assay was performed with 2μM Rnr1 monomer, 2.5μM Rnr2/Rnr4 heterodimer in the presence of 0, 2.1, 4.2, 8.3 or 16.6μM of Sml1 monomer. Each data point represents average and standard deviation of duplicates. The data was normalized by taking the specific activity in the absence of Sml1 as 100%. (B) Comparison of RNR inhibition by wild type and E71A/R72A Sml1. RNR activity assay was performed with 1μM Rnr1 dimer and 2.5μM Rnr2/Rnr4 in the absence or presence of 4.2μM of wild type or E71A/R72A Sml1 monomer. Each data point represents average and standard deviation of duplicates. The data was normalized by taking the specific activity in the absence of Sml1 as 100%.

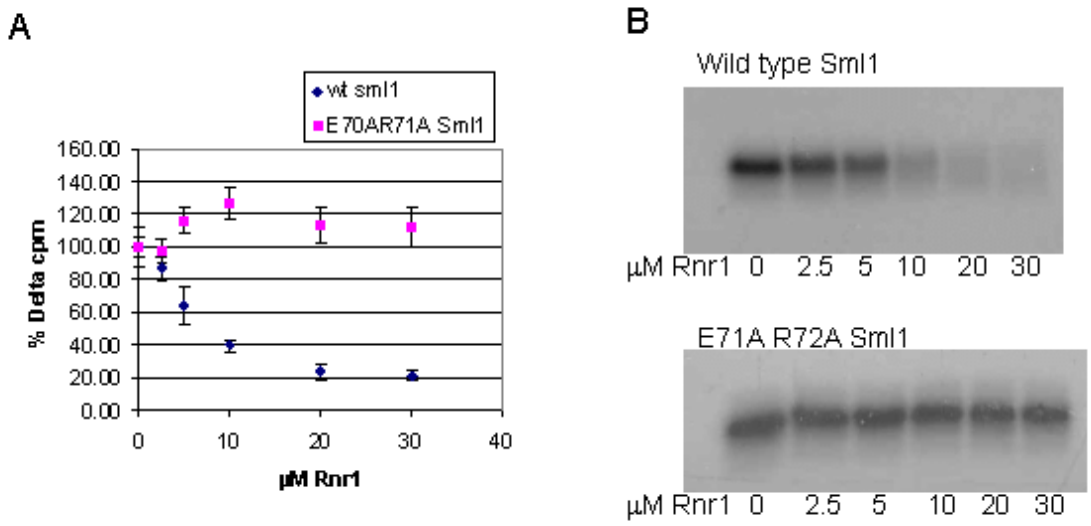

Figure 40. Phosphorylation of Sml1 in the presence of Rnr1.

30µl of 250mM ATP, 0.06µCi/ul γ[^{32}P]ATP, 3µM wild type or E71A/R72A Sml1 monomer and Rnr1 monomer ranging from 0 to 30µM were incubated with 15µl bed volume of GST-Dun1 bound to glutathione resin in 30ºC for 30 minutes. **(A)** P81-cellulose filter assay. After the reaction, 20µl of the supernatant was applied to P81 phospho-cellulose filter and the radioactivity of ^{32}P covalently attached to Sml1 was measured by liquid scintillation counter. Amount of radioactivity from the protein (delta cpm = cpm of protein sample – cpm of background) was determined first. The data was normalized by taking the delta cpm in the absence of Rnr1 as 100%. **(B)** Autoradiography in the experiment with wild type Sml1. 10µl of the supernatant was applied to SDS-PAGE gel and the phosphorylated Sml1 was visualized.

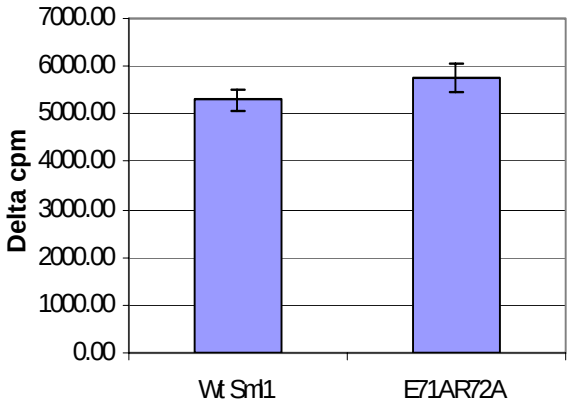

Figure 40 continued
(C) Phosphorylation of wild type Sml1 and E71A/R72A Sml1 was performed on the same day under the same circumstances. Using GST-Dun1 from the same batch of preparation, phosphorylation of wild type and E71A/R72A Sml1 in the absence of Rnr1 was simultaneously measured as in (A).

between phosphorylation of E71A/R72A Sml1 and the concentration of Rnr1 ($p=0.259$). On the other hand, there is a significant correlation between phosphorylation of wild type Sml1 and the concentration of Rnr1 ($p<0.001$). Autoradiography showed consistent results. The experiments with wild type Sml1 and E71A/R72A Sml1 were performed on different days, and the overall level of phosphorylation in E71A/R72A was higher than that in wild type Sml1. However, when phosphorylation of wild type and E71A/R72A Sml1 were performed on the same day under the same circumstances, they are phosphorylated in the same extent (Figure 40C).

Next, to test whether the decrease of Sml1 phosphorylation in the presence of Rnr1 is due to the reduced accessibility caused by the Sml1-Rnr1 interactions, the following experiments were performed. A fluorescence probe, *N*,*N*'-dimethyl-*N*-(iodoacetyl)-*N*'-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)ethylenediamine (IANBD amide), was covalently conjugated to residue 14 of wild type Sml1 (IANBD-wt Sml1) and residue 60 of the C14S/S60C Sml1 (IANBD-S60C Sml1), respectively. A characteristic of IANBD amide is that its emission intensity increases when the probe moves to a less solvent accessible environment. Fluorescence from respective proteins was monitored before and after addition of Rnr1.

Only a small increase of fluorescence intensity was observed when Rnr1 was added to IANBD-wt Sml1 (Figure 41A). On the other hand, more than a three-fold increase of fluorescence intensity was observed in the same experiment with IANBD-S60C Sml1 (Figure 41B). Previously, we identified residue 60 as the primary

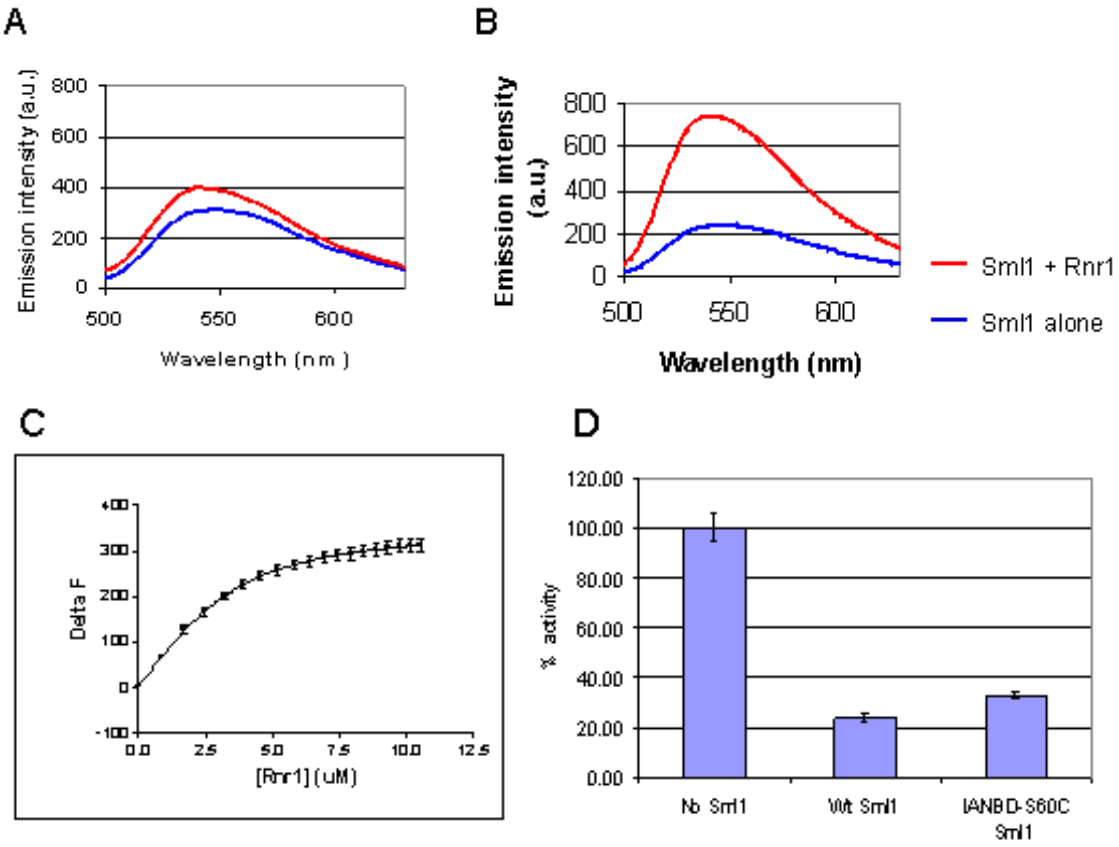

Figure 41. Fluorescence based assay of the Sml1-Rnr1 interaction.

(A) Fluorescence emission scan of wild type Sml1 conjugated with IANBD amide at residue 14 (IANBD-Sml1). Emission scan was performed on 5μM of IANBD-Sml1 monomer first (lower spectrum shown in red). Next, Rnr1 was added to a final concentration of 16μM, while maintaining Sml1 concentration at 5μM. Then, emission scan was performed on the Sml1 / Rnr1 mixture. **(B)** Emission scan of C14S / S60C Sml1 conjugate with IANBD (IANBD-S60C Sml1) at residue 60 before and after addition of Rnr1 performed as in (A). (**C)** Titration of IANBD- S60C Sml1 with Rnr1 (detailed procedure is described in Chapter 2). Each data point represents average and the standard. **(D)** RNR inhibition by IANBD-S60C Sml1. RNR activity assay was performed with 1μM Rnr1 dimer and 2.5μM Rnr2/Rnr4 heterodimer in the absence or presence of 4.2μM of wild type and IANBD-S60C Sml1 monomer. Each data point represents average and standard deviation of triplicates (No Sml1 and wt Sml1) or duplicates (IANBD S60C Sml1).

phosphorylation site of Sml1 (Uchiki et al., 2004). Therefore, our data indicates that the phosphorylation sites of Sml1 become less solvent accessible when Sml1 binds to Rnr1. As a negative control, IANBD-amide was conjugated with beta-mercaptoethanol, and it was mixed with Rnr1 in the same manner. No significant change of fluorescence intensity was observed, demonstrating that increased fluorescence intensity was not due to non-specific interaction between IANBD and Rnr1 or causes other than the Sml1-Rnr1 interactions.

To quantify the Sml1-Rnr1 interactions, IANBD-S60C Sml1 was titrated with Rnr1 and the emission intensities after each titration were recorded (Figure 41C). In a buffer used for the preparation of Rnr1 and the phosphorylation of Sml1, we obtained a dissociation constant (K_d) of 1.07±0.48μM, which is within the range of the value (0.4±0.1μM) reported by Thelander and co-workers (Chabes, Domkin et al. 1999). However, comparison of RNR activity in the presence of IANBD-S60C Sml1 and wild type Sml1 shows that IANBD-S60C Sml1 has slightly less ability to inhibit RNR (Figure 41D). This indicates that IANBD-S60C Sml1 has less affinity than wild type Sml1.

To test the effects of Sml1 phosphorylation on inhibiting RNR, an RNR activity assay was performed in the presence of S60D phospho-mimic Sml1 as well as wild type Sml1 (Figure 42). S60D Sml1 showed a decreased level of RNR inhibition than wild type Sml1, although it still inhibits RNR significantly.

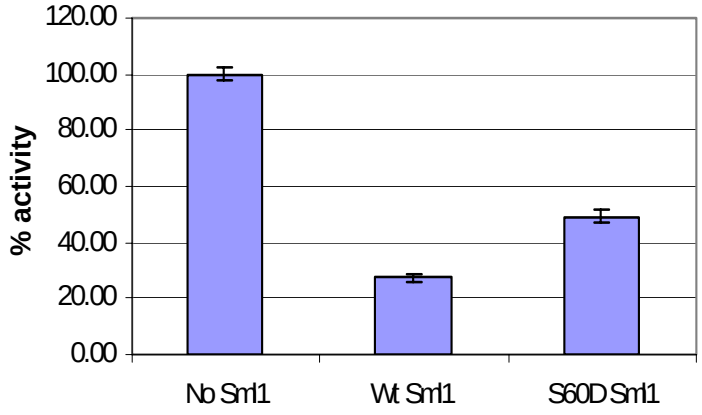

Figure 42. Inhibition of RNR by S60D Sml1.

RNR activity assay was performed with 1μM Rnr1 dimer and 2.5μM Rnr2/Rnr4 heterodimer in the absence or presence of 4.2μM of wild type or S60D Sml1 monomer. Each data point represents average and standard deviation of triplicates.

Discussion

Our study demonstrated that Rnr1 could significantly reduce phosphorylation of Sml1 by the Dun1 kinase through its interaction with Sml1. On the other hand, it still remains uncertain whether or not Sml1 bound to Rnr1 can be phosphorylated. The data presented above indicate that Rnr1 cannot completely compete out Dun1. Sml1 was still weakly phosphorylated in the presence of Rnr1 at the concentration (30μM Rnr1 monomer in Figure 40A) where the descending curve of Sml1 phosphorylation *appears* to reach a plateau. In addition, if we take the dissociation constant (K_d) for the Sml1-Rnr1 interactions to be approximately 1.1μM (as shown in Figure 41C), approximately 0.1μM of unbound Sml1 must be present at 3μM of Sm1 and 30μM Rnr1[26]. Dun1 kinase assay with various concentrations of Sml1, no phosphorylation of Sml1 was detected at ~0.1μM of Sml1 (data not shown). These observations suggest that Sml1 may be phosphorylated weakly when it is bound to Rnr1. However, more observations of Sml1 phosphorylation at higher concentration of Rnr1 will be needed to conclude this issue.

Our study also suggests that unbound Sml1 is largely phosphorylated by Dun1. As shown in Figure 40A, the extent of Sml1 phosphorylation can be dependent on the concentration of Rnr1. In addition to the data shown above, when we initially conducted

---

[26] Assuming that stoichiometry of Sml1 and Rnr1 in the complex is 1 to 1, K_d of the Sml1-Rnr1 interaction is expressed as a quadratic equation of the concentration of the Sml1-Rnr1 complex in terms of total concentration of Sml1 and Rnr1:

$K_d=([Sml1_{total}]-[Complex]) \times ([Rnr1_{total}] - [Complex]) / [Complex]$

Where $[Sml1_{total}\}$=Total concentration of Sml1, $[Rnr1_{total}]$=Total concentration of Rnr1. and [Complex}=Concentration of the Sml1-Rnr1 complex.

For example, when Kd is 1.1μM and total concentration of Sml1 and Rnr1 is 3μM and 30μM respectively, 2.88μM of the Sml1-Rnr1 concentration is obtained by solving the above equation. Therefore, the concentration of unbound Sml1 in this condition is 0.12μM ($[Sml1_{total}]$-[Complex]=3μM-2.88μM=0.12μM).

phosphorylation of Sml1 at 8μM (>2.5 times higher than the concentration in Figure 40A & B), no significant reduction of Sml1 phosphorylation was observed in the presence of Rnr1 at low concentrations (2 or 5μM). This was probably because the concentration of free Sml1 was well above the Km of Sml1 for its phosphorylation by the Dun1 kinase, and the Dun1 kinase was saturated with Sml1 at these conditions. Generally in Michaelis-Menten kinetics, at a range of substrate concentrations well above Km of the reaction, change in substrate concentration has only a small effect on the velocity of the reaction. When the total concentration of Sml1 is 8μM and that of Rnr1 is 2 to 5μM, the concentration of unbound Sml1 is estimated to be 4 to 7μM[26]. Km of Sml1 for its phosphorylation is possibly well below 4 to 7μM. Although it is difficult to accurately determine Km of Sml1 phosphorylation by Dun1 in our current experimental scheme, in the future, it will be important to, at least roughly, estimate the Km of the reaction.

