



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Structural Analysis of the CDK-Cyclin Complex of Pho85-Pho80 and Genome-Wide Characterization of the Phosphate Starvation Response in *Schizosaccharomyces pombe*

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Carter-O'Connell, Ian O'Brien. 2012. Structural Analysis of the CDK-Cyclin Complex of Pho85-Pho80 and Genome-Wide Characterization of the Phosphate Starvation Response in <i>Schizosaccharomyces pombe</i> . Doctoral dissertation, Harvard University.
Accessed	April 17, 2018 3:30:10 PM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:9414562
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

© 2012 – *Ian O'Brien Carter-O'Connell*
All rights reserved.

Advisor: Professor Erin K. O'Shea

Author: Ian O'Brien Carter-O'Connell

Structural Analysis of the CDK-Cyclin Complex of Pho85-Pho80 and Genome-Wide Characterization of the Phosphate Starvation Response in *Schizosaccharomyces pombe*

Abstract

Inorganic phosphate is an essential nutrient required by all organisms for optimal growth. During phosphate starvation, *Saccharomyces cerevisiae* induces a set of genes responsible for the regulation of inorganic phosphate acquisition. The phosphate-responsive signaling (PHO) pathway controls this response, with the CDK-cyclin complex Pho85-Pho80 playing a prominent role. Here we report the X-ray structure of the Pho85-Pho80 complex, identifying the unique structural features that distinguish it from other cell cycle associated CDK-cyclin complexes. The structure reveals a specific salt bridge between a Pho85 arginine and a Pho80 aspartate that maintains a Pho80 loop conformation important for substrate recognition and makes phosphorylation of the Pho85 activation loop dispensable. We show that a cluster of residues distal to the kinase active site are involved in a high affinity interaction between the Pho80 cyclin and the transcription factor substrate (Pho4). The structure also reveals a separate high affinity binding site for the CDK inhibitor (Pho81).

The fission yeast, *Schizosaccharomyces pombe*, regulates expression of the secreted acid phosphatase (*pho1*⁺) via a non-orthologous PHO pathway. The genes induced by phosphate limitation and the molecular mechanism by which the genetically identified positive (*pho7*⁺) and negative (*csk1*⁺) regulators function are not known. Here we use a combination of molecular biology, expression microarrays,

Advisor: Professor Erin K. O'Shea

Author: Ian O'Brien Carter-O'Connell

chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq), and global transcriptome sequencing (RNA-Seq) to characterize the role of *pho7*⁺ and *csk1*⁺ in the PHO response. We show that there is a fast and slow response to phosphate starvation, each with defined regulatory roles. We use ChIP-Seq to identify members of the Pho7 regulon and characterize Pho7 binding dynamics in response to phosphate-limitation and Csk1 activity. We identify a conserved PHO response for the *PHO5* (*pho1*⁺), *PHO84* (*spbc8e4.01c*⁺), and *GIT1* (*spbc1271.09*⁺) orthologs. We show that activation of *pho1*⁺ requires Pho7 binding to a UAS in the *pho1*⁺ promoter and that a URS is necessary for Csk1 repression. We find that Pho7-dependent activation is not limited to phosphate-starvation, as additional environmental stress response pathways require *pho7*⁺ for maximal induction. Using RNA-Seq we show that Pho7 is also involved in regulating non-coding transcription and is a bi-functional transcription factor.

Table of contents

ABSTRACT.....	iii
LIST OF TABLES AND FIGURES.....	ix
STATEMENT OF CONTRIBUTIONS.....	xi
ACKNOWLEDGEMENTS.....	xii
CHAPTER 1: AN INTRODUCTION TO THE PHOSPHATE-RESPONSIVE SIGNALLING (PHO) PATHWAY IN SACCHAROMYCES CEREVISIAE AND SCHIZOSACCHAROMYCES POMBE.....	1
MAINTAINING INTERNAL BALANCE DURING STRESS: MICROORGANISM HOMEOSTASIS.....	2
THE <i>S. CEREVISIAE</i> PHO REGULATORY PROTEINS.....	6
<i>The Transcription Factors: Pho4 and Pho2.....</i>	<i>6</i>
<i>The Pho4 Negative Regulator: Pho85-Pho80.....</i>	<i>8</i>
<i>The Pho85-Pho80 Inhibitory Complex: Pho81 and IP₇.....</i>	<i>11</i>
<i>Upstream of Pho81-IP₇: PHO84, SPL2, ADO1, and ADK1.....</i>	<i>13</i>
PHO IN ASCOMYCOTA: FROM SACCHAROMYCES PARADOXUS TO NEUROSPORA CRASSA.....	15
PHOSPHATE STARVATION IN <i>S. POMBE</i> : A PARALLEL PHO PATHWAY.....	18
REFERENCES.....	22
CHAPTER 2: STRUCTURE OF THE PHO85-PHO80 CDK-CYCLIN COMPLEX OF THE PHOSPHATE-RESPONSIVE SIGNAL TRANSDUCTION PATHWAY.....	32
AUTHOR CONTRIBUTIONS.....	33
ABSTRACT.....	34
INTRODUCTION.....	35
RESULTS AND DISCUSSION.....	37

<i>Structure of the Pho85-Pho80 Complex</i>	37
<i>Discovery of a Specific Mechanism for Circumventing Activation Loop Phosphorylation</i>	42
<i>Potential Involvement of the D Loop in Substrate Recognition</i>	44
<i>A Unique Site on Pho80 for Tight Binding of Pho4 Substrate</i>	49
<i>Location of a Site on Pho80 for Tight Binding of Pho81 CKI</i>	54
CONCLUSIONS.....	55
EXPERIMENTAL PROCEDURES.....	56
<i>Bacterial Expression and Purification of Pho85-Pho80</i>	56
<i>Crystallization, Diffraction Data Collection, and Structure Determination</i>	56
<i>Probing the Roles of the F138 Residue Close to the Active Site and the Distal Site on Pho80 in Substrate Recognition</i>	57
ACKNOWLEDGEMENTS.....	57
REFERENCES.....	57
ACCESSION NUMBERS.....	60
CHAPTER 3: GENOME-WIDE CHARACTERIZATION OF THE PHOSPHATE STARVATION RESPONSE IN <i>SCHIZOSACCHAROMYCES POMBE</i>	61
AUTHOR CONTRIBUTIONS.....	62
INTRODUCTION.....	63
RESULTS.....	66
<i>Identification of the PHO Responsive Genes in <i>S. pombe</i></i>	66
<i>pho7⁺ and csk1⁺ Regulate a Core Subset of the PHO Response in <i>S. pombe</i></i>	71
<i>Pi-Starvation Causes Pho7 Enrichment at the PHO Core Promoters</i>	77

<i>Identification of the Pho7 Upstream Activating Sequence (UAS) in the $pho1^+$ Promoter</i>	82
<i>Attempts at Demonstrating a Direct Link Between Csk1 Kinase Activity and Pho7 Repression</i>	87
<i>Csk1 Represses Pho7 Through an Upstream Repressor Sequence (URS)</i>	88
<i>Pho7 Regulates Expression in Response to Additional Stressors Besides Pi Limitation</i>	93
<i>Pho7 is a Bi-functional Regulator of Coding, Non-Coding, and Anti-Sense Transcription</i>	107
CONCLUSIONS.....	116
MATERIALS AND METHODS.....	118
<i>Growth Conditions and Strains</i>	118
<i>Microarray Analysis and Data Processing</i>	119
<i>Chromatin Immunoprecipitation</i>	121
<i>$pho1^+$ Promoter Deletion Analysis</i>	122
<i>Chromatin Immunoprecipitation of Pho7-TAP with High-Throughput Sequencing (ChIP-Seq)</i>	123
<i>In Silico Motif Discovery</i>	124
<i>$pho7^+$ Regulation in Additional Stress Response Pathways</i>	124
<i>High-Throughput Sequencing of the <i>S. pombe</i> Transcriptome (RNA-Seq)</i>	125
ACKNOWLEDGEMENTS.....	127
REFERENCES.....	128
CHAPTER 4: CONCLUSIONS REGARDING THE <i>S. CEREVISIAE</i> PHO85-PHO80 STRUCTURE AND THE PHO PATHWAY IN <i>S. POMBE</i>	135
STRUCTURE AND FUNCTION OF THE PHO85-PHO80 COMPLEX.....	136