The experiment involving fluorescence spectroscopy showed that the primary phosphorylation site of Sml1 becomes less solvent accessible when Sml1 is bound to Rnr1. This explains our first observation that Rnr1 reduces Sml1 phosphorylation. Accessibility of substrate residues to protein kinases is an important factor that affects phosphorylation. Recent statistical analysis of the specificity of protein kinases show that 65-70% of phosphorylated residues are located on the surface of the protein, while the probability of Ser/Thr/Tyr residues occurring in the interior of a protein is almost equal to its occurrence in the exterior of a protein (Kreegipuu, Blom et al. 1998; Weckwerth and Selbig 2003). Our data show that the accessibility of phosphorylation sites is a major factor for protein phosphorylation.

An important issue is whether phosphorylation of Sml1 *in vivo* is significantly reduced by its interaction with Rnr1. Rothstein and co-workers (Zhao, Chabes et al. 2001) reported that there is no significant difference between the degradation profiles of wild type versus the I76T and S87P Sml1 mutants[27] in the presence of DNA damage and replication blocks. Based on a yeast two hybrid assay, binding of I76T and S87P Sml1 to Rnr1 was found to be significantly impaired (Zhao, Georgieva et al. 2000). These studies concluded that the degradation of Sml1 *in vivo* is triggered by phosphorylation of Sml1 (Zhao, Chabes et al. 2001; Zhao and Rothstein 2002) and that the Sml1-Rn1 interaction has no significant effect on Sml1 degradation *in vivo* (Zhao, Chabes et al. 2001). Based on these findings, one can possibly conclude that the Sml1-Rnr1 interactions are not a significant determinant of phosphorylation of Sml1 *in vivo*. However, the data presented here shows that *in vitro* phosphorylation of Sml1 is significantly reduced by the Sml1-Rnr1 interactions. A possible explanation for the discrepancy between their *in vivo* study and our *in vitro* study is that the population of free Sml1 *in vivo* may be much larger than that of Sml1 bound to Rnr1, and the mutation to disrupt the Sml1-Rnr1 interactions may not cause significant change in the overall concentration of free Sml1. However, since the comparison of the levels of *in vivo* phosphorylation in wild type versus the I76T and S87P Sml1 mutants has not been investigated, we cannot exclude the possibility that the Sml1-Rnr1 interactions may still affect Sml1 phosphorylation *in vivo* to an extent.

Finally, we addressed whether the phosphorylation of Sml1 potentially affects the ability of Sml1 to inhibit RNR. S60D phospho-mimic Sml1 showed reduced levels of

---

[27] These mutations were induced in the chromosomal copy of SML1 gene.

RNR inhibition[28], suggesting that the phosphorylation of Sml1 may also impair activity of Sml1. However, from our current data, we cannot conclude whether the reduced RNR inhibition by S60D Sml1 is due to weaker affinity of S60D Sml1 or any other unknown cause. Further experiments need to be performed to unequivocally address this issue.

Further experiments

This section proposes the experiments that should be done to address issues that could not be fully addressed in Chapter 6. These issues are:

1. Whether phosphorylation of Sml1 by the Dun1 kinase partially takes place when Sml1 is bound to Rnr1,
2. Whether phosphorylation of Sml1 affects binding of Sml1 to Rnr1,
3. Whether phosphorylation of Sml1 affects its ability to inhibit RNR.

Although these issues have been addressed to a certain extent, further experiments are necessary. For each issue, problems associated with our current experimental scheme will be pointed out first. Then, possible solutions and alternative strategies will be discussed.

Unequivocally determine if Sml1 bound to Rnr1 can be weakly phosphorylated

The data presented above certainly showed that the Sml1-Rnr1 interactions reduce phosphorylation of Sml1 *in vitro*. However, it remains uncertain if Sml1 is partially

[28] Thelander and co-workers also performed RNR activity assay in the presence of S60D, S58D/S60D, and S56D/S58D/S58D Sml1. Their result showed that all these mutants are less active than wild type Sml1 and the more acidic residues are introduced at the phosphorylation sites, the less activity of Sml1 was observed. (Unpublished information obtained through personal communications.)

phosphorylated when it is bound to Rnr1. Sml1 phosphorylation was still observed in the presence of Rnr1 at the concentration where the descending curve of Sml1 phosphorylation *appears* to reach a plateau. However, more observations of Sml1 phosphorylation should be made at higher concentration of Rnr1 to conclude that Sml1 is still phosphorylated when it is saturated with Rnr1.

Using the experimental scheme presented in this study, it is still possible to conclude that the phosphorylation of Sml1 takes place when Sml1 is bound to Rnr1 provided we can demonstrate the following:

A. Sml1 is phosphorylated in the presence of Rnr1 at a concentration of Sml1 and Rnr1 where a majority of Sml1 is bound to Rnr1.
B. At a concentration equivalent to the concentration of unbound Sml1 in A (in the absence of Rnr1), Sml1 is phosphorylated to a significantly lesser extent than A.

One difficulty in demonstrating these points is the inability to accurately estimate the unbound Sml1 concentration in the presence of Rnr1. Unbound Sml1 concentration is calculated based on the dissociation constant of the Sml1-Rnr1 interactions. Currently, error associated with dissociation constant (K_d) of the Sml1-Rnr1 interactions is relatively high (K_d=1.07±0.47μM), when the K_d is determined by the fluorescent based binding assay (problems associated with this assay are further discussed in the Appendix). In addition, comparison of RNR activity in the presence of IANBD-S60C Sml1 and wild type Sml1 shows that IANBD-S60C Sml1 has slightly less ability to inhibit RNR (Figure 41D). This indicates that IANBD-S60C Sml1 has less affinity than

wild type Sml1. Currently, IANBD-S60C Sml1 is used for the binding assay and wild type Sml1 is used for the Dun1 kinase assay. Therefore, the concentration of unbound Sml1 in the kinase assay (in the presence of Rnr1) may be underestimated.

An alternative strategy is to covalently tether Sml1 to Rnr1. If Sml1 and Rnr1 can be successfully tethered, it should be possible to specifically observe phosphorylation of Sml1 tethered to Rnr1 by autoradiography, because molecular weights of Sml1 and Rnr1 are 12 kDa and 100 kDa respectively. This experiment is rather qualitative, and the inaccuracy of the protein concentration in the reaction mixture would not be a serious problem. The difficult step for this strategy will be to find an appropriate method of cross-linking. Moreover, it will be necessary to ensure that Sml1 conjugated to a cross linker (but not Rnr1) is phosphorylated in the same extent as un-conjugated Sml1. It will be necessary to find sites on Sml1 that is distant from the phosphorylation site that can be used for cross-linking. A strategy to find such sites is to introduce a cysteine within a region spanning between residues 72-104 on C14S Sml1 and conjugate cross linkers that react with cysteine at one end. In the mass spectrometric analysis presented in chapter 5, we have demonstrated that Sml1 can be phosphorylated without residues 72 to 104. So, it is unlikely that the conjugation of the cross-linker within these residues interferes with the phosphorylation of Sm1l. On the other hand, such conjugations would potentially interfere with binding of Sml1 to Rnr1, and screening several residues may be necessary. The type and length of cross-linker reagents may be another consideration for successfully tethering the two proteins and not interfering with phosphorylation due to the conjugation of the reagent to Sml1.

Test if phosphorylation of Sml1 affects binding of Sml1 to Rnr1.

Our current fluorescence based assay to monitor the Sml1-Rnr1interactions requires conjugation of the fluorescence probe to residue 60 of Sml1, which is the primary phosphorylation site. Therefore, this method cannot be used to monitor binding of either phospho-mimic Sml1 or phosphorylated Sml1 to Rnr1. In order to do so, the site of the fluorescence probe attachment needs to be moved to another location of Sml1 distant from the phosphorylation sites.

An alternative strategy is to use different techniques for monitoring the Sml1-Rnr1 interactions. For example, measuring fluorescence anisotropy might be one such choice. Since the Sml1 dimer is 24 kDa and Rnr1 dimer is 200 kDa, the tumbling rate of Sml1 molecule may be significantly reduced upon binding of Sml1 to Rnr1. When using conjugated Sml1 with a fluorescence probe, fluorescence polarization within a certain time window (fluorescence anisotropy) might be significantly increased upon the binding event. An advantage of this method is that it may be appropriate to conjugate the fluorescence tag away from the region involved in the Sml1-Rnr1 interactions and from the phosphorylation sites. For example, Cys14 on wild type Sml1 may be a candidate site for the conjugation. In this way, the potential problem of interfering with the Sml1-Rnr1 interactions or with phosphorylation may be possibly avoided.

Nevertheless, it will be still difficult to test the binding of phosphorylated Sml1 to Rnr1 (but not the phospho-mimic Sml1 mutant). Fluorescence based assays described above may require significant amounts of Sml1 separated from unphosphorylated Sml1.

However, it may be possible to monitor binding of phospho-Sml1 to Rnr1 in a more qualitative method as described below.

GST-pull down assays using GST fusion Rnr1 (GST-Rnr1) may be such a possible qualitative method. We have obtained a yeast expression plasmid for GST-Rnr1 from Dr. Rodney Rothstein as a personal gift. This plasmid was previously used to demonstrate the Sml1-Rnr1 interactions *in vivo* (Zhao, Muller et al. 1998). First, by applying ^{32}P labeled phospho-Sml1 to GST-Rnr1 bound to the glutathione resin, it may be possible to determine whether or not phospho-Sml1 binds to Rnr1. Second, it may be possible to determine the relative affinity of phosphorylated and unphosphorylated Sml1 by measuring the relative ratio of phosphorylated and unphosphorylated Sml1 in the following samples; (1) A mixture of phosphorylated and unphosphorylated Sml1 before being applied to GST-Rnr1, (2) Sml1 bound to GST-Rnr1, and (3) Sml1 in the flow-through from the GST-Rnr1 column. Recently, we found that the phosphorylated Sml1 can be separated from the unphosphorylated form on SDS-PAGE[29]. It is also possible to analyze the relative amount of Sml1 by quantitative immunoblotting using an antibody against Sml1. Although we have not tested the antibody against phosphorylated Sml1, we have successfully performed immunoblots to detect free Sml1. (This antibody is a personal gift from Dr. Rodney Rothstein.) In these GST pull-down assays, it will be also

[29] SDS-PAGE gel used for autoradiography shown in Figure 38B was stained with Coomassie Blue. Two bands were observed; the slowly migrating band corresponding to migration of phosphorylated Sml1 and fast migrating band corresponding to the unphosphorylated Sml1. The slowly migrating band was observed in only the sample containing phospho-Sml1 while another band was present in all the samples including the negative control which does not contain phospho-Sml1 (data not shown).

important to use proper negative controls to take account for non-specific binding of proteins to glutathione resin or GST.

Test if phosphorylation of Sml1 affects inhibition of RNR

The bottleneck for measuring activity of phospho-Sml1 is to produce a sufficient amount of phospho-Sml1. In our current method to phosphorylate Sml1 *in vitro*, at best, about 40% of 8μM Sml1 can be phosphorylated in about 500μl, which may provide the amount of Sml1 required for measuring inhibition of RNR by Sml1. The total amount of phospho-Sml1 produced in this method depends on the amount of GST-Dun1[30]. Our data have shown that 50% inhibition of RNR activity can be seen with approximately 2μM of Sml1 (Figure 39), and the difference between inhibition of RNR at 2μM and 4μM of Sml1 can be clearly observed. Therefore, if the ability of phosphorylated Sml1 to inhibit RNR is largely different from that of the unphosphorylated form, we may be able to qualitatively observe the difference in inhibition of RNR between the sample containing only unphosphorylated Sm1l versus the sample in which a fraction (30 to 40%) Sml1 is phosphorylated.

Another strategy to produce a large quantity of phospho-Sml1 is to use other kinases that can be produced in bulk. As discussed in chapter 5, Ser60 is mainly phosphorylated by the Dun1 kinase, and it is also a potential phosphorylation site for Casein Kinase II (CKII). We tested this possibility by performing *in vitro*

[30] In our current experimental scheme, *in vitro* phosphorylation of Sml1 levels off approximately one hour after the reaction is started even thorugh Sml1 and ATP are not depeleted. The reason for the decrease of GST-Dun1 activtiy is currently unknown. Even though GST-Dun1 is supposed to be catalytic in the reaction, yield of phospho-Sml1 depends on amount of GST-Dun1 in the reaction mixture.

phosphorylation of wild type and S60A Sml1 by CKII (Figure 43). Although the activity of CKII towards Sml1 was more than 10-fold weaker than its activity towards a synthetic peptide (RRREEETEEE), our result clearly shows that CKII phosphorylates Ser60. CKII (purchased from CALBIOCHEM) used for this study is expressed in *E.coli*. As compared to the Dun1 kinase expressed in S.cerevisiae, it will be easier to prepare CKII expressed in bacteria. Several groups have shown *in vitro* kinase activity of non-commercial CKII expressed in bacteria (Antonelli, Daniotti et al. 1996; Benetti, Kim et al. 1998). Separation of phosphorylated and unphosphorylated Sml1 may be necessary to quantitatively compare their ability to inhibit RNR or binding to Rnr1. As described above, we have observed separation of phosphorylated and unphosphorylated Sml1 on SDS-PAGE gel. Although SDS-PAGE is not a choice for preparative separation of proteins, this suggests that phosphorylated and unphosphorylated Sml1 have significantly different physical properties. Methods based on charge differences of proteins such as preparative electrophoresis or ion exchange chromatography may be used to purify phosphorylated Sml1.

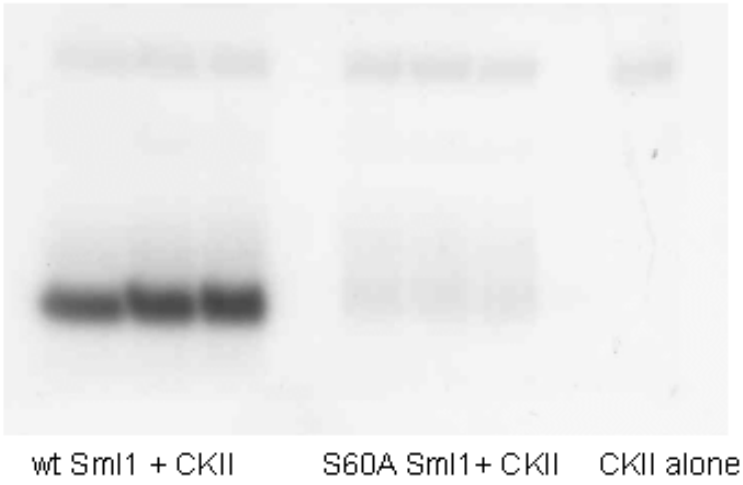

Figure 43. Phosphorylation of Sml1 by Casein kinase II (CKII).

In vitro phosphorylation of Sml1 was performed by incubating 17μM wild type or S60A Sml1, 25U/ml CKII, 60μM ATP and 0.1μCi/ul [γ-^{32}P]ATP in 30°C for 30min. The sample was resolved on 15% polyacrylamide SDS-PAGE gel. The gel was exposed to X-ray film for 24 hours. Lane 1-3. Wild type Sml1 incubated with CKII. Lane 5-7. S60A Sml1 incubated with CKII. Lane 9. A negative control in which Sml1 was omitted from the reaction mixture.

Summary

Phosphorylation of wild type Sml1 in the presence of increased concentrations of Rnr1 shows that Rnr1 reduces phosphorylation of Sml1 *in vitro*. The same experiment using E71A/R72A Sml1, which does not interact with Rnr1, further supports that reduced phosphorylation of wild type Sml1 in the presence of Rnr1 was due to interaction between Sml1 and Rnr1. A fluorescence study conjugating a solvent-accessibility-sensitive probe (IANBD-amide) to the primary phosphorylation site of Sml1 (Ser60) showed that the phosphorylation site becomes less accessible when Sml1 is bound to Rnr1. This indicates that the decreased phosphorylation of Sml1 in the presence of Rnr1 is due to reduced accessibility of Sml1's phosphorylation sites. Our data also suggests a possibility that Sml1 is phosphorylated to a lesser extent when it is bound to Rnr1, although our current data is not sufficiently precise to quantitatively conclude this issue. Experiments to further address this issue were discussed.

The RNR activity assay with S60D phospho-mimic Sml1 exhibited reduced level of RNR inhibition, suggesting that the inhibitory activity of Sml1 is likely to be reduced by phosphorylation of Ser60. Experiments to further address the effects of phosphorylation on activity of Sml1 and the Sml1-Rnr1 interactions were also discussed.

Chapter 7. Conclusions

Biological impact and technical advance made by this dissertation research

The overall goal of this dissertation research was to elucidate the structure-function and regulation of Sml1. More specifically, this work addressed the following three aims: (1) the biochemical characterization of recombinant Sml1, (2) the elucidation of Sml1 phosphorylation, and (3) the investigation of the effects of Rnr1 on Sml1 phosphorylation. The work in this thesis provides structural insights towards the oligomerization and phosphorylation of Sml1. Furthermore, this study addresses the broader issue of DNA damage and replication blocks with respect to the specificity of the Dun1 kinase and the effects of protein-protein interactions on protein phosphorylation.