THE <i>S. POMBE</i> PHO PATHWAY AND BEYOND: THE <i>PHO7</i> ⁺ REGULATORY MECHANISM.....	139
REFERENCES.....	146
APPENDIX A: CHAPTER 2 SUPPLEMENTAL DATA.....	148
SUPPLEMENTAL EXPERIMENTAL PROCEDURES.....	149
<i>Bacterial Expression and Purification of Pho85-Pho80.....</i>	<i>149</i>
<i>Crystallization and Diffraction Data Collection.....</i>	<i>149</i>
<i>Determination of the Crystal Structure of Pho85-Pho80.....</i>	<i>150</i>
<i>Site-Directed Mutagenesis and Kinase Assays.....</i>	<i>151</i>
<i>Identification of a Shorter Segment of the 80 Residue Fragment of Pho81 that Binds Pho85-Pho80.....</i>	<i>154</i>
SUPPLEMENTAL REFERENCES.....	160
APPENDIX B: CHAPTER 3 SUPPLEMENTAL DATA.....	163
SUPPLEMENTAL EXPERIMENTAL PROCEDURES.....	164
<i>Pho7-TAP Expression and Purification.....</i>	<i>164</i>
SUPPLEMENTAL REFERENCES.....	177

List of tables and figures

Figure 1.1.	<i>Saccharomyces cerevisiae</i> Phosphate-Responsive Signal Transduction (PHO) Pathway.....	5
Figure 1.2.	<i>Schizosaccharomyces pombe</i> PHO Pathway.....	21
Table 2.1.	Data Collection and Refinement Statistics.....	38
Figure 2.1.	Structure of Pho85-Pho80.....	40
Figure 2.2.	A Salt Link Between Pho85 R132 and Pho80 D136 that Makes T Loop Phosphorylation Dispensable.....	43
Figure 2.3.	Three Functional Sites of the Pho80 Cyclin Subunit.....	46
Table 2.2.	Probing the Role of F138 in Pho80 Substrate Recognition.....	49
Table 2.3.	Role of Distal Site Cluster on Pho80 in Pho4 Substrate Recognition.....	52
Figure 3.1.	Temporal Dynamics of the Phosphate Starvation Response in <i>S. pombe</i>	69
Figure 3.2.	Pho7 and Csk1 Regulate a Core Set of PHO Responsive Genes.....	74
Table 3.1.	Microarray Results for Genes Regulated by Pi-Starvation, <i>pho7</i> ⁺ , and/or <i>csk1</i> ⁺	76
Figure 3.3.	Pho7 is Recruited to the Promoter During Pi Starvation.....	80
Figure 3.4.	Pho7 Alignment to Other Zn ₂ Cys ₆ Binuclear Cluster Proteins and the Identification of the Pho7 Upstream Activating Sequence in the <i>pho1</i> ⁺ Promoter.....	85
Figure 3.5.	Csk1 Represses Pho7 Activity Through An Upstream Repressor Sequence.....	91
Figure 3.6.	<i>pho7</i> ⁺ Functions in Additional Stress Responses.....	96
Table 3.2.	Gene Ontology Enrichment for Pho7-Dependent Genes Identified via Microarray and CHIP-Seq Analysis.....	98
Figure 3.7.	Microarray Analysis of the Pho7-Dependent Stress Response.....	102

Table 3.3.	Microarray Results for Genes Regulated by Phosphate, Iron, Copper, and/or Osmotic Stress in a Pho7-Dependent Manner.....	104
Figure 3.8.	<i>pho7+</i> Regulates Non-Coding Transcription in Both a Pi-dependent and Pi-independent Manner.....	112
Figure 3.9.	RNA-Seq Profiles for Representative <i>pho7+</i> Regulated Transcripts.....	114
Table S2.1.	Substrates of Pho85-Pcl Complexes and the D Loop Sequences of the Pho80/Pcl Subunits.....	155
Figure S2.1.	Amino Acid Sequence Alignment of Pho80 and the Other Nine <u>Pho85 Cyclins</u> (Pcls).....	156
Figure S2.2.	Hydrophobic Interactions of the Aliphatic Residues on the PSTAIRE/PSSALRE Helices of Pho85, CDK2, and CDK5 with Residues of Their Respective Cyclins.....	158
Figure S2.3.	The Relationship of the $\alpha 5$ Helices with $\alpha 3$ and $\alpha 4$ of Pho80 or Cyclin A and with $\alpha 3$ of p25.....	159
Figure S2.4.	Identification of a Shorter Segment of the 80 Residue Fragment of Pho81 that Binds Pho85-Pho80.....	160
Figure S3.1.	Microarray Epistasis Analysis for <i>pho7</i> ⁺ -Regulated Genes.....	166
Figure S3.2.	Secondary Structure Prediction for Pho7.....	167
Figure S3.3.	Conservation of Residue T463 in Pho7.....	170
Figure S3.4.	Purification of Pho7-TAP.....	171
Figure S3.5.	ChIP-Seq Binding Profiles for Pho7-TAP at Non-PHO Promoters.....	173
Figure S3.6.	Fluorescent Microscopy of the <i>pho1</i> ⁺ <i>pr-yfp</i> ⁺ Constructs.....	175
Table S3.1.	Primers Used in Chapter 3 Study.....	176

Statement of contributions

Professor Erin K. O'Shea (EKO) supervised all work.

Chapter 1

Ian O'Brien Carter-O'Connell (ICO) wrote the text.

Chapter 2

Kexin Huang (KH), ICO, Wei Zhang (WZ), and Gordan A. Leonard (GAL) contributed to this work. Florante A. Quioco (FAQ) is the corresponding author. See chapter title page for more information.

Chapter 3

ICO and Michael Peel (MP) contributed to this work. EKO and Dennis D. Wykoff (DDW) are the corresponding authors. See chapter title page for more information.

Chapter 4

ICO wrote the text.

Funding

Generous funding was provided by the Howard Hughes Medical Institute.

Acknowledgements

My doctoral training benefited tremendously from an array of interactions with colleagues. First and foremost, I would like to thank my graduate advisor, Dr. Erin K. O'Shea, for providing me with support and an intellectually stimulating environment for biological research. Our scientific discussions challenged me to think critically about all the aspects of my research project, which will be incredibly useful as I continue my post-doctoral career.

In the course of my graduate studies I was able to collaborate with a number of excellent individuals. The Pho85-Pho80 structure that facilitated our biochemical analysis was solved by Dr. Kexin Huang, a former member of the Quioco Group at Baylor University. Drs. Huang and Florante Quioco were instrumental in designing the Pho80 mutants and our discussions regarding the Pho85-Pho80 structure were immensely helpful in beginning my thesis work. Experimentation with the *S.pombe* PHO system would not have been possible without the support of Dr. Dennis Wykoff at Villanova University. Dennis let me run wild using his initial observations regarding phosphate starvation and he was always willing to discuss my newest, often markedly different, hypothesis. Christine Kerwin and Michael Peel from the Wykoff lab provided experimental support for the *S.pombe* project. Although we never were able to get Pho85-Pho80, Pho81, and IP₇ to stay in solution long enough to collect enough NMR data, Drs. Michael Durney and Victoria D'Souza patiently provided their time and structural expertise whenever asked, for which I remain grateful.

I would like to thank Drs. Andrew Murray, Rachelle Gaudet, Vlad Denic, and Victoria D'Souza for serving on my dissertation advisory committee. They provided assistance in navigating the challenges inherent with scientific pursuit and their advice proved incredibly helpful. Kevin Monahan, Elsa Beyer, Emiko Morimoto, Jordan Krall, Abbas Rizvi, Andrian Gutu, Tim Peterson, and Catherine Leimkuhler Grimes were always up for a scientific dialogue or a pint when necessary; I was fortunate to befriend such a talented group of colleagues. In the later stages of my work, Vikram Vijayan, Xu Zhou, and Brian Zid helped push me from

biochemistry into systems biology. Their able coding skills helped me get out of a number of high-throughput data jams.

On a personal note, I never would have had the opportunity to pursue a doctoral degree if it weren't for the unfailing support of my family. My parents purchased my first chemistry kit when I was just eight years old, and I believe that the completion of my graduate work is a testament to the excitement they fostered within me for the natural sciences. My sisters, grandparents, aunts, uncles, and cousins all put up with my never-ending prattle about the newest scientific discovery without complaint. They also made sure I never took myself too seriously, for which they have my thanks. The Carter family opened their home to me, granting me unofficial membership in the Yeasty Boys Brewing Company; thanks for teaching me the real purpose of yeast based research.