The first aim was to characterize the molecular nature of Sml1. Our initial goal was to use mass spectrometry and gel filtration chromatography to analyze the histag, wild type, and C14S Sml11 for crystallization. These analytical techniques are commonly to identify crystallization problems such as proteolytic degradation and non-specific oligomerization of proteins, which deter crystallization. At the commencement of this work, only a few publications were available describing some level of biochemical characterization of Sml1 (Zhao, Muller et al. 1998; Chabes, Domkin et al. 1999). The following biochemical aspects of Sml1 were demonstrated:

1. Sml1 can form an intermolecular disulfide linked dimer through Cys 14
2. Sml1 carries Na^+/K^+ adducts, which are most prominent in Sml1-histag
3. Sml1 can also form an oligomer through a non-covalent interaction.

Furthermore, a truncated form of C14S Sml1 was identified, and this was later used to gather important information regarding Sml1 phosphorylation. Further work by other members of our group concluded that Sml1 forms dimer through non-covalent interactions involving the region spanning residues 8 to 20 (Gupta, Peterson et al. 2004).

The second aim of this research was to elucidate the regulation of Sml1 through phosphorylation. Rothstein and co-workers (Zhao, Georgieva et al. 2000; Zhao, Chabes et al. 2001) reported that Sml1 is phosphorylated by the Dun1 kinase, and the phosphorylation possibly triggers degradation of Sml1 during DNA damage / replication block responses. Following these studies, phosphorylation sites of Sml1 was identified by mass spectrometry and confirmed by site-directed mutagenesis. Based on conventional biochemical techniques, substrate specificity of the Dun1 kinase was also characterized. The work led to the following conclusions:

1. Ser 56, Ser58 and Ser60 are phosphorylated by the Dun1 kinase, among which Ser60 is the primary phosphorylation site
2. Phosphorylation of Ser60 possibly enhances phosphorylation of other residues
3. Phosphorylation patterns of doubly and triply phosphorylated Sml1 as well as the cooperativity between phosphorylation sites suggests that phosphorylation of Sml1 follows a hierarchical phosphorylation mechanism
4. Unlike a previous report (Sanchez, Zhou et al. 1997), substrate specificity of the the Dun1 kinase is not similar to cAMP dependent kinase.
5. Instead, the requirement of Glu at the +3 position of the phospho-acceptor (Ser60) on Sml1 for phosphorylation by the Dun1 kinase as well as the *in*

vitro phosphorylation of Ser60 by Casein kinase II indicates that substrate specificity of the Dun1 kinase may be similar to that of Casein kinase II.

This was the first biochemical characterization of the Dun1 kinase using one of its natural substrates. The phosphorylation motif SASASSLE is a unique motif not previously observed. The data provided show not only the phosphorylation sites on Sml1, but also some details of the mechanism of how Dun1 phosphorylates Sml1. These results will also provide a basis for studying the degradation of Sml1. Although Sml1 is the only natural substrate of Dun1 identified to date, it is likely that there are other substrates of Dun1 involved in RNR transcriptional regulation, G2/M cell cycle arrest, and DNA repair. Information obtained in this research will be applicable to studying phosphorylation of other substrates of Dun1.

The third aim of this research was to investigate the relationship of the Rnr1-dependent phosphorylation of Sml1. First, it was important to know whether the Sml1-Rnr1 interactions affect phosphorylation of Sml1. Second, whether or not phosphorylation of Sml1 affects Sml1's ability to inhibit RNR activity. Here is a summary of the results obtained from the above experiments:

1. Phosphorylation of Sml1 by the Dun1 kinase is reduced by the Sml1-Rnr1 interactions.
2. The phosphorylation site (Ser60) of Sml1 becomes less accessible when Sml1 is bound to Rnr1.

3. The phospho-mimic S60D Sml1 has less ability to inhibit RNR activity than wild type Sml1, suggesting that the phosphorylation of Sml1 also reduces RNR inhibition by Sml1.

Combining these observations with the previously reported results by Rothstein and co-workers (Zhao, Chabes et al. 2001), our study implies that unbound Sml1 is mainly phosphorylated *in vivo*. However, the possibility that Sml1 bound to Rnr1 can be partially phosphorylated by the Dun1 kinase cannot be completely ruled out. However, the effect of Sml1 phosphorylation on Sml1's ability to inhibit RNR requires further investigation.

Based on this dissertation research combined with previous reports by other groups (Zhao, Muller et al. 1998; Chabes, Domkin et al. 1999; Zhao, Chabes et al. 2001; Zhao and Rothstein 2002), a model for phosphorylation and regulation of Sml1 is proposed as follows (Figure 44): (1) Upon DNA damage or replication blocks, free Sml1 is phosphorylated by the Dun1 kinase at Ser60 (2) Phosphorylation of Sml1 at Ser60 enhances phosphorylation at Ser56 and Ser58, (3) Phosphorylation of Sml1 weakens its ability to inhibit RNR and (4) Phosphorylation of Sml1 leads to its degradation. Based on personal communication with Dr. Rodney Rothstein at Columbia University, degradation of S56A/S58A/S60A/S61A Sml1 takes place significantly slower than that of wild type Sml1. On the other hand, degradation profiles of S60A and wild type Sml1 do not differ significantly. These observations indicate that phosphorylation of all phospho-acceptors Ser56, Ser58 and Ser60 are possibly important for degradation of Sml1. While our data showed that Ser61 is not likely to be phosphorylated *in vitro*, *in vivo* phosphorylation of Ser61 has not been fully tested. In addition, our observation in RNR

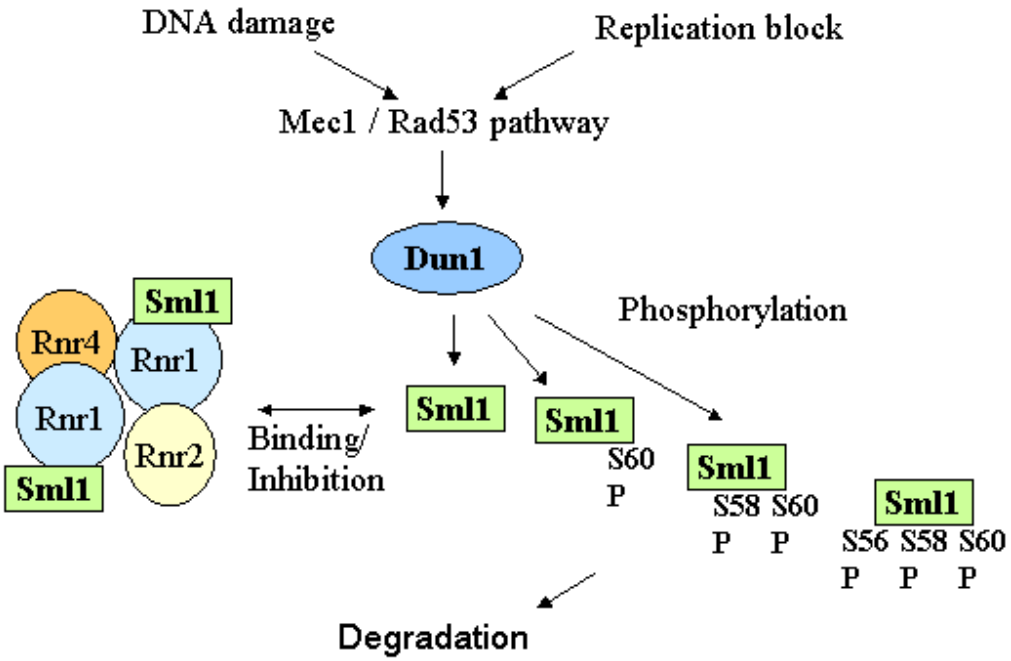

Figure 44. A new model for phosphorylation and regulation of Sml1.

This model is proposed based on this dissertation research and previous reports from other groups (Zhao, Muller et al. 1998; Chabes, Domkin et al. 1999; Zhao, Chabes et al. 2001; Zhao and Rothstein 2002).

inhibition by S60D Sml1 indicates that phosphorylation of Sml1 weakens its ability to inhibit RNR. Most likely that is because the Sml1-Rnr1 interactions are weakened by Sml1 phosphorylation. On the other hand, the Sml1-Rnr1 interactions potentially stabilize Sml1 *in vivo* (Zhao, Chabes et al. 2001). Therefore, it is possible that Sml1 phosphorylation causes dissociation of Sml1 from Rnr1 and the dissociation makes Sml1 more prone to degradation.

Our recent study (Gupta, Peterson et al. 2004) also showed that Sml1 forms a dimer through a region spanning residues 8 to 20. Previously Rothstein and co-workers reported that this region is not involved in binding of Sml1 to Rnr1 (Zhao, Georgieva et al. 2000). However, dimerization of Sml1 may have consequences in Sml1's biological function which has not been fully investigated. For example, dimerization of Sml1 may increase the effective concentration of Sml1 to interact with Rnr1. It is likely that the dimer of Sml1 interacts with the dimer of Rnr1, because the active form of the RNR large subunit is generally a dimer or a larger oligomer (Jordan and Reichard 1998) and the stoichiometry of Sml1 and Rnr1 in their complex is roughly one to one (Chabes, Domkin et al. 1999). Although RNR inhibition by Δ38 and Δ50 Sml1 has been shown, the concentration dependence of RNR inhibition by these mutants has not been fully investigated. In addition, the dimerization domain of Sml1 might have a biological function other than RNR inhibition such as interactions with other proteins or is involved in subcellular localization of Sml1.

Another important aspect of this research is the establishment of experimental schemes within our group. Although some of the techniques used in this study have been

developed elsewhere, the research enabled the group to employ more advanced techniques and handle more complex biological systems.

Here is a brief summary of the new techniques developed during this dissertation. These techniques will be employed by the Dealwis and Hettich groups: (1) basic experimental schemes to analyze proteins by ESI-FTICR mass spectrometry (FTICR-MS), (2) peptide fingerprinting and SORI-CAD of the peptides to unequivocally verify protein species, and (3) using intact proteins to analyze by FTICR-MS with high resolution and high mass accuracy. These techniques enable the identification of chemical and posttranslational modifications such as disulfide bonds, metal adducts and phosphorylation sites as demonstrated with Sml1.

Moreover, the method for conducting the RNR activity assay was learned during a visit to the Stubbe laboratory at MIT. These assays were then used to demonstrate that the Rnr1 and Sml1 prepared in our laboratory were biologically active. Based on fluorescence spectroscopy using a solvent-accessibility sensitive probe (IANBD), a preliminary method to determine the dissociation constant of the Sml1-Rnr1 interactions has been developed. This work has led to the design of several new biochemical experiments involving RNR.

Future directions

This project has reached a stage to address problems with more biological significance and more challenging technicalities. The following biological problems should be addressed as an extension of this research.

First, the *in vivo* phosphorylation of Sml1 needs to be studied in depth. So far, Rodney Rothstein's group at Columbia University has studied the degradation profile of Sml1 mutants in which putative phosphorylation sites of Sml1 was mutated (unpublished information obtained through personal communications). Nevertheless, mass spectrometry is a more direct approach to unequivocally identify *in vivo* phosphorylation sites of Sml1. A challenge for the mass spectrometric approach will be to prepare phosphorylated Sml1 that can be applied to mass spectrometry (i.e. in sufficient quantity and purity[31]). In addition, *in vivo* Sml1 phosphorylation can be quantitatively analyzed by conventional biochemical methods, such as P32 labeling, which is more reliable than methods used by other groups (Zhao, Georgieva et al. 2000; Zhao, Chabes et al. 2001).

Second, degradation of Sml1 needs to be studied in detail. For example, it is currently unknown whether Sml1 is degraded by a proteasomal dependent pathway or a vacuolar pathway. There has been no study addressing ubiquitination of Sml1. The link between phosphorylation and degradation of Sml1 also needs to be addressed further. As discussed in Chapter 1, there are at least two possible mechanisms for how phosphorylation of Sml1 can cause its degradation. They are:

1. Phosphorylation of Sml1 causes the dissociation of Sml1 from Rnr1 making it more prone to degradation.
2. There are components of the degradation machinery, such as F box proteins of ubiquitin ligases, which specifically recognize phosphorylated forms of Sml1.

---

[31] "Purity" here concerns not only presence of other proteins but also presence of small molecules such as salt, buffer and detergent.

Although it is technically challenging, these hypothesis should be tested. It should be noted that these two possible mechanims are not exclusive to each other.

Third, phosphorylation of Sml1 by the Dun1 kinase needs to be further characterized. Chapter 5 illustrated the difference in the migration on a SDS-PAGE gel between Sml1 phosphorylated by the Dun1 kinase expressed in the presence and absence of DNA damage. One possible explanation for this result is that the relative abundance of singly, doubly, and triply phosphorylated Sml1 may be different when Dun1 expressed during or in the absence of DNA damage phosphorylates Sml1. In addition, unlike the study by Elledge and co-workers (Sanchez et al., 1997), no significant difference in the overall rate of Sml1 phosphorylation by the Dun1 kinase was observed, when Dun1 is expressed in the presence and absence of DNA damage. Further experiments that address this discrepancy might provide insights into the activation and substrate specificity of Dun1. Furthermore, it is worth investigating the effects of DNA replication blocks (e.g. exposure of cells to HU) or other types of DNA damage (e.g. exposure of cells to γ-rays) on the activity and specificity of Dun1.

Fourth, as discussed in Chapter 6, the relationship between phosphorylation of Sml1 and the Sml1-Rnr1 interactions need to be further addressed. Researching this issue will provide biochemical aspects of Sml1 phosphorylation such as: (1) can phospho-Sml1 inhibit Rnr1, (2) can Sml1 complexed to Rnr1 be phosphorylated by Dun1, and (3) will the complexed Sm1 dissociate from Rnr1 upon phosphorylation.

LIST OF REFERENCES

List of references

Abraham, R. T. (2003). "Checkpoint signaling: epigenetic events sound the DNA strand-breaks alarm to the ATM protein kinase." Bioessays **25**(7): 627-30.

Abraham, R. T. (2004). "PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways." DNA Repair (Amst) **3**(8-9): 883-7.

Adamczyk, M., J. C. Gebler, et al. (2001). "Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry." Rapid Commun Mass Spectrom **15**(16): 1481-8.

Amatayakul-Chantler, S., S. W. Qian, et al. (1994). "[Ser77]transforming growth factor-beta 1. Selective biological activity and receptor binding in mink lung epithelial cells." J Biol Chem **269**(44): 27687-91.

Andersson, L. and J. Porath (1986). "Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography." Anal Biochem **154**(1): 250-4.

Antonelli, M., J. L. Daniotti, et al. (1996). "Cloning, expression and properties of the alpha' subunit of casein kinase 2 from zebrafish (Danio rerio)." Eur J Biochem **241**(1): 272-9.

Banin, S., L. Moyal, et al. (1998). "Enhanced phosphorylation of p53 by ATM in response to DNA damage." Science **281**(5383): 1674-7.

Bashkirov, V. I., E. V. Bashkirova, et al. (2003). "Direct kinase-to-kinase signaling mediated by the FHA phosphoprotein recognition domain of the Dun1 DNA damage checkpoint kinase." Mol Cell Biol **23**(4): 1441-52.

Bashkirov, V. I., J. S. King, et al. (2000). "DNA repair protein Rad55 is a terminal substrate of the DNA damage checkpoints." Mol Cell Biol **20**(12): 4393-404.

Benetti, P. H., S. I. Kim, et al. (1998). "Expression and characterization of the recombinant catalytic subunit of casein kinase II from the yeast Yarrowia lipolytica in Escherichia coli." Protein Expr Purif **13**(3): 283-90.

Benito, J., C. Martin-Castellanos, et al. (1998). "Regulation of the G1 phase of the cell cycle by periodic stabilization and degradation of the p25rum1 CDK inhibitor." Embo J **17**(2): 482-97.

Bhaumik, D. and T. S. Wang (1998). "Mutational effect of fission yeast polalpha on cell cycle events." Mol Biol Cell **9**(8): 2107-23.

Biemann, K. (1988). "Contributions of mass spectrometry to peptide and protein structure." Biomed Environ Mass Spectrom **16**(1-12): 99-111.

Bjorklund, S., S. Skog, et al. (1990). "S-phase-specific expression of mammalian ribonucleotide reductase R1 and R2 subunit mRNAs." Biochemistry **29**(23): 5452-8.

Canman, C. E., D. S. Lim, et al. (1998). "Activation of the ATM kinase by ionizing radiation and phosphorylation of p53." Science **281**(5383): 1677-9.