Finally, I wish to thank my wife, Joanna Carter-O'Connell, for agreeing to join me in this grad school adventure. Whenever my experiments failed or I began to doubt my ability, she invariably knew what needed to be said to set things right. I truly believe that none of this work would have been possible without her unending wellspring of support and encouragement; you have my deepest thanks, Little 'a'.

Ian Carter-O'Connell

March 2012

CHAPTER 1

An Introduction to the Phosphate-Responsive Signaling (PHO) Pathway in

Saccharomyces cerevisiae* and *Schizosaccharomyces pombe

Maintaining Internal Balance During Stress: Microorganism Homeostasis

Microorganisms exist within complex and rapidly changing environments. Often the flux in nutrient availability results in sub-optimal conditions for the proper growth of the organism. To successfully navigate unpredictable circumstances, microorganisms have amassed a diverse array of regulatory networks capable of sensing stressful conditions, transmitting a signal identifying the stressor, and inducing a transcriptional response to appropriately modify cellular behavior. Regulatory events in response to stress take many forms; during glucose limitation in *Saccharomyces cerevisiae*, for example, the transcription factor Msn2 is activated to induce a plethora of stress responsive genes. As glucose levels drop, Msn2 is dephosphorylated in a tightly coordinated manner (Gorner et al., 2002). This causes Msn2 to be imported into and exported from the nucleus with a frequency dependent on the magnitude of the glucose limitation (Hao and O'Shea, 2011). This behavior results in different profiles of transcriptional up-regulation tuned specifically to the intensity of glucose limitation. Other stresses regulate transcription factor binding more directly. In *Escherichia coli*, the *fur* gene is responsible for regulating transcription of iron uptake genes in response to iron availability (reviewed in (Escolar et al., 1999)). Given that iron is toxic at higher concentrations, this system has to be tightly controlled to keep internal iron concentrations within a very narrow range (Crosa, 1997). During iron repletion Fur binds Fe^{2+} , which causes a conformational change in the protein allowing it to bind DNA (Bagg and Neilands, 1987). Bound Fur- Fe^{2+} represses transcription of iron uptake genes and transcription is only activated when iron concentrations drop enough to allow Fur and Fe^{2+} to dissociate.

Inorganic phosphate homeostasis presents an interesting challenge to microorganisms, as inorganic phosphate is required for the maintenance of many different biologically important molecules, proper signal transduction, and energy storage/transfer. It is essential that organisms keep a constant pool of phosphate available for biochemistry at all times. In *S. cerevisiae* the phosphate-responsive signaling (PHO) pathway is responsible for regulating phosphate homeostasis. In this work we use the PHO pathway to explore nutrient homeostasis at a molecular and global level.

The first step in responding to phosphate starvation is harvesting as much phosphate as possible. *S. cerevisiae* accomplishes this goal by inducing the expression of the secreted acid phosphatase, *PHO5*, and the high-affinity inorganic phosphate transporter, *PHO84*, in response to phosphate starvation (Bostian et al., 1980; Bun-Ya et al., 1991; Lemire et al., 1985). The expression of *PHO5* and *PHO84* is controlled by the transcription factor Pho4 (Ogawa and Oshima, 1990), along with the constitutively nuclear transcription factor Pho2 (Vogel et al., 1989). Phosphate dependent regulation is accomplished by controlling the phosphorylation state of Pho4 (Komeili and O'Shea, 1999). During growth in conditions with high levels of external phosphate Pho4 is multiply phosphorylated by the cyclin-dependent kinase-cyclin (CDK-cyclin) pair Pho85-Pho80 (Kaffman et al., 1994). This phosphorylation causes the export of Pho4 from the nucleus (Kaffman et al., 1998a) and the loss of Pho2 cooperation (Komeili and O'Shea, 1999) preventing the activation of the PHO response. During phosphate starvation the CDK-inhibitor (CKI) Pho81, and the small metabolite *myo*-D-inositol heptakisphosphate (IP₇), combine to inhibit the Pho85-Pho80 complex (Kaffman et al., 1994; Lee et al.,

2007). The loss of kinase activity allows Pho4 to become dephosphorylated, enter the nucleus (Kaffman et al., 1998b), interact with Pho2, and induce the transcription of the PHO regulon. This process is summarized in Figure 1.1.

The work presented here explores the *S. cerevisiae* PHO response in structural and biochemical detail before expanding to a systems level analysis of the PHO pathway in the distantly related yeast, *Schizosaccharomyces pombe*. In chapter 2, we detail the structural features of the Pho85-Pho80 complex that contribute to Pho4 regulation along with the unique structural aspects that distinguish Pho85-Pho80 from other CDK-cyclin pairs. Chapter 3 characterizes the PHO pathway in *S. pombe*, in which the orthologous components to this system have either been lost (Pho4, Pho2, Pho81) or are not involved in the PHO response (Pho85, Pho80). Instead, a parallel PHO pathway has evolved using a putative transcription factor (*pho7*⁺) and a CDK-activating kinase (*csk1*⁺). To provide a framework for this analysis I will first review each of the *S. cerevisiae* PHO components described above before reviewing the PHO pathway in other yeasts.

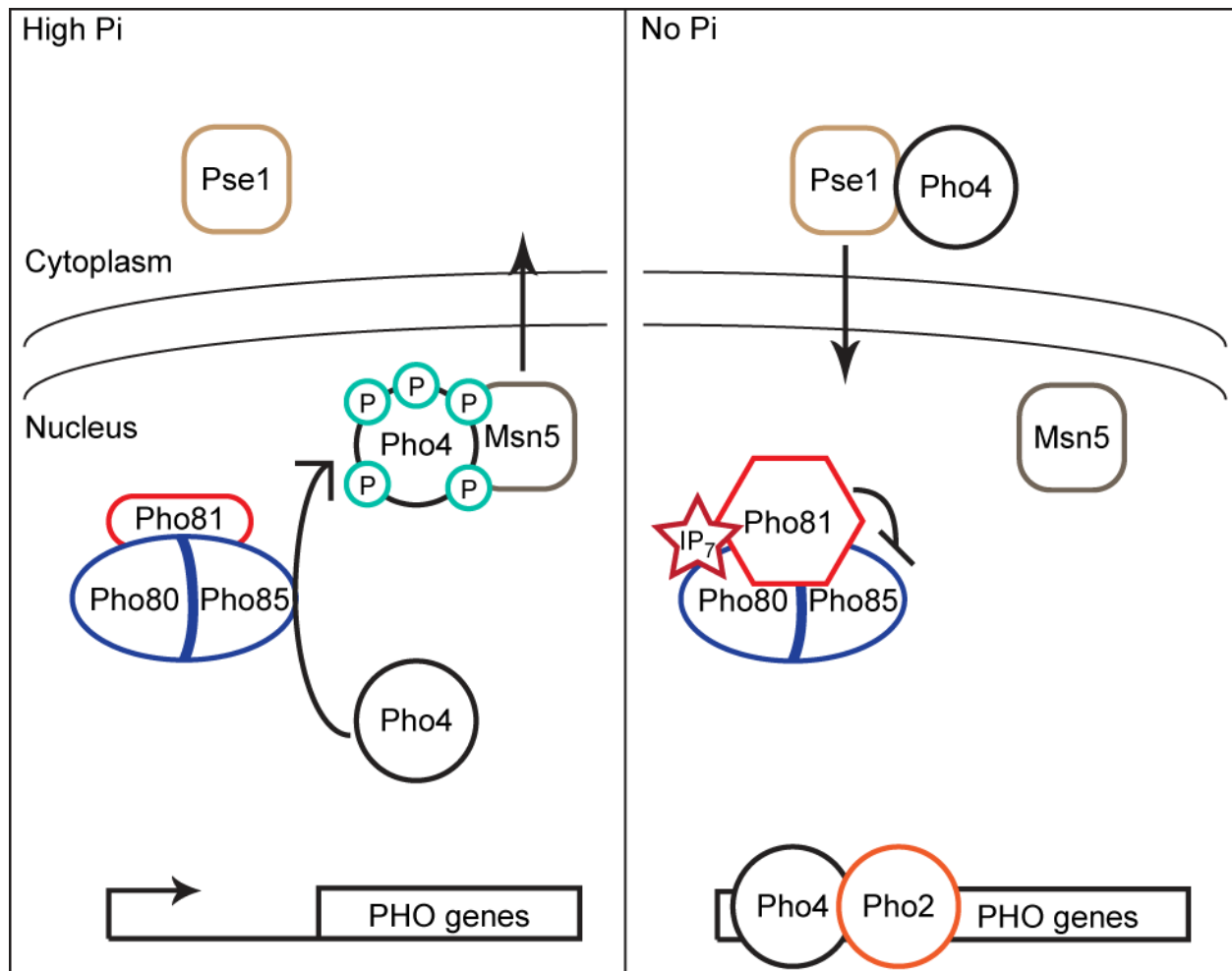


Figure 1.1. *Saccharomyces cerevisiae* Phosphate-Responsive Signal Transduction (PHO) Pathway.

In high phosphate conditions, Pho85-Pho80 phosphorylates Pho4 at five different residues. Hyperphosphorylated Pho4 associates with Msn5 and is exported from the nucleus. Depletion of inorganic phosphate leads to the IP_7 -Pho81-dependent inhibition of the Pho85-Pho80 complex. Pse1 imports unphosphorylated Pho4 into the nucleus; Pho4 interacts with the transcription factor, Pho2, at PHO-specific genes, inducing transcription. The mechanism for Pho4 dephosphorylation is unknown.

The *S. cerevisiae* PHO Regulatory Proteins

The Transcription Factors: Pho4 and Pho2

The *PHO4* and *PHO2* genes were first identified in a deletion screen looking for positive and negative regulators of the secreted acid phosphatase, *PHO5* (To et al., 1973). Pho4 contains a basic helix-loop-helix (bHLH) domain (Berben et al., 1990; Shimizu et al., 1997) and binds as a homodimer to the consensus sequence 5'-CACGTG-3' (Maerkl and Quake, 2007; Ogawa and Oshima, 1990; Shao et al., 1998). Pho2 is a homeodomain transcription factor that is involved in the regulation of several metabolic pathways (PHO, purine biosynthesis, histidine biosynthesis) (Berben et al., 1988; Brazas and Stillman, 1993; Daignan-Fornier and Fink, 1992; Vogel et al., 1989). The activation of PHO genes requires the co-operative binding of both Pho2 and Pho4, and this interaction is inhibited in high phosphate growth (Vogel et al., 1989; Zhou and O'Shea, 2011).

Pho4 regulation occurs via the phosphorylation of five distinct serine residues by the Pho85-Pho80 complex (Komeili and O'Shea, 1999). These sites are SP1, SP2, SP3, SP4 and SP6. During phosphate-rich growth the phosphorylation of these sites impairs three separate biological functions of Pho4 that combine to inactivate it: (1) the phosphorylation of SP2 and SP3 results in the export of Pho4 from the nucleus by the receptor Msn5 (Kaffman et al., 1998a); (2) SP4 phosphorylation prevents the interaction between Pho4 and the nuclear import complex Pse1/Kap121 (Kaffman et al., 1998b; O'Neill et al., 1996); and (3) SP6 phosphorylation prevents the association of Pho4 with Pho2 (Komeili and O'Shea, 1999). There is no known function of the SP1 phosphorylation site. During phosphate starvation these sites become

dephosphorylated through the action of a phosphatase that has not been identified and Pho4 is free to activate transcription.

Historically, expression of *PHO5* and *PHO84* were primarily used as a proxy for studying the PHO response. The advent of whole-genome technologies allowed for the definition of the entire PHO regulon. cDNA microarray comparison of cells grown in high-Pi and no-Pi conditions revealed that 22 genes are regulated by phosphate starvation (Ogawa et al., 2000). In addition to finding previously defined PHO regulated genes with known roles in phosphate uptake (i.e. *PHO81*, *PHO84*, *PHO5*, *PHO89*), this study identified a number of phosphate metabolism (PHM) genes that are involved in polyphosphate maintenance through vacuolar components (Cohen et al., 1999; Sethuraman et al., 2001) and alternative organic phosphate harvesting (Reddy et al., 2008). In a separate study, the magnitude of phosphate starvation was found to influence Pho4 phosphorylation, with intermediate phosphate levels triggering preferential phosphorylation of SP6 with lower levels of SP1-SP4 modification (Springer et al., 2003). Differential phosphorylation activated only a subset of the PHO regulon, providing a mechanism by which the modification of a single transcription factor could lead to combinatorial output. Together, these experiments provide a list of the genes that are responsible for adapting the cell's behavior to decreasing levels of external phosphate.

Recent work developed this analysis further by outlining the global mechanism of Pho4-Pho2 regulation at the promoter level. The Pho4 binding motif occurs throughout the *S. cerevisiae* genome (843 sites), yet only a fraction of these sites (115) are involved in Pho4-dependent regulation (Zhou and O'Shea, 2011). Four factors ensure

that Pho4 only activates the specific PHO components detailed above. First, nucleosomal positioning in *S. cerevisiae* occludes a number of Pho4 binding sites and chromatin structure is itself regulated by phosphate starvation (Lam et al., 2008; Svaren et al., 1994; Venter et al., 1994). Second, Pho4 binds to non-occluded CACGTG core motifs *in vivo* with different binding affinities based on the identity of the 5' flanking nucleotide (i.e. NCACGTG, with a preference for G) (Maerkl and Quake, 2007; Ogawa et al., 1995). Third, this preference is enforced by competition with Cbf1; a bHLH transcription factor that binds to CACGTG motifs flanked by a 5' T (Harbison et al., 2004; MacIsaac et al., 2006; Zhu et al., 2009) and prevents inappropriate binding of Pho4 to non-PHO promoters (Zhou and O'Shea, 2011). Fourth, 96% of the genes that are Pho4-regulated *in vivo* require Pho2 cooperation to achieve maximal induction (Zhou and O'Shea, 2011). In chapter 3, the PHO regulon and the transcription factor dynamics described in *S. cerevisiae* will serve as a point of comparison to the functionally analogous PHO pathway in *S. pombe*.

The Pho4 Negative Regulator: Pho85-Pho80

The Pho85-Pho80 CDK-cyclin complex is responsible for negatively regulating Pho4 activity in response to high levels of inorganic phosphate (Kaffman et al., 1994). CDK family members have long been studied for their role in cell-cycle progression and metabolism (Morgan, 1997; Nigg, 1995). The defining feature of this type of kinase is its requirement for cyclin association to achieve maximal phosphorylation efficiency (De Bondt et al., 1993; Sherr, 1993). Upon cyclin binding the kinase rearranges its catalytic core for optimal alignment of ATP and the substrate peptide, facilitating phosphate

transfer (Jeffrey et al., 1995; Radzio-Andzelm et al., 1995). A single CDK can bind one of several cyclins, and the identity of the cyclin determines substrate specificity for that particular CDK-cyclin pair (Miller and Cross, 2001). Substrate recognition occurs primarily at two sites: (1) the sequence motif targeted for phosphorylation (Songyang et al., 1994; Srinivasan et al., 1995) and (2) a site distal to the modified S/T residue that is bound by the CDK-cyclin (Adams et al., 1999; Schulman et al., 1998; Takeda et al., 2001).

A family of ten Pho85-cyclins (Pcls) regulates Pho85 substrate recognition (Huang et al., 2007a; Measday et al., 1997). The Pcl family is further sub-divided based on sequence similarity to either *PHO80* (*PHO80*, *PCL6*, *PCL7*, *PCL8*, *PCL10*) or *PCL1, 2* (*PCL1*, *PCL2*, *PCL5*, *PCL9*, *CLG1*) (Measday et al., 1997). The Pho80-like cyclins target Pho85 towards substrates involved in stress responsive pathways (Pho80) or glycogen regulation (Pcl6, 7, 8, 10) (Huang et al., 2007a; Huang et al., 1998; Timblin et al., 1996; Wang et al., 2001; Wilson et al., 2002, 2005). The Pcl1, 2-like cyclins play a role in cell-cycle progression and cell integrity (Pcl1, 2, 9, Clg1) (Espinoza et al., 1994; Huang et al., 2002; Measday et al., 1994; Measday et al., 1997) or amino acid starvation (Pcl5) (Bomeke et al., 2006). The sequence recognition motifs that are phosphorylated by each Pho85-Pcl complex in these diverse situations are known in some cases (Friesen et al., 2003; Meimoun et al., 2000; Nishizawa et al., 1998; Reinders et al., 1998; Tan et al., 2003; Wilson et al., 1999), but the biochemical and structural mechanisms dictating specificity are not. For the Pho85-Pho80 complex both the substrate recognition motif and the distal binding site on Pho4 are known (Byrne et

al., 2004; Komeili and O'Shea, 1999; O'Neill et al., 1996), making the Pho85-Pho80 structure an ideal candidate for exploring substrate specificity in this system.