Chabes, A., V. Domkin, et al. (2000). "Yeast ribonucleotide reductase has a heterodimeric iron-radical-containing subunit." Proc Natl Acad Sci U S A **97**(6): 2474-9.

Chabes, A., V. Domkin, et al. (1999). "Yeast Sml1, a protein inhibitor of ribonucleotide reductase." J Biol Chem **274**(51): 36679-83.

Chabes, A., B. Georgieva, et al. (2003). "Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase." Cell **112**(3): 391-401.
Chabes, A. and L. Thelander (2000). "Controlled protein degradation regulates ribonucleotide reductase activity in proliferating mammalian cells during the normal cell cycle and in response to DNA damage and replication blocks." J Biol Chem **275**(23): 17747-53.
Chow, J. P., W. Y. Siu, et al. (2003). "Differential contribution of inhibitory phosphorylation of CDC2 and CDK2 for unperturbed cell cycle control and DNA integrity checkpoints." J Biol Chem **278**(42): 40815-28.
Climent, I., B. M. Sjoberg, et al. (1991). "Carboxyl-terminal peptides as probes for Escherichia coli ribonucleotide reductase subunit interaction: kinetic analysis of inhibition studies." Biochemistry **30**(21): 5164-71.
Coulonval, K., L. Bockstaele, et al. (2003). "Phosphorylations of cyclin-dependent kinase 2 revisited using two-dimensional gel electrophoresis." J Biol Chem **278**(52): 52052-60.
Cuesta, R., Q. Xi, et al. (2004). "Structural basis for competitive inhibition of eIF4G-Mnk1 interaction by the adenovirus 100-kilodalton protein." J Virol **78**(14): 7707-16.
Cutler, R. L., G. J. Pielak, et al. (1987). "Replacement of cysteine-107 of Saccharomyces cerevisiae iso-1-cytochrome c with threonine: improved stability of the mutant protein." Protein Eng **1**(2): 95-9.
Dass, C. (2001). Principles and Practice of Biological Mass Spectrometry. New Tork, John Wiley & Sons, Inc.
Davis, R., M. Thelander, et al. (1994). "Purification, characterization, and localization of subunit interaction area of recombinant mouse ribonucleotide reductase R1 subunit." J Biol Chem **269**(37): 23171-6.
Desany, B. A., A. A. Alcasabas, et al. (1998). "Recovery from DNA replicational stress is the essential function of the S-phase checkpoint pathway." Genes Dev **12**(18): 2956-70.
Dice, L. T. (2003). Structure function studies of Sml1, an inhibitor of DNA synthesis in Saccharomyces cerevisiae. Bioehcmistry and Cellular and Molecular Biology. Knoxville, The University of Tennessee.
Domkin, V., L. Thelander, et al. (2002). "Yeast DNA damage-inducible Rnr3 has a very low catalytic activity strongly stimulated after the formation of a cross-talking Rnr1/Rnr3 complex." J Biol Chem **277**(21): 18574-8.
Durocher, D., J. Henckel, et al. (1999). "The FHA domain is a modular phosphopeptide recognition motif." Mol Cell **4**(3): 387-94.
Durocher, D. and S. P. Jackson (2002). "The FHA domain." FEBS Lett **513**(1): 58-66.
Dyson, H. J. and P. E. Wright (2002). "Coupling of folding and binding for unstructured proteins." Curr Opin Struct Biol **12**(1): 54-60.
Elledge, S. J. (1996). "Cell cycle checkpoints: preventing an identity crisis." Science **274**(5293): 1664-72.

Elledge, S. J. and R. W. Davis (1987). "Identification and isolation of the gene encoding the small subunit of ribonucleotide reductase from Saccharomyces cerevisiae: DNA damage-inducible gene required for mitotic viability." Mol Cell Biol **7**(8): 2783-93.

Elledge, S. J. and R. W. Davis (1989). "DNA damage induction of ribonucleotide reductase." Mol Cell Biol **9**(11): 4932-40.

Elledge, S. J. and R. W. Davis (1990). "Two genes differentially regulated in the cell cycle and by DNA-damaging agents encode alternative regulatory subunits of ribonucleotide reductase." Genes Dev **4**(5): 740-51.

Elledge, S. J., Z. Zhou, et al. (1993). "DNA damage and cell cycle regulation of ribonucleotide reductase." Bioessays **15**(5): 333-9.

Eriksson, S., A. Graslund, et al. (1984). "Cell cycle-dependent regulation of mammalian ribonucleotide reductase. The S phase-correlated increase in subunit M2 is regulated by de novo protein synthesis." J Biol Chem **259**(19): 11695-700.

Fan, H., C. Villegas, et al. (1996). "Ribonucleotide reductase R2 component is a novel malignancy determinant that cooperates with activated oncogenes to determine transformation and malignant potential." Proc Natl Acad Sci U S A **93**(24): 14036-40.

Feldman, R. M., C. C. Correll, et al. (1997). "A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p." Cell **91**(2): 221-30.

Ficarro, S. B., M. L. McCleland, et al. (2002). "Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae." Nat Biotechnol **20**(3): 301-5.

Filatov, D., R. Ingemarson, et al. (1992). "The role of herpes simplex virus ribonucleotide reductase small subunit carboxyl terminus in subunit interaction and formation of iron-tyrosyl center structure." J Biol Chem **267**(22): 15816-22.

Fiol, C. J., A. Wang, et al. (1990). "Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates." J Biol Chem **265**(11): 6061-5.

Flotow, H. and P. J. Roach (1989). "Synergistic phosphorylation of rabbit muscle glycogen synthase by cyclic AMP-dependent protein kinase and casein kinase I. Implications for hormonal regulation of glycogen synthase." J Biol Chem **264**(16): 9126-8.

Fraczkiewicz, R. and W. Braun (1998). "Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules." J. Comp. Chem. **19**: 319-333.

Frimpong, K., B. G. Darnay, et al. (1993). "Syrian hamster 3-hydroxy-3-methylglutaryl-coenzyme A reductase expressed in Escherichia coli: production of homogeneous protein." Protein Expr Purif **4**(4): 337-44.

Gabrielli, B. G., M. S. Lee, et al. (1992). "Cdc25 regulates the phosphorylation and activity of the Xenopus cdk2 protein kinase complex." J Biol Chem **267**(25): 18040-6.

Gardner, R., C. W. Putnam, et al. (1999). "RAD53, DUN1 and PDS1 define two parallel G2/M checkpoint pathways in budding yeast." Embo J **18**(11): 3173-85.
Gauthier, J. W., T. R. Trauman, et al. (1991). "Sustained off-resonance irradiation for collision activated dissociation involving FTMS:Collision activated techniques that emulates infared multiphoton dissociation." Annal. Chem. Acta **246**: 211-225.
Gavin, A. C., M. Bosche, et al. (2002). "Functional organization of the yeast proteome by systematic analysis of protein complexes." Nature **415**(6868): 141-7.
Ge, J., D. L. Perlstein, et al. (2001). "Why multiple small subunits (Y2 and Y4) for yeast ribonucleotide reductase? Toward understanding the role of Y4." Proc Natl Acad Sci U S A **98**(18): 10067-72.
Ge, Y., B. G. Lawhorn, et al. (2002). "Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry." J Am Chem Soc **124**(4): 672-8.
Georgieva, B., X. Zhao, et al. (2000). "Damage response and dNTP regulation: the interaction between ribonucleotide reductase and its inhibitor, Sml1." Cold Spring Harb Symp Quant Biol **65**: 343-6.
Gilbert, C. S., C. M. Green, et al. (2001). "Budding yeast Rad9 is an ATP-dependent Rad53 activating machine." Mol Cell **8**(1): 129-36.
Goldberg, A. L. (2003). "Protein degradation and protection against misfolded or damaged proteins." Nature **426**(6968): 895-9.
Gonzales, T. and J. Robert-Baudouy (1996). "Bacterial aminopeptidases: properties and functions." FEMS Microbiol Rev **18**(4): 319-44.
Gu, Y., J. Rosenblatt, et al. (1992). "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15." Embo J **11**(11): 3995-4005.
Gupta, V., C. B. Peterson, et al. (2004). "Sml1p Is a Dimer in Solution: Characterization of Denaturation and Renaturation of Recombinant Sml1p." Biochemistry **43**(26): 8568-8578.
Hammet, A., B. L. Pike, et al. (2002). "Posttranscriptional regulation of the RAD5 DNA repair gene by the Dun1 kinase and the Pan2-Pan3 poly(A)-nuclease complex contributes to survival of replication blocks." J Biol Chem **277**(25): 22469-74.
Hanks, S. K. and A. M. Quinn (1991). "Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members." Methods Enzymol **200**: 38-62.
Hartwell, L., T. Weinert, et al. (1994). "Cell cycle checkpoints, genomic integrity, and cancer." Cold Spring Harb Symp Quant Biol **59**: 259-63.
Hartwell, L. H., J. Culotti, et al. (1974). "Genetic control of the cell division cycle in yeast." Science **183**(120): 46-51.
Hartwell, L. H. and T. A. Weinert (1989). "Checkpoints: controls that ensure the order of cell cycle events." Science **246**(4930): 629-34.
Hejazi, M., K. Piotukh, et al. (2002). "Isoaspartyl dipeptidase activity of plant-type asparaginases." Biochem J **364**(Pt 1): 129-36.
Hendrickson, C. L. and M. R. Emmett (1999). "Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry." Annu Rev Phys Chem **50**: 517-36.

Ho, Y., A. Gruhler, et al. (2002). "Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry." Nature **415**(6868): 180-3.
Hoffmann, I., G. Draetta, et al. (1994). "Activation of the phosphatase activity of human cdc25A by a cdk2-cyclin E dependent phosphorylation at the G1/S transition." Embo J **13**(18): 4302-10.
Horn, D. M., R. A. Zubarev, et al. (2000). "Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules." J. Am. Soc. Mass Spectrom. **11**: 320-32.
Huang, M. and S. J. Elledge (1997). "Identification of RNR4, encoding a second essential small subunit of ribonucleotide reductase in Saccharomyces cerevisiae." Mol Cell Biol **17**(10): 6105-13.
Huang, M., Z. Zhou, et al. (1998). "The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor." Cell **94**(5): 595-605.
Hunter, T. (1991). "Protein kinase classification." Methods Enzymol **200**: 3-37.
Hunter, T. (1994). "1001 protein kinases redux--towards 2000." Semin Cell Biol **5**(6): 367-76.
Imai, J., H. Yashiroda, et al. (2003). "Proteasomes and molecular chaperones: cellular machinery responsible for folding and destruction of unfolded proteins." Cell Cycle **2**(6): 585-90.
Janin, J. and B. Seraphin (2003). "Genome-wide studies of protein-protein interaction." Curr Opin Struct Biol **13**(3): 383-8.
Jordan, A. and P. Reichard (1998). "Ribonucleotide reductases." Annu Rev Biochem **67**: 71-98.
Kaldis, P. (1999). "The cdk-activating kinase (CAK): from yeast to mammals." Cell Mol Life Sci **55**(2): 284-96.
Kashlan, O. B. and B. S. Cooperman (2003). "Comprehensive model for allosteric regulation of mammalian ribonucleotide reductase: refinements and consequences." Biochemistry **42**(6): 1696-706.
Kashlan, O. B., C. P. Scott, et al. (2002). "A comprehensive model for the allosteric regulation of mammalian ribonucleotide reductase. Functional consequences of ATP- and dATP-induced oligomerization of the large subunit." Biochemistry **41**(2): 462-74.
Kato, R. and H. Ogawa (1994). "An essential gene, ESR1, is required for mitotic cell growth, DNA repair and meiotic recombination in Saccharomyces cerevisiae." Nucleic Acids Res **22**(15): 3104-12.
Kim, A. S., L. T. Kakalis, et al. (2000). "Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein." Nature **404**(6774): 151-8.
Kim, D. H., K. F. Faull, et al. (2004). "Borate-nucleotide complex formation depends on charge and phosphorylation state." J Mass Spectrom **39**(7): 743-51.
Knight, Z. A., B. Schilling, et al. (2003). "Phosphospecific proteolysis for mapping sites of protein phosphorylation." Nat Biotechnol **21**(9): 1047-54.

Kondo, T., T. Wakayama, et al. (2001). "Recruitment of Mec1 and Ddc1 checkpoint proteins to double-strand breaks through distinct mechanisms." Science **294**(5543): 867-70.
Kreegipuu, A., N. Blom, et al. (1998). "Statistical analysis of protein kinase specificity determinants." FEBS Lett **430**(1-2): 45-50.
Krishna, R. G. and F. Wold (1993). "Post-translational modification of proteins." Adv Enzymol Relat Areas Mol Biol **67**: 265-98.
Kriwacki, R. W., L. Hengst, et al. (1996). "Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity." Proc Natl Acad Sci U S A **93**(21): 11504-9.
Kumar, A. and M. Snyder (2002). "Protein complexes take the bait." Nature **415**(6868): 123-4.
Lapko, V. N., X. Y. Jiang, et al. (1997). "Posttranslational modification of oat phytochrome A: phosphorylation of a specific serine in a multiple serine cluster." Biochemistry **36**(34): 10595-9.
Liu, C., K. A. Powell, et al. (2003). "Cop9/signalosome subunits and Pcu4 regulate ribonucleotide reductase by both checkpoint-dependent and -independent mechanisms." Genes Dev **17**(9): 1130-40.
Liu, V. F., D. Bhaumik, et al. (1999). "Mutator phenotype induced by aberrant replication." Mol Cell Biol **19**(2): 1126-35.
Loeb, L. A. and T. A. Kunkel (1982). "Fidelity of DNA synthesis." Annu Rev Biochem **51**: 429-57.
Lopes, M., C. Cotta-Ramusino, et al. (2001). "The DNA replication checkpoint response stabilizes stalled replication forks." Nature **412**(6846): 557-61.
Mann, M., S. E. Ong, et al. (2002). "Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome." Trends Biotechnol **20**(6): 261-8.
Mann, M. and A. Pandey (2001). "Use of mass spectrometry-derived data to annotate nucleotide and protein sequence databases." Trends Biochem Sci **26**(1): 54-61.
Mann, M. and M. Wilm (1994). "Error-tolerant identification of peptides in sequence databases by peptide sequence tags." Anal Chem **66**(24): 4390-9.
Manning, G., G. D. Plowman, et al. (2002). "Evolution of protein kinase signaling from yeast to man." Trends Biochem Sci **27**(10): 514-20.
Marshall, A. G., C. L. Hendrickson, et al. (1998). "Fourier transform ion cyclotron resonance mass spectrometry: a primer." Mass Spectrom Rev **17**(1): 1-35.
Marshall-Batty, K. R. and H. Nakai (2003). "Trans-targeting of the phage Mu repressor is promoted by conformational changes that expose its ClpX recognition determinant." J Biol Chem **278**(3): 1612-7.
Medzihradszky, K. F., Z. Darula, et al. (2004). "O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes." Mol Cell Proteomics **3**(5): 429-40.
Melo, J. A., J. Cohen, et al. (2001). "Two checkpoint complexes are independently recruited to sites of DNA damage in vivo." Genes Dev **15**(21): 2809-21.