Pho85 shows significant similarity to a number of CDKs involved in cell-cycle progression (i.e. Cdc28/CDK2) (Toh-e et al., 1988) and neuronal migration (CDK5) (Lew and Wang, 1995). In fact, overexpression of the murine CDK5 partially complements the phenotypes caused by the loss of *PHO85* in *S. cerevisiae* (Huang et al., 1999; Nishizawa et al., 1999a). Yet there are a number of unique features specific to the Pho85-Pho80 complex that are interesting. First, in many CDK-cyclin complexes phosphorylation of a conserved threonine or serine residue on the activating loop (T-loop) of the CDK is required to achieve maximal activation (Morgan, 1996; Russo et al., 1996b). The same residue in the Pho85 T-loop is dispensable when measuring Pho85-Pho80 kinase activity (Nishizawa et al., 1999b). Second, as discussed above, CDK-cyclin complexes bind and modify a specific sequence motif in the target substrate. The consensus sequence for this modification is commonly S/TPXK/R (Holmes and Solomon, 1996; Songyang et al., 1994). For Pho85-Pho80 the sequence from the SP sites in Pho4 is SPXI/L (O'Neill et al., 1996), requiring a different mode of substrate recognition. Third, unlike other CDK-cyclin kinase reactions, the phosphorylation of Pho4 by Pho85-Pho80 occurs in a processive manner (Jeffery et al., 2001). This means that for each Pho4 binding event an average of 2.1 phosphates are transferred. The nature of this processivity has been attributed to a distal Pho80 binding site on Pho4 (Byrne et al., 2004), though the site on Pho80 responsible for this interaction with Pho4 is not known. Finally, the biochemical mechanism for Pho81-mediated inhibition (see below) is remarkably different for Pho85-Pho80 than that for canonical CKI-CDK-cyclin

complexes. In chapter 2 we use our crystal structure of Pho85-Pho80 to investigate each of these observations from a structural and biochemical viewpoint.

The Pho85-Pho80 Inhibitory Complex: Pho81 and IP₇

Pho81 is a large protein (134 kDa) originally identified as a positive regulator for *PHO5* expression (To et al., 1973). Pho81 acts as negative regulator of *PHO85* and *PHO80* (Ueda et al., 1975). Pho81 binds Pho85-Pho80 in both high-Pi and no-Pi conditions and inhibits kinase activity in response to phosphate starvation (Schneider et al., 1994). Inhibition of the CDK-cyclin Pho85-Pho80 places Pho81 in the class of proteins known as CDK-inhibitors (CKIs). CKIs function in eukaryotic cell cycle control and oncogenesis (reviewed in (Grana and Reddy, 1995; Harper and Elledge, 1996; Peter and Herskowitz, 1994)). They serve as checkpoints for cell cycle progression by inhibiting their regulated CDK-cyclin pairs in response to developmental cues designed to pause cell growth and proliferation (Parker et al., 1995; Serrano et al., 1993). CKIs can be grouped into two families based on sequence identity and functional similarities. The Kip/Cip family includes p27^{Kip1}, p21^{Cip1}, and p57^{Kip2} (Harper et al., 1993; Matsuoka et al., 1995; Toyoshima and Hunter, 1994; Xiong et al., 1993), which preferentially inhibit CDKs in the G₁/S phase (Harper and Elledge, 1996). The crystal structure of the 69-amino-acid N-terminal region of p27^{Kip1} bound to Cdk2-cyclin A revealed the basis for inhibition (Russo et al., 1996a). The CKI in the structure contacts both Cdk2 and cyclin A and inserts a pseudo-substrate that mimics ATP into Cdk2's kinase site. This causes a rearrangement of the catalytic site, locking the CDK-cyclin in an inactive state, and only occurs with preformed CDK-cyclin complex.

The INK4 (inhibitor of Cdk4) family comprises the other class of CKIs. INK4s contain an ankyrin repeat structure important for inhibition (discussed below) and include the proteins p16^{INK4a}, p15^{INK4b}, p18^{INK4c}, and p19^{INK4d} (Hannon and Beach, 1994; Hirai et al., 1995; Jen et al., 1994; Serrano et al., 1993). Crystal structures of Cdk6 bound by either p16^{INK4a} or p19^{INK4d} reveal the ankyrin repeat domain making extensive contact with the CDK (Russo et al., 1998). This causes a major re-alignment of Cdk6 that prevents Cdk6 from associating with its cognate cyclin. Unlike the Kip/Cip family, the INK4 family inhibits CDK-cyclin activity by competing with the free cyclin for CDK binding. Expression of both CKI families is induced transcriptionally when the cell requires inhibition of the correct CDK-cyclin complex. CKIs are also regulated post-translationally by a number of phosphorylation and degradation pathways (Morgan, 1997).

Pho81 does not fit within either family of canonical CKIs. First, Pho81 does not compete with Pho80 for Pho85 association (like an INK4 CKI) nor is binding sufficient to inhibit the Pho85-Pho80 complex (as it is for a Kip/Cip CKI). Second, although expression of *PHO81* is increased ~ 2-fold during phosphate starvation (Ogawa et al., 2000), similar levels of Pho81 are bound to the Pho85-Pho80 complex irrespective of external phosphate conditions (Schneider et al., 1994). Third, the domain responsible for Pho81-based inhibition shares no sequence similarity to the other CKIs. Huang et al. identified a novel 80 amino-acid segment of Pho81 (residues 644-723) termed the Pho81 minimal domain (Pho81MD) that is necessary and sufficient for Pho81 based inhibition (Huang et al., 2001). Fourth, inhibition mediated by Pho81 requires the small molecule *myo*-D-inositol heptakisphosphate (IP₇) (Lee et al., 2007). During phosphate

starvation the enzyme Vip1 converts pools of *myo*-D-inositol hexakisphosphate (IP₆) into IP₇, and IP₇ mediates the switch between inactive Pho81 and inhibitory Pho81. Pho81MD, IP₇ and Mg²⁺ are the minimal components required to inhibit Pho85-Pho80 *in vitro* and inhibition proceeds in a reversible, non-covalent manner (Lee et al., 2008). Pho81MD can be broken down even further into three separate peptides: S2, a 21 amino-acid peptide (residues 644-665) dispensable for inhibition; S3, a 35 amino-acid peptide (residues 666-701) that binds constitutively in a non-inhibitory manner to Pho85-Pho80; and S1, a 21 amino-acid peptide (residues 702-723) that binds only in the presence of IP₇ and inhibits Pho85-Pho80 activity (Huang et al., 2007b; Lee et al., 2008). These differences between Pho81 and canonical CKI inhibitory mechanisms led us to ask: what are the structural and biochemical factors that facilitate Pho81-IP₇ inhibition? To address this question, we have identified Pho81 interaction sites on Pho85-Pho80 and speculate on how this might influence Pho81-IP₇ inhibition in chapter 2.

Upstream of Pho81-IP₇: *PHO84*, *SPL2*, *ADO1* and *ADK1*

The first step in activating the PHO pathway occurs upstream of Pho81 and IP₇. We know that IP₇ is not the initial internal signal alerting the cell to inorganic phosphate limitation because in the inositol biosynthetic pathway the activity of Vip1 is held in check by the IP₇ phosphatase Ddp1 (Lee et al., 2007; Safrany et al., 1999; Yang et al., 1999). In high-Pi conditions, IP₇ stores are constantly being converted into IP₆, preventing the activation of the PHO pathway. The molecular mechanism regulating the ratio of IP₇ to IP₆ and the switch increasing IP₇ production is unknown. A clue to this

process may come from the genes *ADO1* and *ADK1*. Huang and O'Shea identified adenylate kinase (*ADK1*) and adenosine kinase (*ADO1*) in a systematic deletion screen for PHO regulatory components upstream of *PHO81* (Huang and O'Shea, 2005). Deletion of either gene causes the constitutive expression of *PHO5* (Huang and O'Shea, 2005). The biochemical details of this interaction are unknown; future experiments should explain how proteins associated with nucleotide metabolism are involved in regulation of the PHO response.

Although the mechanism linking external phosphate limitation to Pho81-IP₇ inhibition remains unclear, much is known about the upstream process of importing inorganic phosphate into the cell. In high-Pi conditions *S. cerevisiae* constitutively expresses a host of low-affinity inorganic phosphate transporters that transport phosphate into the cell (Ghillebert et al., 2011; Nieuwenhuis and Borst-Pauwels, 1984; Wykoff and O'Shea, 2001). Thus, phosphate transport is achieved without activating the PHO regulon. During starvation these transporters are repressed through the activity of *SPL2* (Ghillebert et al., 2011; Hurlimann et al., 2009; Wykoff et al., 2007), causing positive feedback in the PHO pathway by temporarily lowering the phosphate transport rate and enhancing perceived starvation. Starvation induces the expression of the high-affinity inorganic transporter, *PHO84* (Ogawa et al., 1995). Pho84 increases the phosphate transfer rate causing a dampening of the PHO response and providing negative feedback to the system (Wykoff and O'Shea, 2001). Positive and negative feedback is processed through the Pho81-IP₇-Pho85-Pho80 complex, creating the conditions necessary for bistability in intermediate phosphate conditions (Wykoff et al., 2007). Bistability occurs because each individual cell can accomplish sufficient

phosphate uptake by expressing either many low-affinity transporters or one high-affinity transporter and the feedback system effectively “locks” in this choice. Future experimentation in different fungal species should reveal whether this bistability is a conserved feature of the PHO pathway.

PHO in Ascomycota: From *Saccharomyces paradoxus* to *Neurospora crassa*

Members of the ascomycota phylum have experienced diverse environmental conditions throughout evolutionary history. They have evolved gene genetic expression profiles to optimally exploit their environmental niches; yet, they all face the challenges described above for maintaining nutrient homeostasis. Given that successfully managing environmental flux places a strong selective pressure on an organism, signaling components required for a given stress response are often conserved through evolutionary time (Pinter et al., 2005). However, shared signaling components between related species can differ in how they integrate upstream signals that generate downstream transcriptional output (Gasch, 2007). Rewiring of signal transduction cascades can occur through *cis*-changes at the promoter – affecting transcription rate, transcription initiation, and transcript stability – or through *trans*-changes in the factors responsible for interacting with *cis*-elements – affecting transcription factor expression/activity (Chang et al., 2008; Wittkopp et al., 2004). Re-wiring allows each species to adapt a core stress response network to the unique challenges they face within their niche. The PHO pathway in *S. cerevisiae* is well suited for understanding how stress responsive pathways have evolved over time because we have a

comprehensive list of the *cis*- and *trans*-regulatory elements governing the PHO response that might have diverged in other ascomycetes.

The sequencing of many ascomycete genomes has greatly facilitated this analysis (Dujon et al., 2004). Using the available orthogroup assignments for the 23 curated ascomycete genomes reveals a high degree of conservation for the signaling components in the PHO pathway (Wapinski et al., 2007). Orthologs for *PHO80*, *PHO81*, and *PHO2* are found in 20 species, orthologs for *PHO4* in 19, and orthologs for *PHO85* in all 23. The two primary downstream effectors, *PHO5* and *PHO84*, are found in 17 and 23 species, respectively.

In some cases, it appears that the conserved signaling components share functional roles between species. In *Kluyveromyces lactis* the *PHO5* ortholog, *KIPHO5*, is induced ~500-fold during phosphate starvation (Ferminan and Dominguez, 1997). The *cis*-regulatory elements upstream of *KIPHO5* are conserved, as both Pho4 and Pho2 binding sites occur in the promoter. A 2.5Kb segment of *KIPHO5*, along with its native promoter, fully complements the *pho5Δ* phenotype when transformed into *S. cerevisiae* (Ferminan and Dominguez, 1997). In *Candida glabrata* the roles for *PHO81*, *PHO80*, *PHO4*, and *MSN5* were all confirmed by deletion analysis (Kerwin and Wykoff, 2009). As in *S. cerevisiae*, the loss of *PHO81* and *PHO4* results in a loss of phosphate-responsive secreted acid phosphatase activity, while the loss of *PHO80* and *MSN5* results in constitutive secreted acid phosphatase activity (Kerwin and Wykoff, 2009). Finally, in *Neurospora crassa* the orthologs to *PHO81* (NUC-2), *PHO80* (PREG), *PHO85* (PGOV), and *PHO4* (NUC-1) have been identified and their roles in phosphate

responsive signaling are functionally equivalent to those in *S. cerevisiae* (Metzenberg and Chia, 1979).

There is increasing evidence, however, that the PHO signaling components in the ascomycetes have been subjected to functional alterations during speciation. The PHO response in *C. glabrata* has undergone two such events at different levels of the signaling pathway. First, while the ortholog of *PHO2* is present in *C. glabrata*, *PHO2* deletion does not lead to an inability to activate phosphate responsive genes as it does in *S. cerevisiae* (Kerwin and Wykoff, 2009). This dispensability of *PHO2* for PHO activation is also seen in *Candida albicans* (Kerwin and Wykoff, 2009). Furthermore, despite extensive genetic analysis completed in *N. crassa*, the *PHO2* ortholog has never been implicated in the *N. crassa* PHO response (Metzenberg and Chia, 1979; Peleg et al., 1996). Second, *C. glabrata* does not possess an ortholog of *PHO5*, which resulted in the niche specialization of the phosphomutase-like domain protein, *PMU2* (Orkwis et al., 2010). *PMU2* displays a phosphatase activity and its expression is regulated in a phosphate- and *PHO4*-dependent manner. Apart from the examples mentioned above, little experimental work has been completed comparing the functional roles of the PHO components between different yeast species. Future experiments should reveal the extent of PHO pathway re-wiring due to speciation and the role different environmental niches have played in phosphate responsive signal transduction in various ascomycetes.

Phosphate Starvation in *S. pombe*: A Parallel PHO Pathway

S. pombe presents an interesting system for the study of the PHO pathway because the bulk of the PHO signaling components have been lost, yet *S. pombe* still senses and responds to inorganic phosphate starvation through some unknown mechanism. There are no clear orthologs of *PHO4*, *PHO2*, or *PHO81* in *S. pombe*, and the closest orthologs to *PHO80* and *PHO85* are not involved in the phosphate starvation response (Tanaka and Okayama, 2000). Despite this lack of clear PHO involvement, the *PHO5* ortholog, *pho1*⁺, remains phosphate responsive, with its expression induced 2-5 fold during phosphate limitation (Henry et al., 2011; Maundrell et al., 1985; Schwaninger et al., 1990). *pho1*⁺ is also regulated by adenine starvation, though its role in nucleotide metabolism is unclear (Schweingruber et al., 1992). Prior to the implementation of a deletion library screen for regulators of *pho1*⁺ activity (discussed below) only a few aspects of *pho1*⁺ regulation had been reported, and they were limited primarily to chromatin remodeling. The ATP-dependent nucleosome remodeling complex, Ino80-*lec1*, evicts histones from the *pho1*⁺ promoter in response to phosphate starvation, but it is unknown what factors activate Ino80-*lec1* during phosphate limitation (Hogan et al., 2010). Likewise, the SWI/SNF complex members *snf5*⁺ and *snf22*⁺ bind the *pho1*⁺ promoter and are required for *pho1*⁺ expression, but the mechanism controlling their response to phosphate starvation is unknown (Monahan et al., 2008).