Meyer, H. E., E. Hoffmann-Posorske, et al. (1991). "Determination and location of phosphoserine in proteins and peptides by conversion to S-ethylcysteine." Methods Enzymol **201**: 169-85.
Moore, K. L. (2003). "The biology and enzymology of protein tyrosine O-sulfation." J Biol Chem **278**(27): 24243-6.
Mortz, E., P. B. O'Connor, et al. (1996). "Sequence tag identification of intact proteins by matching tanden mass spectral data against sequence data bases." Proc Natl Acad Sci U S A **93**(16): 8264-7.
Murray, A. W. (2004). "Recycling the cell cycle: cyclins revisited." Cell **116**(2): 221-34.
Nash, P., X. Tang, et al. (2001). "Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication." Nature **414**(6863): 514-21.
Nguyen, H. H., J. Ge, et al. (1999). "Purification of ribonucleotide reductase subunits Y1, Y2, Y3, and Y4 from yeast: Y4 plays a key role in diiron cluster assembly." Proc Natl Acad Sci U S A **96**(22): 12339-44.
Nickerson, D. P. and L. L. Wong (1997). "The dimerization of Pseudomonas putida cytochrome P450cam: practical consequences and engineering of a monomeric enzyme." Protein Eng **10**(12): 1357-61.
Oda, Y., T. Nagasu, et al. (2001). "Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome." Nat Biotechnol **19**(4): 379-82.
Olsen, H. B., S. Ludvigsen, et al. (1996). "Solution structure of an engineered insulin monomer at neutral pH." Biochemistry **35**(27): 8836-45.
Osborn, A. J., S. J. Elledge, et al. (2002). "Checking on the fork: the DNA-replication stress-response pathway." Trends Cell Biol **12**(11): 509-16.
Ouspenski, II, S. J. Elledge, et al. (1999). "New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability." Nucleic Acids Res **27**(15): 3001-8.
Paciotti, V., M. Clerici, et al. (2001). "Characterization of mec1 kinase-deficient mutants and of new hypomorphic mec1 alleles impairing subsets of the DNA damage response pathway." Mol Cell Biol **21**(12): 3913-25.
Pati, D., C. Keller, et al. (1997). "Reconstitution of a MEC1-independent checkpoint in yeast by expression of a novel human fork head cDNA." Mol Cell Biol **17**(6): 3037-46.
Pawson, T. (2004). "Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems." Cell **116**(2): 191-203.
Posewitz, M. C. and P. Tempst (1999). "Immobilized gallium(III) affinity chromatography of phosphopeptides." Anal Chem **71**(14): 2883-92.
Raffen, R. and F. J. Stevens (1999). "Small zone, high-speed gel filtration chromatography to detect protein aggregation associated with light chain pathologies." Methods Enzymol **309**: 318-32.
Reichard, P. (1988). "Interactions between deoxyribonucleotide and DNA synthesis." Annu Rev Biochem **57**: 349-74.
Rhind, N. and P. Russell (2000). "Checkpoints: it takes more than time to heal some wounds." Curr Biol **10**(24): R908-11.

Roepstorff, P. and J. Fohlman (1984). "Proposal for a common nomenclature for sequence ions in mass spectra of peptides." Biomed Mass Spectrom **11**(11): 601.
Rossig, L., A. S. Jadidi, et al. (2001). "Akt-dependent phosphorylation of p21(Cip1) regulates PCNA binding and proliferation of endothelial cells." Mol Cell Biol **21**(16): 5644-57.
Sanchez, Y., J. Bachant, et al. (1999). "Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms." Science **286**(5442): 1166-71.
Sanchez, Y., B. A. Desany, et al. (1996). "Regulation of RAD53 by the ATM-like kinases MEC1 and TEL1 in yeast cell cycle checkpoint pathways." Science **271**(5247): 357-60.
Sanchez, Y., Z. Zhou, et al. (1997). "Analysis of budding yeast kinases controlled by DNA damage." Methods Enzymol **283**: 398-410.
Sandell, L. L. and V. A. Zakian (1993). "Loss of a yeast telomere: arrest, recovery, and chromosome loss." Cell **75**(4): 729-39.
Schwartz, M. F., J. K. Duong, et al. (2002). "Rad9 phosphorylation sites couple Rad53 to the Saccharomyces cerevisiae DNA damage checkpoint." Mol Cell **9**(5): 1055-65.
Sherr, C. J. and J. M. Roberts (1999). "CDK inhibitors: positive and negative regulators of G1-phase progression." Genes Dev **13**(12): 1501-12.
Skowyra, D., K. L. Craig, et al. (1997). "F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex." Cell **91**(2): 209-19.
Sommerhalter, M., W. C. Voegtli, et al. (2004). "Structures of the yeast ribonucleotide reductase Rnr2 and Rnr4 homodimers." Biochemistry **43**(24): 7736-42.
Stensballe, A., O. N. Jensen, et al. (2000). "Electron capture dissociation of singly and multiply phosphorylated peptides." Rapid Commun Mass Spectrom **14**(19): 1793-800.
Stevens, F. J. (1989). "Size-exclusion high-performance liquid chromatography in analysis of protein and peptide epitopes." Methods Enzymol **178**: 107-30.
Sun, Z., D. S. Fay, et al. (1996). "Spk1/Rad53 is regulated by Mec1-dependent protein phosphorylation in DNA replication and damage checkpoint pathways." Genes Dev **10**(4): 395-406.
Syka, J. E., J. J. Coon, et al. (2004). "Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry." Proc Natl Acad Sci U S A **101**(26): 9528-33.
Tanaka, H., H. Arakawa, et al. (2000). "A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage." Nature **404**(6773): 42-9.
Tchelet, A., T. Vogel, et al. (1997). "Selective modification at the N-terminal region of human growth hormone that shows antagonistic activity." Mol Cell Endocrinol **130**(1-2): 141-52.
Terasima, T. and L. J. Tolmach (1963). "X-ray sensitivity and DNA synthesis in synchronous populations of HeLa cells." Science **140**: 490-2.
Thelander, L. (1973). "Physicochemical characterization of ribonucleoside diphosphate reductase from Escherichia coli." J Biol Chem **248**(13): 4591-601.

Thelander, L. and P. Reichard (1979). "Reduction of ribonucleotides." Annu Rev Biochem **48**: 133-58.
Toczyski, D. P., D. J. Galgoczy, et al. (1997). "CDC5 and CKII control adaptation to the yeast DNA damage checkpoint." Cell **90**(6): 1097-106.
Toh, G. W. and N. F. Lowndes (2003). "Role of the Saccharomyces cerevisiae Rad9 protein in sensing and responding to DNA damage." Biochem Soc Trans **31**(Pt 1): 242-6.
Uchiki, T., L. T. Dice, et al. (2004). "Identification of phosphorylation sites on the yeast ribonucleotide reductase inhibitor Sml1." J Biol Chem **279**(12): 11293-303.
Uchiki, T., V. Gupta, et al. (2002). Characterization of Sml1p Protein Oligometruzation by ESI-FTICR-MS. 50th ASMS Conference on Mass Spectrometry and Allied Topics, Orlando, FL.
Uchiki, T., R. Hettich, et al. (2002). "Characterization of monomeric and dimeric forms of recombinant Sml1p-histag protein by electrospray mass spectrometry." Anal Biochem **301**(1): 35-48.
Verma, R., R. S. Annan, et al. (1997). "Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase." Science **278**(5337): 455-60.
Weckwerth, W. and J. Selbig (2003). "Scoring and identifying organism-specific functional patterns and putative phosphorylation sites in protein sequences using mutual information." Biochem Biophys Res Commun **307**(3): 516-21.
Weinert, T. (1998). "DNA damage and checkpoint pathways: molecular anatomy and interactions with repair." Cell **94**(5): 555-8.
Weinert, T. A. and L. H. Hartwell (1988). "The RAD9 gene controls the cell cycle response to DNA damage in Saccharomyces cerevisiae." Science **241**(4863): 317-22.
Weinert, T. A. and L. H. Hartwell (1990). "Characterization of RAD9 of Saccharomyces cerevisiae and evidence that its function acts posttranslationally in cell cycle arrest after DNA damage." Mol Cell Biol **10**(12): 6554-64.
Welcker, M., J. Singer, et al. (2003). "Multisite phosphorylation by Cdk2 and GSK3 controls cyclin E degradation." Mol Cell **12**(2): 381-92.
Wilson, G. D. (2004). "Radiation and the cell cycle, revisited." Cancer Metastasis Rev **23**(3-4): 209-25.
Wright, P. E. and H. J. Dyson (1999). "Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm." J Mol Biol **293**(2): 321-31.
Wu, Q., S. Van Orden, et al. (1995). "Characterization of cytochrome c variants with high-resolution FTICR mass spectrometry: correlation of fragmentation and structure." Anal Chem **67**(14): 2498-509.
Yaffe, M. B. and A. E. Elia (2001). "Phosphoserine/threonine-binding domains." Curr Opin Cell Biol **13**(2): 131-8.
Yan, Z., G. W. Caldwell, et al. (1999). "Identification of a gluconic acid derivative attached to the N-terminus of histidine-tagged proteins expressed in bacteria." Biochem Biophys Res Commun **262**(3): 793-800.

Yan, Z., G. W. Caldwell, et al. (1999). "Mass spectrometric determination of a novel modification of the N-terminus of histidine-tagged proteins expressed in bacteria." Biochem Biophys Res Commun **259**(2): 271-82.
Yang, F. D., R. A. Spanevello, et al. (1990). "The carboxyl terminus heptapeptide of the R2 subunit of mammalian ribonucleotide reductase inhibits enzyme activity and can be used to purify the R1 subunit." FEBS Lett **272**(1-2): 61-4.
Yang, J., Z. P. Xu, et al. (2004). "ATM and ATR: sensing DNA damage." World J Gastroenterol **10**(2): 155-60.
Yao, R., Z. Zhang, et al. (2003). "Subcellular localization of yeast ribonucleotide reductase regulated by the DNA replication and damage checkpoint pathways." Proc Natl Acad Sci U S A **100**(11): 6628-33.
Zakian, V. A. (1995). "ATM-related genes: what do they tell us about functions of the human gene?" Cell **82**(5): 685-7.
Zeng, Y. and H. Piwnica-Worms (1999). "DNA damage and replication checkpoints in fission yeast require nuclear exclusion of the Cdc25 phosphatase via 14-3-3 binding." Mol Cell Biol **19**(11): 7410-9.
Zhang, X., C. J. Herring, et al. (1998). "Identification of phosphorylation sites in proteins separated by polyacrylamide gel electrophoresis." Anal Chem **70**(10): 2050-9.
Zhao, X., A. Chabes, et al. (2001). "The ribonucleotide reductase inhibitor Sml1 is a new target of the Mec1/Rad53 kinase cascade during growth and in response to DNA damage." Embo J **20**(13): 3544-53.
Zhao, X., B. Georgieva, et al. (2000). "Mutational and structural analyses of the ribonucleotide reductase inhibitor Sml1 define its Rnr1 interaction domain whose inactivation allows suppression of mec1 and rad53 lethality." Mol Cell Biol **20**(23): 9076-83.
Zhao, X., E. G. Muller, et al. (1998). "A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools." Mol Cell **2**(3): 329-40.
Zhao, X. and R. Rothstein (2002). "The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1." Proc Natl Acad Sci U S A **99**(6): 3746-51.
Zheng, P., D. S. Fay, et al. (1993). "SPK1 is an essential S-phase-specific gene of Saccharomyces cerevisiae that encodes a nuclear serine/threonine/tyrosine kinase." Mol Cell Biol **13**(9): 5829-42.
Zhou, B. B. and S. J. Elledge (2000). "The DNA damage response: putting checkpoints in perspective." Nature **408**(6811): 433-9.
Zhou, B. P. and M. C. Hung (2002). "Novel targets of Akt, p21(Cipl/WAF1), and MDM2." Semin Oncol **29**(3 Suppl 11): 62-70.
Zhou, Z. and S. J. Elledge (1993). "DUN1 encodes a protein kinase that controls the DNA damage response in yeast." Cell **75**(6): 1119-27.
Zolnierowicz, S. and M. Bollen (2000). "Protein phosphorylation and protein phosphatases. De Panne, Belgium, September 19-24, 1999." Embo J **19**(4): 483-8.

APPENDIX

Appendix

Experimental procedures used in this research are described in Chapter 2. However, calculation and data interpretation involved in some of these experiments are not straightforward. In this Chapter, methods of calculation and data interpretation performed for these experiments will be illustrated. Numbers actually obtained in these experiments will be taken as sample data, and calculation and data interpretation will be conducted on these data. Specifically, experiments discussed in this chapter are:

1. Determination of Rnr1 specific activity
2. P81 phospho-cellulose based assay of Sml1 phosphorylation by the Dun1 kinase
3. Identification of Sml1 phosphorylation sites based on mass spectrometric fragmentation (SORI-CAD) of phospho-peptide derivatives.
4. Fluorescence based assay to determine dissociation constant of the Sml1-Rnr1 interactions.

Determination of Rnr1 specific activity

Preparation of CDP stock

CDP used as the substrate of ribonucleotide reductase is mixture of non-ratio-labeled (cold) and C14 labeled (hot) compounds. The radioactivity (cpm) given by unit quantity (nmol in this case) of CDP is generally called specific activity.[32] The goal of this protocol

[32] In this case, specific activity refers to radioactivity per unit amount of the compound. It is different from specific activity of the enzyme.

is to prepare a desired concentration of CDP with a desired specific activity. In this protocol, specific activity is given by cpm/nmol. An important assumption is that radioactive and non-radioactive CDP are chemically identical. In addition, it should be noted that the molar quantity of non-radioactive CDP is much more (≥1000 x) than the radioactive form. Thus, we can assume that the total amount of CDP is almost the same as the total amount of cold CDP.

This protocol consists of four steps:

1. Deciding required CDP concentration and specific activity of the stock solution.
2. Preparing non-radioactive CDP solution and determining total concentration of cold CDP.
3. Preparing radioactive CDP.
4. Mixing radioactive and non-radioactive CDP and determining specific activity.

To make CDP stock, first it is important to decide how much total amount (concentration and volume) and specific activity of CDP stock will be required. Generally, higher specific activity correlates to the better sensitivity. However, due to cost, health and environmental concern, the specific activity should be limited to a reasonable level. In most cases, 1500 to 2000 cpm/nmol of specific activity is sufficient to measure wild type ribonucleotide reductase activity. However, when activity is expected to be lower due to the presence of inhibitors or mutation of RNR subunits, it is

reasonable to have 2500 to 4000 cpm/nmol of specific activity. The example illustrated here aimed specific activity of 2000 cpm/nmol in 75μl of 30mM CDP.

The second step is to make non-radioactive CDP solution. Normally, we weigh di-sodium salt of CDP and dissolve it in 20mM Hepes-KOH pH 7.0. Although pH does not usually change as 1 to 100mM of CDP solution is added to the buffer, it is better to check the pH. At pH other than ~7, a small amount of HCl or KOH should be added to adjust the pH. In this example, we *tried* to make 30mM CDP. However, it is important to actually check the concentration based on absorbance at 270nm and extinction coefficient of CDP at 270nm (ε_{270}=9.000$mM^{-1}cm^{-1}$), because concentration aimed by weight of CDP is often not accurate. Table 11 shows CDP concentrations determined based on three measurements of absorbance at 270nm.

The third step is to decide appropriate volume of hot and cold CDP solution needed to obtain the desired specific activity. On the supplier's product information or package of hot CDP, the concentration of radioactivity is indicated. In this example, when it is supplied, a total radioactivity of 10μCi is present in a volume of 200μl of hot CDP. Therefore the concentration of radioactivity is 0.05μCi/μl. Unit of radioactivity should be converted from μCi to cpm. The conversion is achieved based on the conversion factor:

$1\mu Ci = 2.22 \times 10^6$ dpm

Table 11. Determination of CDP concentration.

A270 of 1/1000 dilution	
Reading 1	0.2707
Reading 2	0.2813
Reading 3	0.2825
Average	0.2782
Concentration (mM)	**30.57**

A 1/1000 dilution of CDP stock was made in three independent aliquots. Absorbance at 270nm (A_{270}) for each aliquot was measured (reading 1, 2, and 3 on the table), and the average of the three readings was calculated (0.2782). The concentration (mM) of CDP, 30.57mM, was determined based on the average $A_{270,}$ dilution factor and the extinction coefficient (Average A_{270} x dilution factor / extinction coefficient =0.2782 x 1000 / 9.000 (mM^{-1})= 30.57mM).

Thus far, we have assumed that the dpm is approximately or slightly more than cpm[33]. Based on this assumption, 0.05μCi/μl is converted into 1.11 x 10^5 cpm/μl. We are aiming 2000cpm/nmol in 75μl. 75μl of 30.57mM cold CDP contains total 2293nmol CDP [(30.57 x 10^{-3}) mol/L x (75 x 10^{-6}) L]. To have 2000cpm/nmol, total 4585500cpm of radioactivity is required (2000cpm/nmol x 2293nmol = 4585500cpm). 41.3μl of the hot CDP is required to have total 4585500cpm (4585500cpm / 1.11 x 10^5 cpm /μl=41.3μl). Therefore, in this case, 41μl of hot CDP was taken into an Ependorf tube. Since hot CDP is supplied in 50% ethanol, it was completely dried in a lypholizer. Then, 75μl of the cold CDP was added to the tube and mixed thoroughly.

Due to high error of the radioactivity concentration originally given by the supplier, it is important to actually check radioactivity concentration in the newly made CDP stock. In this example, 1/100 dilution of the stock was made first. Then, 4, 6, and 10μl of the diluted stock was mixed with scintillation cocktail, and their radioactivity was measured. Table 12 shows the radioactivity of the diluted stock. The volume of the diluted CDP stock (μl) was plotted with radioactivity (cpm) as shown in Figure 45.

A liner regression of these data point gave a curve, Y = 612.28X – 305.79 (Y: radioactivity (cpm). X: volume of the diluted CDP stock (μl)). The slope of the curve

---

[33] This assumption is not exactly correct. Strictly, dpm should be converted to cpm based on:

dpm = (cpm indicated – cpm background)/detector efficiency

We have not determined detector efficiency. Background cpm is less than 100, and it is negligible in our case. However, it is not practical to convert dpm to cpm in a strict manner, because error of radioactivity concentration given by the supplier is relatively high. When CDP stocks were made from different batch of hot CDP, we obtained both higher and lower specific activity than expected value.

Table 12. Radioactivity of CDP stock diluted 1/100 times.

Volume (µl)	cpm
4	2197.37
6	3286.79
10	5844.00

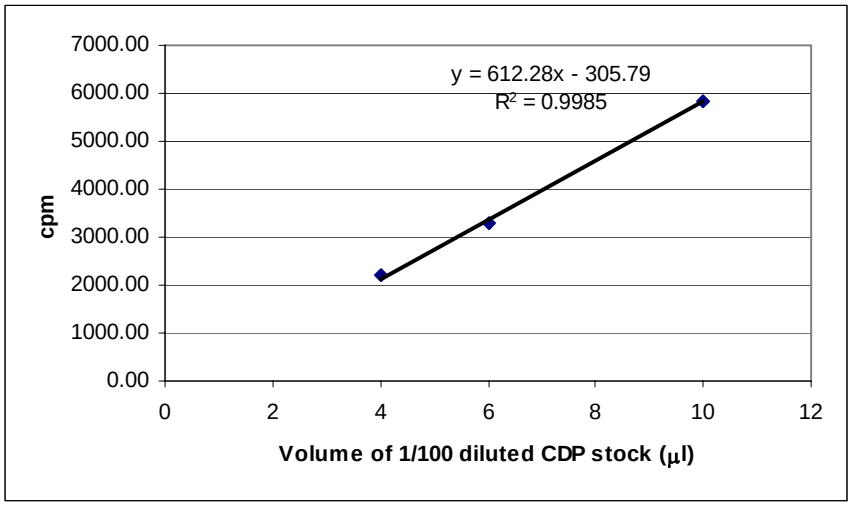

Figure 45. Plot of cpm vs. volume of CDP stock.

(612.28) is radioactivity (cpm) given by 1/100 μl of the CDP stock. Therefore, in this example, 61228cpm is given by 1μl of the CDP stock (i.e. 61228 cpm/μl). On the other hand, 1μl of the CDP stock contains 30.57nmol of CDP (30.57 x 10^{-3} mol/L x 10^{-6}L x 10^{9} nmol/mol=30.57nmol, assuming that total amount of CDP = total amount of cold CDP[34]). From the total amount of radioactivity (61228cpm/μl) and total amount of CDP (30.57nmol/μl) per μl of the stock, the specific activity of the CDP stock is given as 2003cpm/nmol [(61228 cpm/μl / (30.57nmol/μl) = 2003cpm/nmol].

RNR activity assay

The details of the assay procedure are described in Chapter 2. In the following example, the concentration of Rnr1 and H_6Rnr2/Rnr4 was 0.2mg/ml (1μM dimer) and 0.4mg/ml (5μM heterodimer) respectively. CDP stock from the previously described section was used (specific activity of 2003cpm/nmol). During the reaction, 50μl of aliquot was taken every four minutes (0, 4, 8 and 12 minutes in time course of the reaction), and was immediately boiled to quench the reaction. After sedimentation of precipitated materials by centrifugation, 950μl of solution containing alkaline phosphatase and deoxycytidine was added to each aliquot. The aliquots were incubated on 30°C water bath for 3 hours. From each aliquot, 950μl was applied to a borate column, and the column was washed with 8ml of water. All flow-through from the column (with total volume of 8.95ml) was collected in a tube. After mixing thoroughly, 3ml of the flow-through was taken and

[34] We tested if this assumption is correct, and there was no significant contribution of hot CDP to total amount of CDP in this procedure. However, it is also possible to determine exact concentration of CDP after mixing cold and hot CDP instead of determining it before mixing them together.

mixed with >15ml of scintillation cocktail. Radioactivity (cpm) of each aliquot was given as shown in Table 13.

Based on radioactivity and the specific activity, amount of dCDP in each 3ml of the flow-through was calculated [Amount of dCDP (nmol) = radioactivity (cpm) / specific activity of CDP stock (2003 cpm/nmol in this example)] (shown in the third column of Table 13). The amount of dCDP in the 50µl aliquot taken during quenching was calculated by multiplying these values (values in the third column of Table 13) by dilution factor [(8.95/3) x (1/0.95) = 3.14].

The rate of the reaction (nmol CDP produced per minutes) was obtained by plotting the quenching time (first column of Table 13) and nmol of dCDP produced (fourth column of Table 13). Figure 46 represents the x-axis as the quenching time and y-axis as nmol dCDP produced. By linear regression of data points on the plot, a curve, Y = 1.0639X + 0.8071, was obtained. The slope of the curve represents the rate of the reaction, which is 1.064 nmol/min. The rate of the reaction in this case represents the amount of dCDP produced in 50µl of the reaction mixture in every minute. The rate of reaction is defined as "*activity*".

Activity can vary depending on the amount of enzyme in the reaction mixture. To normalize the activity, it needs to be expressed in terms of unit amount of the enzyme. The normalized activity is called "*specific activity*". By convention, amount of enzyme needed to express specific activity is given by mg unit of enzyme, rather than mol of

Table 13. Radioactivity of reaction mixture quenched over the time course

Quenching time (min)	cpm	nmol dCDP in 3ml used for counting	nmol dCDP in 50µl aliquot
0	493.80	0.2465	0.774
4	3320.33	1.6577	5.206
8	5824.00	2.9076	9.131
12	8706.96	4.3469	13.651

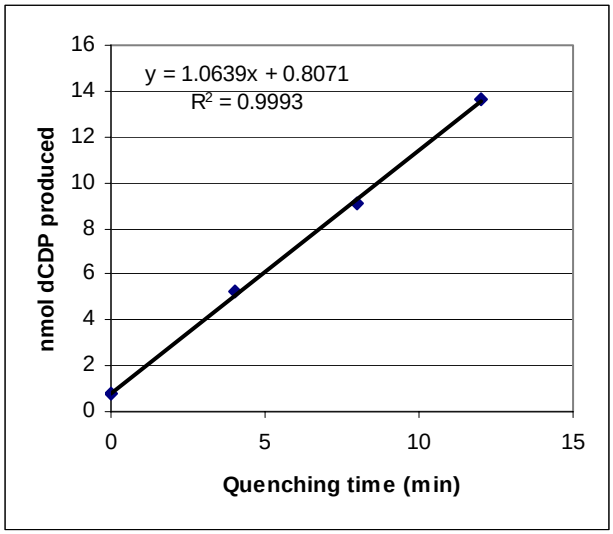

Figure 46. Plot of quenching time vs. nmol dCDP produced in 50µl of the reaction mixture.

enzyme. In this example, total 0.01mg of Rnr1[35] is present in 0.05ml of aliquot taken to quench reaction (Rnr1 concentration x volume of aliquots taken for quenching = 0.2mg/ml x 0.05ml = 0.01mg). By dividing activity by the amount of Rnr1, specific activity of 106.4nmol/min/mg was obtained [specific activity = activity / mg Rnr1 = 1.064 (nmol/min) / 0.01mg = 106.4 nmol/min/mg].

In order to obtain better statistics, the experiment is normally replicated in duplicate or triplicate. In replicated experiment, the same calculation was performed for *each* dataset, and average of specific activity was calculated[36].

P81 phospho-cellulose based assay of Sml1 phosphorylation

"Activity" of the Dun1 kinase in our current experimental scheme is the amount of phosphate incorporated into Sml1 over 30min or 2 hours. Currently, we are simply comparing relative degree of phosphate incorporation between different forms of Sml1 or in the presence and absence of Rnr1.

The following is an example of kinase assay performed by the method described in Chapter 2. In this example, phosphate incorporation was conducted over 30 minutes. Originally, we had 212μM of Sml1, 10mM of ATP (cold ATP) and 7.86μCi/μl of

---

[35] RNR consists of Rnr1 homodimer and Rnr2/Rnr4 heterodimer. As described in Chapter 2, the components in which the activity is *not* measured should have excess molar ratio. In this case, we are measuring activity of Rnr1, and H_6Rnr2/Rnr4 has 5 molar excess of Rnr1 homodimer. So, Rnr1 is the limiting factor in the reaction. (i.e. Rnr2/Rnr4 is saturating Rnr1, and rate of the reaction must be proportional to amount of Rnr1 present in the reaction mixture, but small error in Rnr2/Rnr4 should not affect the activity.)

[36] It is not appropriate to calculate average radioactivity at each quenching time when the reaction is replicated. The 0 min of the quenching time is not necessary to be the same for different replicates. We are concerned with rate of the reaction (i.e. slope of the curve) rather than overall radioactivity at each quenching point.

$\gamma[^{32}P]$ATP (hot ATP) as stock solutions. The concentration of Sml1 stock was determined using Coomassie Plus Assay Kit (PIERCE). The concentration of cold ATP stock was determined based on absorbance at 259nm and the extinction coefficient of ATP at 259nm, $\varepsilon_{259}=14500M^{-1}cm^{-1}$. The concentration of hot ATP stock is estimated based on the label of the original stock (250μCi in 25μl = 10μCi/μl) and the radioactivity decay estimated based on number of dates after it is produced. From these stock solutions, 12μM Sml1, 370μM cold ATP and 1μCi/ul of $\gamma[^{32}P]$ATP was prepared.

Next, specific activity of ATP in this mixture was determined. To determine the concentration of radioactivity (cpm) in this mixture, 1/100 dilution of the mixture was made. 2, 4, and 8μl of 1/100 dilution was taken and mixed with the scintillation cocktail. Table 14 shows their radioactivity determined by the scintillation counter. The volume of the 1/100 dilution was plotted with radioactivity (cpm) as shown in Figure 47. A liner regression of the data points on the plot gave a curve, Y=1660.3X-545.46, (X: volume (μl) of the diluted mixture. Y: radioactivity (cpm)). The slope of the curve (1660.3) is radioactivity (cpm) given by 1/100 μl of the undiluted mixture. Therefore, in this example, 1μl of the mixture gives 166030cpm (i.e. 166030cpm/μl). Assuming that this mixture contains mostly cold ATP, 1μl of the mixture contains 370pmol ATP (370×10^{-6} mol/L $\times 10^{-6}$L $\times 10^{12}$ pmol/mol=370pmol). Therefore, the specific activity of ATP in this mixture is 448.7cpm/pmol [(166030cpm/μl) / (370pmol/μl)].

The assay was performed as follows. 30μl of Sml1 /ATP / $\gamma[^{32}P]$ATP mixture was mixed with 15μl bed volume of GST-Dun1 bound glutathione resin, and it is incubated on 30˚C water bath for 30min. To take background radioactivity into consideration,

Table 14.Volume of Sml1 /ATP / γ[^{32}P]ATP mixture diluted 1/100 times and their corresponding radioactivity.

µl mixture	cpm
2	2849.33
4	5984.45
8	12774.04

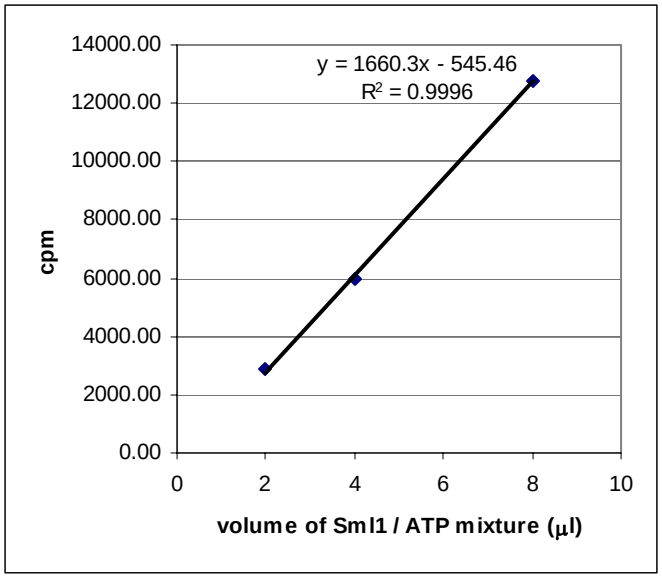

Figure 47. Plot of cpm vs. volume of reaction mixture

negative control was performed. 30μl of Sml1 /ATP / γ[^{32}P]ATP mixture was incubated with glutathione resin without GST-Dun1[37].

Note that when mixed with the resins, Sml1 and ATP are diluted with the buffer contained in the glutathione resin. Throughout this thesis, the concentrations of Sm1 and ATP (and Rnr1 in Chapter 6) are stated based on the dilution factor determined as described below. According to the manufacturer (Amersham Bioscience), approximately 96% bed volume of the resin is occupied with liquid by which the resin is equilibrated. Therefore, 15μl bed volume of the resin contains approximately 14.4μl (15μl x 0.96=14.4μl) of the buffer. When 30μl of the solution containing 12μM Sml1 and 370μM ATP was mixed with 15μl be volume of the resin, the concentration of Sml1 and ATP should be reduced to 8μl and 250μM respectively.

The reaction was stopped[38] by addition of 10μl of 400mM cold ATP and placing the reaction mixture on ice. The reaction vessel was briefly centrifuged, and 30μl of the supernatant was placed on 3cm^3 of P81 phospho-cellulose filter. Supposedly, 198.5pmol of Sml1 was applied to the filter (Sml1 concentration in the sample applied to the filter was 8μM x 44.4μl/(10μl+44.4μl)=6.6μM. 30μl of this solution contains 6.6 x 10^{-6} mol/L x 30 x 10^{-6} mol/L x 10^{12} pmol/mol = 198.5pmol of Sml1). Table 15 shows radioactivity

---

[37] Ideally, it is better to use GST- D328A Dun1 (kinase dead mutant) as a negative control. Nevertheless, as shown in Chapter 5, we have performed negative control experiments with glutathione resin pre-incubated with lysate of yeast cells that is not transformed with GST-DUN1 plasmid. At the same time, another negative control experiment was performed with glutathione resin that is simply equilibrated with reaction mixture. No significant difference was observed between these two. Since then, we have been using glutathione resin simply equilibrated with reaction mixture as negative control.

[38] In this method, phosphorylation still takes place after addition of cold ATP. However, incorporation of ^{32}P to Sml1 is significantly slowed down. Over 30minutes on ice after addition of 500mM ATP, increase of ^{32}P incorporation to Sml1 was indistinguishable from fluctuation of background radioactivity, which is 100-500cpm.

Table 15. Radioactivity of P32 that remains on P81 filters with Sml1.

	cpm	Δ cpm	pmol PO4	mol PO4 / mol Sml1
Sample 1	28780.0	28431.5	63.36	0.32
Sample 2	25192.5	24844.0	55.37	0.28
Sample 3	23466.7	23118.2	51.52	0.26
Sample 4	22466.7	22118.2	49.29	0.25
Negative control 1	347.8			
Negative control 2	349.2			
Negative control average	348.5			

(cpm) counted from the filter after it was washed. Assumption here is that most of Sml1 applied to the filter remains on the filter after it was washed.

To determine the radioactivity associated with Sml1 by phosphorylation, background radioactivity needs to be subtracted. The average cpm of the negative control was calculated as shown in bottom row of the second column in Table 15. The result was then subtracted from the cpm of each sample (shown in the third column of Table 15). This value, termed Δcpm, represents the radioactivity covalently attached to Sml1. By dividing Δcpm by the specific activity of ATP determined as above (448.7cpm/pmol), Δcpm is converted into pmol of phosphate attached to Sml1 (the fourth column of Table 15). By further dividing this value by total moles of Sml1 on the filter (198.5pmol), the molar ratio of phosphate and Sml1 was determined (the fifth column of Table 15).

Identification of Sml1 phosphorylation sites based on mass spectrometric fragmentation (SORI-CAD) of phospho-peptide derivatives.

Based on mass spectrometric analysis of peptides generated by CNBr digestion of phosphorylated Sml1, we found that a CNBr peptide spanning residues 52 to 64 (GSSASASASSLEM) contains three phosphorylation sites (discussed in Chapter 5). By SORI-CAD, we identified phosphorylation sites of doubly and triply phosphorylaed species. In this section, identification of the phosphorylation sites on triply phosphorylated species will be illustrated based on its SORI-CAD data. Although it will be not shown, identification of phosphorylation sites on doubly phosphorylated species was achieved in the same manner.