Recently Henry et al. used a genome-wide set of gene deletions to identify genes required for normal *pho1*⁺ phosphate-responsive regulation in *S. pombe* (Henry et al., 2011; Kim et al., 2010). They describe eleven genes that produce an uninducible *pho1*⁺ phenotype upon deletion and five genes that produce a constitutive phenotype. The

uninducible collection was sorted based on function and two transcription factors are identified. The first, *snf5*⁺ has been discussed previously. The second, *pho7*⁺, is a putative transcription factor containing a Zn₂Cys₆ binuclear cluster (ZC) domain. The ZC domain is often associated with transcription factors, although that role has not been experimentally confirmed for *pho7*⁺ (MacPherson et al., 2006). Of the constitutive mutants, both *ado1*⁺ and *aps1*⁺ are orthologs to *S. cerevisiae* PHO signaling components (*ADO1* and *DDP1*, respectively). The other two genes in the constitutive category, *spcc1393.13*⁺ and *csk1*⁺, have orthologs in *S. cerevisiae* but the *S. cerevisiae* genes have no known role in the PHO pathway. The function of *spcc1393.13*⁺ in *S. pombe* is unknown, though the function of *csk1*⁺ has been well characterized.

csk1⁺ was first identified for its role in activating the CDK-cyclin complex Mcs6-Mcs2 (Hermand et al., 1998). CDK-activating kinases (CAKs) are responsible for phosphorylating the T-loop of their target CDKs, promoting the full activity of the kinase subunit. Consistent with this role, *csk1*⁺ phosphorylates the positive transcription elongation factor b (P-TEFb) ortholog Cdk9, increasing the kinase activity ~10 fold (Pei et al., 2006). Phosphorylation of Mcs6 and Cdk9 by Csk1 causes enhanced transcript elongation and more efficient mRNA capping/processing (Hermand et al., 1998; Pei et al., 2006). Unlike the CDK-cyclin complex responsible for PHO repression in *S. cerevisiae*, *csk1*⁺ functions without a cognate cyclin and operates solely as a CAK. How Csk1 causes inhibition of *pho1*⁺ transcription remains an open question.

From the deletion collection results, Henry et al. conclude that the PHO pathway in *S. pombe* has developed in parallel to the *S. cerevisiae* PHO pathway. The analogy to the *S. cerevisiae* system, where a transcription factor (*PHO4* or *pho7*⁺) activates PHO

responsive genes and a kinase complex (*PHO85-PHO80* or *csk1*⁺) is responsible for repressing the system in high-Pi conditions piqued our interest. This regulatory structure is the subject of our whole-genome characterization of the PHO pathway in *S. pombe*. In chapter 3 we characterize the PHO responsive genes in *S. pombe* that are regulated by phosphate starvation, *pho7*⁺, and *csk1*⁺. We demonstrate the global enrichment of Pho7 throughout the *S. pombe* genome, including regions with no known role in phosphate starvation. We elucidate the mechanism for *pho7*⁺ and *csk1*⁺-dependent regulation at the *pho1*⁺ promoter. Finally, using additional stress conditions (i.e. iron-starvation, copper-starvation, osmotic pressure, and carbon source switching) we demonstrate that, unlike *PHO4*, *pho7*⁺ has a broader role in additional stress response networks. The PHO pathway in *S. pombe* and the elements we explored in chapter 3 are summarized in Figure 1.2.

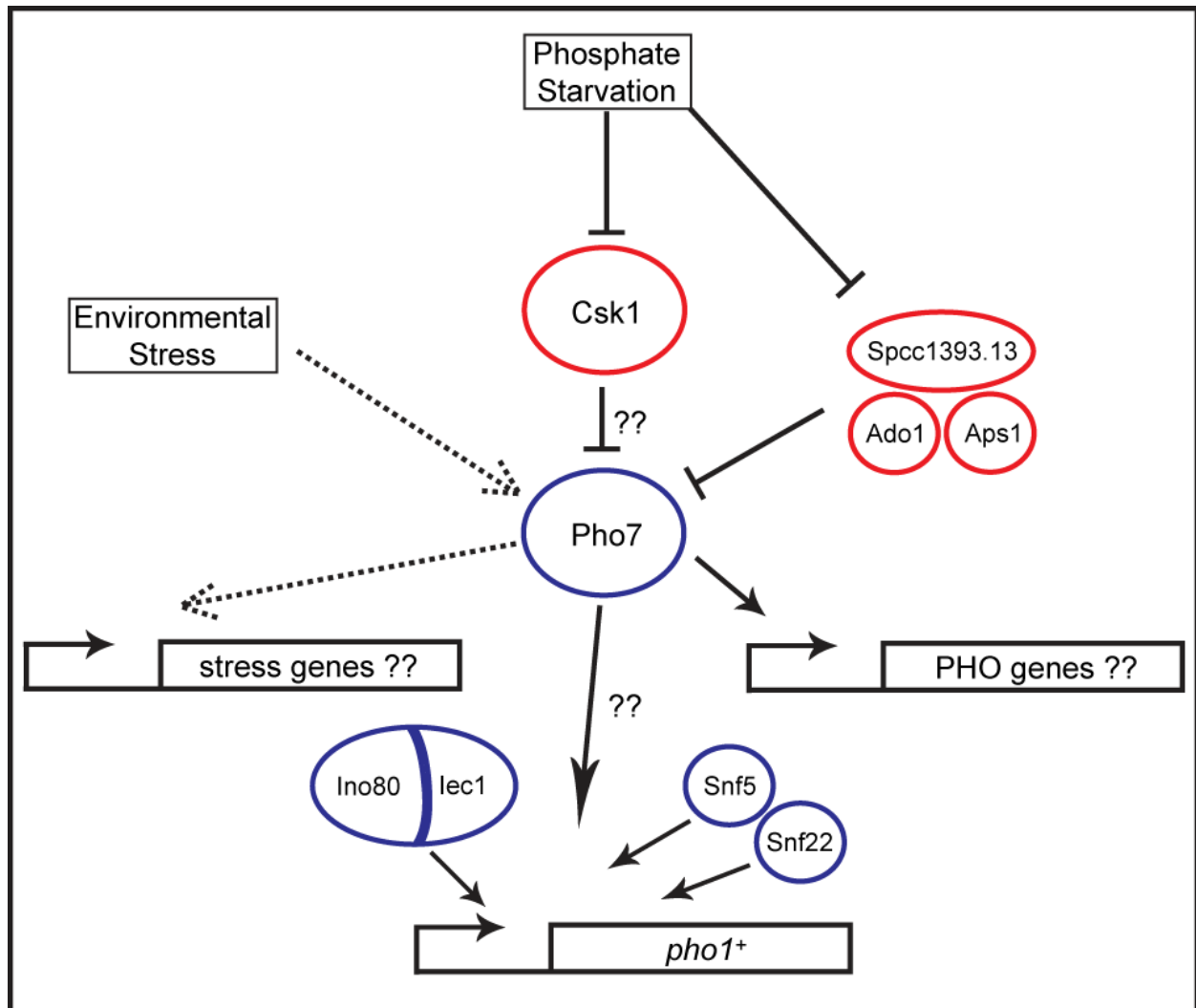


Figure 1.2. *Schizosaccharomyces pombe* PHO Pathway.

In high phosphate conditions Csk1, Ado1, Aps1, and Spcc1393.13 inhibit (red) the activity of the transcription factor Pho7 at the *pho1+* promoter. During phosphate starvation this inhibition is repressed and Pho7, Ino80-*lec1*, Snf5, and Snf22 activate (blue) *pho1+* transcription. Through a separate system, Pho7 is activated during iron starvation, copper starvation, osmotic pressure, and carbon source switching, up-regulating differential sets of stress response genes for each condition (Environmental Stress, dashed lines). Question marks represent genes and mechanisms that we identified and characterized in chapter 3.