It is important to know what kind of modification was made on the Sml1 phospho-peptide spanning residues 52 to 64 during the experiment. Identification of Sml1 phosphorylation sites were achieved by following steps:

1. CNBr digestion of phosphorylated Sml1.
2. Ga(III)IMAC enrichment of Sml1 phospho-peptide.
3. Conversion of phospho-serine on Sml1 phospho-peptide into S-ethycysteine.
4. SORI-CAD of Sml1 peptide in which phospho-serine is converted into S-ethylcysteine.

Sml1 phospho-peptide was chemically modified in step 1 and 3 of the procedure above. First, after CNBr digestion of Sml1, C-terminus methionine of the peptide was converted into homolactone. We have observed that C-terminus homolactone of the CNBr peptide is often converted further into homoserine by intramolecular hydrolysis. C-terminus homolactone is 30.08736Da (in monoisotopic mass) less than C-terminus methionine. The CNBr peptide used for SORI-CAD had homolactone at its C-terminus. Second, in the third step above, phospho-serine was converted into S-ethylcysteine. Mass of S-ethylcysteine is 44.00846Da (in monoisotopic mass) more than seine. In this example, the peptide of interest was originally phosphorylated at three serine residues.

Based on the knowledge of the chemical modification described above as well as previous data showing the triply phosphorylated CNBr peptide, we expected to find a peptide spanning residues 52 to 64 of Sml1 (GSSASASASSLEM) with:

1. Loss of 30.08736Da at C terminus due to CNBr digestion.

2. Gain of 44.00846Da at three serine residues due to conversion of serine into S-ethylcysteine. (Note: total gain is 44.00846Da x 3 = 132.2538Da.)

The mass of such peptide is 1285.441Da. An ion of 1285.528Da (1286.536^{+1}) was found. This was one of a few ions found in this sample, and no other ion close to 1286.436Da was observed. Assuming this ion is the peptide of interest, it was isolated in the mass spectrometer and used as parent ion in SORI-CAD.

The peptide of interest had six serine residues (S53, S54, S56, S58, S60 and S61), three of them converted into S-ethylcysteine. Our objective is to find which three of the six serine residues were converted into S-ethylcysteine. In the SORI-CAD of this peptide (the ion with mass/charge of 1286.536), over 20 peaks were observed. Mass/charge of these fragment ions were compared with calculated mass/charge based on possible locations of S-ethylcysteine (possible choice of three of the six serine residues) on the peptide and fragmentation pattern generally observed in CAD of peptides.

In general, CAD fragmentation of peptides most frequently takes place at the amide bond of the peptide backbone. If a charge is retained on N-terminus side from the amide bond where the fragmentation takes place, the fragment ion is classified as b-type ion. If the charge is retained on C-terminus side from the amide bond where the fragmentation takes place, the fragment ion is classified as y-type ion. b-type or y-type ions are most likely to be observed in CAD of peptides. Figure 48 illustrate b-type and y-type ion possibly produced from the peptide spanning residues 53 to 64 of Sml1.

Given the peptide sequence (GSSASASASSLEM), mass/charge of a singly charged fragment ion (M + H) was calculated by a web-based server, PROWL

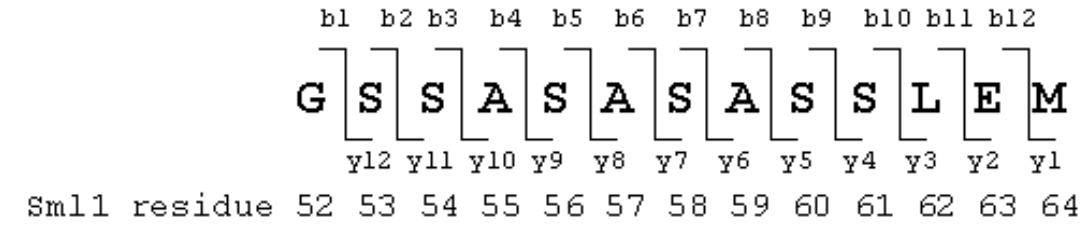

Figure 48. b-type and y-type ions that can be generated from the peptide spanning residues 52 to 64 of Sml1.

(http://prowl.rockefeller.edu/prowl-cgi/sequence.exe). PROWL also allows modifications on particular residues on the peptide in terms of gain or loss of mass associated with the modifications. In this case, the loss of 30.08736Da at the C-terminus and gain of 44.00846Da on three serine residues were included in the calculation. In addition, neutral loss of NH_3, H_2O and CO occurs frequently on b or y type ions resulting loss of 17,02655, 18.0156 or 27.99491Da respectively, and this was taken into consideration.

The observed mass/charge was compared with calculated mass/charge for the given peptide. It should be noted that there are 20 possible ways to replace the six serine residues with S-ethylcysteine. Therefore, there are 20 possibilities to add 44.00846Da on three serine residues on the peptide. Out of the 20 possibilities, the most reasonable match between observed and calculated mass/charge was found when Ser56, Ser58 and Ser60 were replaced with S-ethylcysteine. The following discussion explains why Ser56 Ser58 and Ser60 are the sites of phosphorylation (S-ethylcysteine) and why other possibilities can be excluded.

Table 16 is the list of calculated mass/charge for b-type ions along with observed mass/charge that match the calculated value. In this case, residue 56, 58 and 60 are considered to be S-ehtylcysteine. Some observed mass/charge matched calculated mass/charge of b type ion which have undergone neutral loss of one or two H_2O (loss of 18.0156 or 36.312), and they are denoted as b* or b** ions. Although the data on Table 16 covers only a part of the peptide, it does indicate that S61 is not a phosphorylation site and S60 is a phosphorylation site. Some of b-type ions are unique to the location of S-ethylcysteine, and existence of such ions excludes some possibilities.

Table 16. List of calculated and observed mass/charge of b-type ions generated from the peptide spanning residues 52 to 64 of Sml1.

		Calculated	Observed	Calculated	Observed	Calculated	Observed
Residue	Ion number	b		b*		b**	
G53	1	58.032		40.016		22.001	
S53	2	145.064		127.048		109.033	
S54	3	232.096		214.080		196.065	
A55	4	303.133		285.118		267.102	
S56 +Mod	5	434.174		416.158		398.142	
A57	6	505.211		487.195		469.179	
S58 +Mod	7	636.251		618.235		600.220	
A59	8	707.288		**689.273**	689.280	**671.257**	671.268
S60 +Mod	9	**838.329**	838.332	**820.313**	820.320	802.297	
S61	10	925.361	925.368	907.345	907.355	889.329	
L62	11	1038.445	1038.451	1020.429	1020.440	1002.413	
E63	12	1167.487	1167.497	1149.472	1149.486	1131.456	1131.474
M64 HomoSer	13						

Residue 64 is homoserine (denoted as HomoSer) and residues 56, 58 and 60 are S-ethylcysteine(Denoted as +Mod). b* and b** ions are b type ion that lost one or two H_2O respectively.

Observation of b9 and b*9 (red bold letters on Table 16) shows that S61 is not a phosphorylation site. If S61 was a phosphorylation site, the mass/charge of b9 and b*9 would be 794.320 and 776.305 respectively (44.00846Da less than the values on Table 16). These values did not match any of observed ions. In addition, if S61 was a phosphorylation site, b1 to b9 ions and their neutral loss derivatives have mass/charge 44.00846Da less than values on Table 16. No such ions were observed.

Observation of b*8 and b**8 (blue bold letters on Table 16) shows that S60 is a phosphorylation site. If Ser60 is not a phosphorylation site (but three residues out of S53, S54, S56 and S58 are the sites), b*8 and b**8 would have a mass/charge of 733.281 and 715.265 respectively (44.00846Da more than the values on Table 16). These values did not match any of the observed ions. In addition, if S60 was not a phosphorylation site, b1 to b8 ions and their neutral loss derivatives would have a mass/charge 44.00846Da higher than the values on Table 16. No such ions were observed

Y-type ions provided the information that S53 and S54 are not phosphorylation sites but S56 and S58 are phosphorylation sites. Table 17 is the list of calculated mass/charge for y-type ions along with observed mass/charge that match the calculated values. Like Table 16, residue 56, 58 and 60 are considered to be S-ehtylcysteine. Two observed mass/charge matched calculated values of y type ion which have undergone neutral loss of one H_2O (loss of 18.0156), and it is denoted as y* ions. Only 6 y-type ions covering four residues were observed. However, observations of these y-type ions provided important clues.

Table 17. List of calculated and observed mass/charge of y-type ions generated from the peptide spanning residues 52 to 64 of Sml1.

		Calculated	Observed	Calculated	Observed
Residue	Ion number	y		y*	
G52	13				
S53	12	1229.4187		1211.4031	
S54	11	1142.3867		1124.3711	
A55	10	1055.3547		1037.3391	
S56 +Mod	9	**984.3177**	984.429	966.3021	
A57	8	**853.2772**	853.386	835.2616	
S58 +Mod	7	**782.2401**	782.348	**764.2245**	764.339
A59	6	**651.1997**	651.306	**633.1841**	633.296
S60 +Mod	5	580.1626		562.147	
S61	4	449.1222		431.1066	
L62	3	362.0901		344.0745	
E63	2	249.0062		230.9906	
M64 HomoSer	1	119.9636		101.948	

Observation of y9 (red bold letters on Table 17) excludes the possibility that S53 or S54 is a phosphorylation site. If either S53 or S54 was a phosphorylation site, the mass/charge of y9 ion would be 940.309 (44.00846Da less than the value on Table 17). This value did not match any of observed ions. Similarly, if both S53 and S54 are phosphorylation sites, mass/charge of y9 ion would be 896.301. This value did not matched the observed values.

Observation of y8, y7 and y*7 (blue bold letters on Table 17) shows that S56 is a phosphorylation site. If S56 is not a phosphorylation site (but S56, S58 S61 are the sites), mass/charge of y8, y7 and y*7 would be 897.286, 826.249 and 808.233 respectively (44.00846Da more than the values on Table 17). These values did not match any of observed ions. In addition, if S56 is not a phosphorylation site, y1 to y8 ions and their neutral loss derivatives have mass/charge 44.00846Da more than the values on Table 17. None of such ions were observed.

Observation of y6 and y*6 (green bold letters on Table 17) shows that S58 is a phosphorylation site. If S58 was not a phosphorylation site the mass/charge of y6 and y*6 would be 695.208 and 677.193 respectively (44.00846Da more than the values on Table 17). These values did not match any of observed ions. In addition, if S58 is not a phosphorylation site, y1 to y6 ions and their neutral loss derivatives have mass/charge 44.00846Da more than values on Table 17. None of such ions were observed.

Although the data above show that S56, S58 and S60 are phosphorylation site, we cannot completely exclude minor possibility that S53, S54 or S61 is a phosphorylation site. In mass spectrometry, due to a weak signal or suppression of the signal by more

Residue 64 is homoserine (denoted as HomoSer) and residues 56, 58 and 60 are S-ethylcysteine(Denoted as +Mod). y* is y-type ion that lost one H_2O.

abundant ions, the less abundant ions are sometimes not detected. Therefore, even though ions corresponding to phosphorylation of these sites are not observed, that does not mean that these sites are not phosphorylated. If any of b1 to b4 ions on Table 16 or their neutral loss derivatives were observed, we could exclude the possibility that S53 or S54 is a phosphorylation site. Similarly, if any of y1 to y4 ions on Table 17 or their neutral loss derivatives were observed, we could exclude the possibility that S61 is a phosphorylation site. However, we did not make such observation.

In addition, we observed two ions that suggest a possibility of S53 or S54 to be a phosphorylation site. Table 18 shows mass/charge of y7 and y8 ions that can be generated from the peptide in which two of residues 53, 54 or 56 are S-ethylcysteine. These calculated mass/charge values matched two observed ions. This observation leaves the possibility that S54 and S53 are minor phosphorylation sites. On the other hand, in the analysis of doubly phosphorylated peptide, we did not observe any ions that indicate phosphorylation at S54 or S53. Furthermore, the Dun1 kinase assay based on P32 incorporation into S56A / S58A / S60A showed complete loss of phosphorylation. Therefore, at most, S54 and S53 are possibly a minor phosphorylation sites.

Table 18. Mass/charge of y7 and y8 ions generated from the peptide in which two of residues 53, 54 or 56 are S-ethylcysteine.

	Calculated	Observed
y8	809.26874	809.3786
y7	738.23164	738.3418

Fluorescence based assay to determine dissociation constant of the Sml1-Rnr1 interactions.

As described in Chapter 2, the binding constant of the Sml1-Rnr1 interactions was determined based on an increase of fluorescence intensity from IANBD-S60C Sml1[39] when it was titrated with Rnr1. In this section, the detailed method and data analysis procedure of a recent trial will be described. The problem associated with our current experimental scheme and possible solutions will be also discussed.

IANBD-S60C Sml1 and Rnr1 stored in –80ºC were used for this experiment. β-mercaptoethanol conjugated with IANBD amide (IANBD-BME) stored in –20ºC was used as the negative control. The first step was to determine the concentration of protein and IANBD-BME. Protein concentrations are measured based on Coomassie Plus Assay Kit (PIERCE) using BSA standard supplied with the kit. IANBD-BME concentration was measured based on absorbance at 478nm of IANBD-BME diluted in methanol and extinction coefficient of IANBD at 478nm ($25{,}000 cm^{-1}M^{-1}$). By diluting these samples in 50mM Hepes-KOH pH7.0 100mM KCl, 5% Glycerol, 5mM $MgCl_2$, 5mM DTT, four types of solutions[40] were prepared:

1. 5μM IANBD-S60C Sml1 monomer.
2. 5μM IANBD-S60C Sml1 monomer, 30μM Rnr1 monomer.
3. 5μM IANBD-BME.

---

[39] IANBD-S60C Sml1 is C14S / S60C Sml1 mutant conjugated with IANBD amide at residue 60. Fluorescence intensity of IANBD amide increases when it moves to a less solvent accessible environment. (See Chapter 2 "Fluorescence based assay of the Sml1-Rnr1 interaction".)

[40] 250μM ATP was included in every solution for the data shown in Chapter 6 to be consistent with the kinase assay.

4. 5μM IANBD-BME monomer, 30μM Rnr1 monomer.

100μl of 5μM IANBD-S60C Sml1 monomer (solution 1) was placed in a quartz cuvette (3x3x20mm), and fluorescent intensity was measured for one second at excitation wavelength of 478nm and emission wavelength of 541nm using the fluorometer (LS50B: Perkin Elmer, Wellesley, MA). The slit width was adjusted to obtain emission intensity of 150 to 200 in arbitrary unit. Then, the sample was titrated with the Sml1/Rnr1 mixture (solution 2) at 3μl increments using pipette. After each addition of the Sml1/Rnr1 mixture, the solution in the cuvette was mixed by aspirating and dispensing the solution by pipette. As a negative control, the same procedure was performed with 5μM IANBD-BME (solution 3) and a mixture of 5μM IANBD-BME and 30μM Rnr1 monomer (solution 4). The fluorescence intensity after each titration was manually recorded. These titration procedures were repeated three times.

Table 19 shows fluorescence intensities for these series of titration. Concentration of Rnr1 after each titration was calculated based on the dilution factor obtained from volumes and Rnr1 concentrations of the solution in the cuvette and solution added to the cuvette. Concentration of IANBD-S60C Sml1 and IANBD-BME were assumed to be 5μM throughout the procedure. Average intensities of the negative control series was calculated first (Shown in the 8th column on Table 19 labeled as "NC average").

Then, average intensities of negative control was subtracted from the series of intensity obtained by titration of IANBD-S60C Sml1 with Rnr1 (i.e. The values in 8th column of Table 19 were subtracted from the values in the 2nd, 3rd and 4th columns on Table 19.) These values are now called corrected fluorescence intensity or Delta F

Table 19. Fluorescence emission intensity in titration of IANBD-S60C Sml1 and IANBD-BME with Rnr1.

	Experimental series			Negative control series			
[Rnr1] (μM)	Run1	Run2	Run3	Run1	Run2	Run3	NC average
0.00	183.40	179.00	182.50	180.40	184.60	183.30	182.77
0.87	247.30	244.00	245.90	177.90	184.40	180.30	180.87
1.70	328.80	297.60	296.60	178.20	183.40	180.80	180.80
2.48	368.90	333.10	338.50	179.00	185.00	180.50	181.50
3.21	398.20	380.50	371.80	178.70	186.30	181.70	182.23
3.91	427.90	407.70	392.70	179.30	186.80	181.20	182.43
4.58	450.70	420.90	413.50	178.80	185.20	182.10	182.03
5.21	462.70	433.60	427.90	179.60	186.10	181.60	182.43
5.81	474.70	441.40	443.10	179.80	185.90	183.20	182.97
6.38	483.50	445.30	447.80	179.30	187.30	181.40	182.67
6.92	492.10	456.20	459.60	181.40	184.50	182.60	182.83
7.44	500.90	461.40	460.40	181.60	185.60	182.40	183.20
7.94	502.50	459.90	466.40	180.90	186.10	183.90	183.63
8.42	509.50	469.30	468.50	179.60	185.20	183.00	182.60
8.87	511.90	472.20	471.70	178.80	184.40	183.80	182.33
9.31	517.20	477.30	472.10	180.60	185.90	182.50	183.00
9.73	520.60	483.70	477.30	179.10	184.90	183.40	182.47
10.13	520.60	483.80	480.70	178.80	186.30	184.10	183.07
10.52	520.70	483.80	481.70	179.30	185.90	183.90	183.03

(Shown in Table 20). Three series of delta F as dependent variables (the 2nd, 3rd and 4th column on Table 20) and corresponding Rnr1 concentrations (the first column on Table 20) as dependent valuables were inputted to GraphPad Prism software (GraphPad, San Diego, CA). By nonlinear regression, the series of data were analyzed by fitting them to equation below (Derivation of this equation is described in the next section):

$$DeltaF = DeltaF \max \frac{[Sml1] + [Rnr1] + Kd - \sqrt{([Sml1] + [Rnr1] + Kd)^2 - 4[Sml1][Rnr1]}}{2[Sml1]}$$

(Equation 1)

(Delta F_{max} : Delta F when Sml1 is saturated with Rnr1. K_d: dissociation constant of the Sml1-Rnr1 interactions. [Sml1]: total concentration of IANBD-S60C Sml1. [Rnr1]: total concentration of Rnr1).

The analysis was performed in two slightly different methods. In the first method, in Equation 1, K_d and Delta F_{max} were considered to be unknown valuables (i.e. they were not fixed). In this case, concentration of Sml1 monomer was specified to be 5μM[41] (i.e. Sml1 concentration was fixed at 5μM). 300 and 0.4μM[42] were chosen as initial values of Delta F_{max} and K_d respectively. In the second method, in the equation above, K_d, Delta

[41] In this case, Equation 1 was written as Y=Bmax * (1/5) *0.5*((5+X+Kd)-((5+X+Kd)^2-4*5*X)^0.5 on GraphPad software. In this equation X is Rnr1 concentration as independent valuable, Y is Delta F as dependent valuable and Bmax is Delta F when Sml1 is saturated with Rnr1.
[42] Based on Chabes et al., 1999

Table 20. Delta F during titration of IANBD-S60C Sml1 with Rnr1

[Rnr1] (μM)	Delta F-1	Delta F -2	Delta F -3
0.00	0.63	-3.77	-0.27
0.87	66.43	63.13	65.03
1.70	148.00	116.80	115.80
2.48	187.40	151.60	157.00
3.21	215.97	198.27	189.57
3.91	245.47	225.27	210.27
4.58	268.67	238.87	231.47
5.21	280.27	251.17	245.47
5.81	291.73	258.43	260.13
6.38	300.83	262.63	265.13
6.92	309.27	273.37	276.77
7.44	317.70	278.20	277.20
7.94	318.87	276.27	282.77
8.42	326.90	286.70	285.90
8.87	329.57	289.87	289.37
9.31	334.20	294.30	289.10
9.73	338.13	301.23	294.83
10.13	337.53	300.73	297.63
10.52	337.67	300.77	298.67

F_{max} and [Sml1] were considered to be unknown valuables[43]. 300, 0.4μM and 5μM were chosen as initial values of Delta F_{max}, K_d and [Sml1] respectively. The results of these analyses are shown in Table 21 and Figure 49.

Two method of analysis gave a different result regarding how well the model fit to the data points. In the first method, when Sml1 monomer concentration is specified to be 5μM, a systematic deviation of data point from the model was observed (see Figure 49B). In the second method, when Sml1 concentration is not specified, deviation of the data point from the model was randomized (see Figure 49D) as P value[44] in the first method was 0.67 and that for the second method was 0.99 (see Runs Test on Table 21). In addition, the overall deviation of data point from the model was reduced in the second method (see Figure 49B and Goodness of fit on Table 21). Therefore, model given by the second method was likely to be closer to what was actually happening. K_d±Std error given by the model in the second method (1.07±0.47μM) was chosen.

However, the second method still had a problem associated with concentration of IANBD-S60C Sml1. IANBD-S60C Sml1 monomer concentration for the best fitting equation in the second method was 3.25±0.65μM, while 5μM was aimed in the experiment. A possible explanation of this discrepancy may be that actual Sml1 concentration in the cuvette was less than the concentration we aimed. Especially in a sample with lower concentration of protein, protein concentration is significantly

---

[43] In this case, Equation 1 was written as Y=Bmax * (1/S) *0.5*((S+X+Kd)-((S+X+Kd)^2-4*S*X)^0.5 on GraphPad software. In this equation S is unspecified Sml1 concentration.

[44] P-value is probability that systematic deviation of the data point from the model occurs by random chance. The greater P-value it is less likely that there is a systematic deviation of the data point from the model.

Table 21. Result of data fitting

	[Sml1]=5μM	[Sml1] not specified
Best-fit values		
BMAX	326.1	356.4
KD	0.2828	1.072
S		3.248
Std. Error		
BMAX	10.51	19.77
KD	0.1139	0.4793
S		0.6528
95% Confidence Intervals		
BMAX	305.0 to 347.1	316.8 to 396.1
KD	0.05450 to 0.5112	0.1101 to 2.033
S		1.938 to 4.558
Goodness of Fit		
Degrees of Freedom	55	54
R^2	0.9558	0.9669
Absolute Sum of Squares	20060	15010
Sy.x	19.1	16.67
Runs test		
Points above curve	28	18
Points below curve	29	39
Number of runs	30	36
P value (runs test)	0.607	0.9999
Deviation from Model	Not Significant	Not Significant

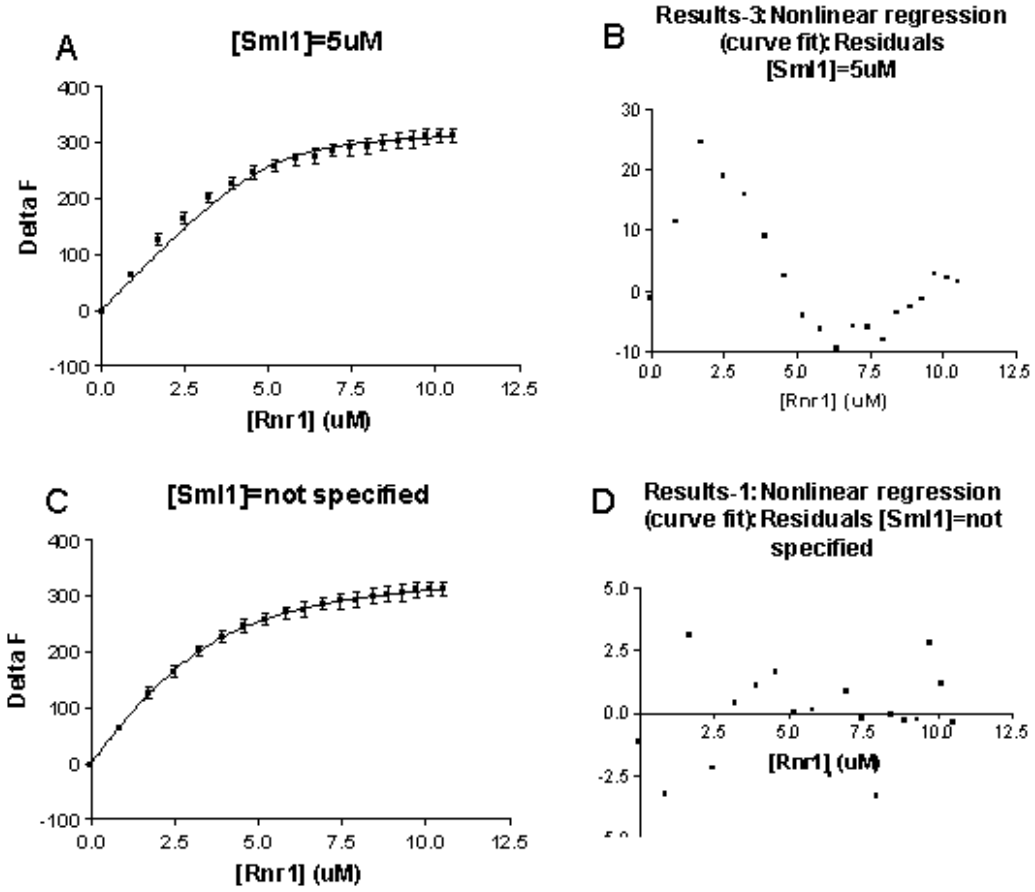

Figure 49. Plot of Rnr1 concentration and Delta F fitted to the one-to-one binding model.

(A) Fitting when Sml1 concentration was specified to be 5μM. **(B)** Residual plot of (A). **(C)** Fitting when Sml1 concentration was not specified. **(D)** Residual plot of (C).

reduced whenever the sample is transferred from vessel to vessel. This is mainly due to adsorption of protein on the wall of vessel and pipette tip. In case of Rnr1, effect of sample loss over its concentration might be relatively small, because Rnr1 was more concentrated (30μM of Rnr1 monomer concentration was aimed in the solution added to the cuvette). In addition, K_d±Std error for the best-fit equation given by the second method was 1.07±0.47μM, while that given by the first method was 0.28±0.11μM. The error associated with K_d given by the second method (±0.47μM) was higher than error given by the first method (±0.11μM). Possibly, the unspecified Sml1 concentration in the second method allows more error for K_d during the fitting procedure. In fact, when Sml1 concentration is specified to be 3.25μM (i.e. fixed at 3.25μM), K_d±Std error for the best-fit equation was 1.07±0.18μM, exactly the same average of K_d with significantly smaller error. However, given the fact that there is uncertainty in the concentration of IANBD-S60C Sml1, this model was not chosen.

Another possible explanation for the discrepancy between IANBD-S60C Sml1 concentration given by the data fitting procedure (3.25±0.65μM) and the concentration we aimed (5μM) is that Rnr1 concentration may be underestimated. For example, if we assume that actual Rnr1 concentration in the cuvette is 1.55 times more than we aimed, the same data fitting procedure (the second method, not specifying Sml1 concentration) gives 5.03±1.03μM of IANBD-S60C Sml1 concentration, which is close to the concentration of IANBD-S60C Sml1 we initially aimed to have in the cuvette (5μM). In this case, K_d±Std error for the best-fit equation was 1.66±0.74μM. However, this model may not be correct because it is unlikely that actual concentration of Rnr1 in the cuvette

is considerably higher (0.55 times more) than the concentration we aimed at (although the opposite situation can happen).

To improve accuracy of our model, it is important to have accurate concentration of proteins in the cuvette. A strategy to obtain accurate concentration of IANBD-S60C Sml1 is to measure IANBD-S60C Sml1 concentration when it is in the cuvette. For example, it may be possible to measure concentration of IANBD-S60C Sml1 in the cuvette based on absorbance at 280nm or 478nm[45]. In this way, potential loss of IANBD-S60C Sml1 during sample transfer would be taken into consideration. Another strategy may be to include carrier proteins such as BSA in the sample to minimize loss of IANBD-Sml1 as well as Rnr1 due to adsorption of these proteins to wall of the vessel and pipette tips, which occurs during sample transfer. In addition, it is important to more accurately measure concentration of proteins in the beginning. Our current method to measure protein concentration using Coomassie Protein Assay Kit may not provide accurate values for concentration of Sml1 or Rnr1 (or both). Using more accurate method of protein concentration measurement may be necessary in the future.

Providing dissociation constant of the Sml1-Rnr1 interactions reported by Thelander and co-workers (0.4±0.1μM), the dissociation constant we obtained (1.07±0.47μM) is probably within the range of (or fairly close to) actual dissociation constant. However, current method needs improvement to obtain more accurate value of dissociation constant.

---

[45] Absorbance maximum of IANBD is 478nm.

Derivation of the equation for one to one binding model

The assumptions taken in the derivation of Equation 1 are:

1. Molar ratio of Sml1 and Rnr1 in the complex is one to one.
2. The fluorescence intensity is proportional to fraction of Sml1 that is bound to Rnr1.

Based on the first assumption, concentration of Sml1 bound to Rnr1 is regarded as:

[Sml1 bound] = [Complex] (Equation 2)

([Sml1 bound]: Concentration of IANBD-S60C Sml1 bound to Rnr1. [Complex]: Concentration of IANBD S60C Sml1-Rnr1 complex.)

Based on the second assumption and Equation 2, relation between fluorescence intensity and total concentration of Sml1 can be written as:

$$\frac{DeltaF}{DeltaF\,\max} = \frac{[Sml1bound]}{[Sml1]} = \frac{[Complex]}{[Sml1]} \qquad \text{(Equation 3)}$$

(Delta F max: Delta F when IANBD-S60C Sml1 is saturated with Rnr1. [Sml1]: *Total* concentration of IANBD-S60C Sml1.)

On the other hand, K_d can be written as:

$$K_d = \frac{[Sml1\,free][Rnr1\,free]}{[Complex]} = \frac{([Sml1]-[Complex])([Rnr1]-[Complex])}{[Complex]}$$ (Equation 4-1)

([Sml1free]: Concentration of free Sml1. [Rnr1free]: Concentration of free Rnr1. [Sml1]: *Total* concentration of IANBD-S60C Sml1. [Rnr1]: *Total* concentration of Rnr1)

Equation 3 can be rearranged as a quadratic equation of concentration of Sml1-Rnr1 complex ([Complex]):

$[Complex]^2 - ([Sml1]+[Rnr1]+K_d)[Complex] + [Sml1][Rnr1] = 0$ (Equation 4-2)

The solutions of this equation are:

$$[Complex] = \frac{([Sml1]+[Rnr1]+K_d) \pm \sqrt{([Sml1]+[Rnr1]+K_d)^2 - 4[Sml1][Rnr1]}}{2}$$

(Equation 4-3)

One solution

$$[Complex] = \frac{([Sml1]+[Rnr1]+K_d) + \sqrt{([Sml1]+[Rnr1]+K_d)^2 - 4[Sml1][Rnr1]}}{2}$$

is wrong. If this solution is correct, it leads to a premise:

[Complex] > ([Sml1] + [Rnr1])/2

(Note: [Sml1] and [Rnr1] are total concentration of Sml1 and Rnr1 respectively.)

This premise is wrong because; concentration of the Sml1-Rnr1 complex is written as:

[Complex] = ([$Sml1_{bound}$][$Rnr1_{bound}$])/2.

([$Sml1_{bound}$]: Concentration of Sml1 in the complex. [$Rnr1_{bound}$]: Concentration of Rnr1 in the complex).

And

[$Sml1_{bound}$] < [Sml1]

[$Rnr1_{bound}$] < [Rnr1]

(Note: [Sml1] and [Rnr1] are total concentration of Sml1 and Rnr1 respectively.)

Therefore

[Complex] < ([Sml1][Rnr1])/2.

Assuming that another solution is correct, concentration of the Sml1-Rnr1 complex ([Complex]) can be written as:

$$[Complex] = \frac{([Sml1] + [Rnr1] + K_d) - \sqrt{([Sml1] + [Rnr1] + K_d)^2 - 4[Sml1][Rnr1]}}{2}$$

(Equation 5)

Replacing concentration of the Sml1-Rnr1 complex in Equation 3 ([Complex]) with the terms on the right side of Equation 5, Equation 3 can be written as:

$$\frac{DeltaF}{DeltaF\max} = \frac{([Sml1]+[Rnr1]+K_d)-\sqrt{([Sml1]+[Rnr1]+K_d)^2-4[Sml1][Rnr1]}}{2}\times\frac{1}{[Sml1]}$$

(Equation 6)

Rearrangement of Equation 6 yields Equation 1:

$$DeltaF = DeltaF\max\frac{([Sml1]+[Rnr1]+K_d)-\sqrt{([Sml1]+[Rnr1]+K_d)^2-4[Sml1][Rnr1]}}{2[Sml1]}$$

Vita

Tomoaki Uchiki was born in Saitama, Japan on February 21, 1974. He was raised in Saitama, Japan. He graduated from Jiyu Gakuen High School in Tokyo, Japan in 1992, and graduated from Jiyu Gakuen Saiko Gakubu (a post-high school training school) in Tokyo, Japan in 1996. He enrolled Earlham College in Richmond, IN and received B.A. in Biology in 1999.

He enrolled University of Tennessee-Oak Ridge National Laboratory Graduate School of Genome Science and Technology in 1999 to pursue his doctorate in Life Sciences.