REFERENCES

- Adams, P.D., Li, X., Sellers, W.R., Baker, K.B., Leng, X., Harper, J.W., Taya, Y., and Kaelin, W.G., Jr. (1999). Retinoblastoma protein contains a C-terminal motif that targets it for phosphorylation by cyclin-cdk complexes. *Mol Cell Biol* *19*, 1068-1080.
- Bagg, A., and Neilands, J.B. (1987). Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* *26*, 5471-5477.
- Berben, G., Legrain, M., Gilliquet, V., and Hilger, F. (1990). The yeast regulatory gene PHO4 encodes a helix-loop-helix motif. *Yeast* *6*, 451-454.
- Berben, G., Legrain, M., and Hilger, F. (1988). Studies on the structure, expression and function of the yeast regulatory gene PHO2. *Gene* *66*, 307-312.
- Bomeke, K., Pries, R., Korte, V., Scholz, E., Herzog, B., Schulze, F., and Braus, G.H. (2006). Yeast Gcn4p stabilization is initiated by the dissociation of the nuclear Pho85p/Pcl5p complex. *Mol Biol Cell* *17*, 2952-2962.
- Bostian, K.A., Lemire, J.M., Cannon, L.E., and Halvorson, H.O. (1980). In vitro synthesis of repressible yeast acid phosphatase: identification of multiple mRNAs and products. *Proc Natl Acad Sci U S A* *77*, 4504-4508.
- Brazas, R.M., and Stillman, D.J. (1993). The Swi5 zinc-finger and Grf10 homeodomain proteins bind DNA cooperatively at the yeast HO promoter. *Proc Natl Acad Sci U S A* *90*, 11237-11241.
- Bun-Ya, M., Nishimura, M., Harashima, S., and Oshima, Y. (1991). The PHO84 gene of *Saccharomyces cerevisiae* encodes an inorganic phosphate transporter. *Mol Cell Biol* *11*, 3229-3238.
- Byrne, M., Miller, N., Springer, M., and O'Shea, E.K. (2004). A distal, high-affinity binding site on the cyclin-CDK substrate Pho4 is important for its phosphorylation and regulation. *J Mol Biol* *335*, 57-70.
- Chang, Y.W., Robert Liu, F.G., Yu, N., Sung, H.M., Yang, P., Wang, D., Huang, C.J., Shih, M.C., and Li, W.H. (2008). Roles of cis- and trans-changes in the regulatory evolution of genes in the gluconeogenic pathway in yeast. *Mol Biol Evol* *25*, 1863-1875.
- Cohen, A., Perzov, N., Nelson, H., and Nelson, N. (1999). A novel family of yeast chaperons involved in the distribution of V-ATPase and other membrane proteins. *J Biol Chem* *274*, 26885-26893.
- Crosa, J.H. (1997). Signal transduction and transcriptional and posttranscriptional control of iron-regulated genes in bacteria. *Microbiol Mol Biol Rev* *61*, 319-336.

Daignan-Fornier, B., and Fink, G.R. (1992). Coregulation of purine and histidine biosynthesis by the transcriptional activators BAS1 and BAS2. *Proc Natl Acad Sci U S A* **89**, 6746-6750.

De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O., and Kim, S.H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595-602.

Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., De Montigny, J., Marck, C., Neueglise, C., Talla, E., *et al.* (2004). Genome evolution in yeasts. *Nature* **430**, 35-44.

Escolar, L., Perez-Martin, J., and de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* **181**, 6223-6229.

Espinoza, F.H., Ogas, J., Herskowitz, I., and Morgan, D.O. (1994). Cell cycle control by a complex of the cyclin HCS26 (PCL1) and the kinase PHO85. *Science* **266**, 1388-1391.

Ferminan, E., and Dominguez, A. (1997). The KIPHO5 gene encoding a repressible acid phosphatase in the yeast *Kluyveromyces lactis*: cloning, sequencing and transcriptional analysis of the gene, and purification and properties of the enzyme. *Microbiology* **143** (Pt 8), 2615-2625.

Friesen, H., Murphy, K., Breikreutz, A., Tyers, M., and Andrews, B. (2003). Regulation of the yeast amphiphysin homologue Rvs167p by phosphorylation. *Mol Biol Cell* **14**, 3027-3040.

Gasch, A.P. (2007). Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast* **24**, 961-976.

Ghillebert, R., Swinnen, E., De Snijder, P., Smets, B., and Winderickx, J. (2011). Differential roles for the low-affinity phosphate transporters Pho87 and Pho90 in *Saccharomyces cerevisiae*. *Biochem J* **434**, 243-251.

Gorner, W., Durchschlag, E., Wolf, J., Brown, E.L., Ammerer, G., Ruis, H., and Schuller, C. (2002). Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor. *EMBO J* **21**, 135-144.

Grana, X., and Reddy, E.P. (1995). Cell cycle control in mammalian cells: role of cyclins, cyclin dependent kinases (CDKs), growth suppressor genes and cyclin-dependent kinase inhibitors (CKIs). *Oncogene* **11**, 211-219.

Hannon, G.J., and Beach, D. (1994). p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* **371**, 257-261.

Hao, N., and O'Shea, E.K. (2011). Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nat Struct Mol Biol* **19**, 31-39.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., *et al.* (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99-104.

Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., and Elledge, S.J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**, 805-816.

Harper, J.W., and Elledge, S.J. (1996). Cdk inhibitors in development and cancer. *Curr Opin Genet Dev* **6**, 56-64.

Henry, T.C., Power, J.E., Kerwin, C.L., Mohammed, A., Weissman, J.S., Cameron, D.M., and Wykoff, D.D. (2011). Systematic screen of *Schizosaccharomyces pombe* deletion collection uncovers parallel evolution of the phosphate signal transduction pathway in yeasts. *Eukaryot Cell* **10**, 198-206.

Hernand, D., Pihlak, A., Westerling, T., Damagnez, V., Vandenhoute, J., Cottarel, G., and Makela, T.P. (1998). Fission yeast Csk1 is a CAK-activating kinase (CAKAK). *EMBO J* **17**, 7230-7238.

Hirai, H., Roussel, M.F., Kato, J.Y., Ashmun, R.A., and Sherr, C.J. (1995). Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Mol Cell Biol* **15**, 2672-2681.

Hogan, C.J., Aligianni, S., Durand-Dubief, M., Persson, J., Will, W.R., Webster, J., Wheeler, L., Mathews, C.K., Elderkin, S., Oxley, D., *et al.* (2010). Fission yeast Iec1-ino80-mediated nucleosome eviction regulates nucleotide and phosphate metabolism. *Mol Cell Biol* **30**, 657-674.

Holmes, J.K., and Solomon, M.J. (1996). A predictive scale for evaluating cyclin-dependent kinase substrates. A comparison of p34cdc2 and p33cdk2. *J Biol Chem* **271**, 25240-25246.

Huang, D., Friesen, H., and Andrews, B. (2007a). Pho85, a multifunctional cyclin-dependent protein kinase in budding yeast. *Mol Microbiol* **66**, 303-314.

Huang, D., Moffat, J., and Andrews, B. (2002). Dissection of a complex phenotype by functional genomics reveals roles for the yeast cyclin-dependent protein kinase Pho85 in stress adaptation and cell integrity. *Mol Cell Biol* **22**, 5076-5088.

Huang, D., Moffat, J., Wilson, W.A., Moore, L., Cheng, C., Roach, P.J., and Andrews, B. (1998). Cyclin partners determine Pho85 protein kinase substrate specificity in vitro and in vivo: control of glycogen biosynthesis by Pcl8 and Pcl10. *Mol Cell Biol* **18**, 3289-3299.

Huang, D., Patrick, G., Moffat, J., Tsai, L.H., and Andrews, B. (1999). Mammalian Cdk5 is a functional homologue of the budding yeast Pho85 cyclin-dependent protein kinase. *Proc Natl Acad Sci U S A* **96**, 14445-14450.

- Huang, K., Ferrin-O'Connell, I., Zhang, W., Leonard, G.A., O'Shea, E.K., and Quioco, F.A. (2007b). Structure of the Pho85-Pho80 CDK-cyclin complex of the phosphate-responsive signal transduction pathway. *Mol Cell* 28, 614-623.
- Huang, S., Jeffery, D.A., Anthony, M.D., and O'Shea, E.K. (2001). Functional analysis of the cyclin-dependent kinase inhibitor Pho81 identifies a novel inhibitory domain. *Mol Cell Biol* 21, 6695-6705.
- Huang, S., and O'Shea, E.K. (2005). A systematic high-throughput screen of a yeast deletion collection for mutants defective in PHO5 regulation. *Genetics* 169, 1859-1871.
- Hurlimann, H.C., Pinson, B., Stadler-Waibel, M., Zeeman, S.C., and Freimoser, F.M. (2009). The SPX domain of the yeast low-affinity phosphate transporter Pho90 regulates transport activity. *EMBO Rep* 10, 1003-1008.
- Jeffery, D.A., Springer, M., King, D.S., and O'Shea, E.K. (2001). Multi-site phosphorylation of Pho4 by the cyclin-CDK Pho80-Pho85 is semi-processive with site preference. *J Mol Biol* 306, 997-1010.
- Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N.P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376, 313-320.
- Jen, J., Harper, J.W., Bigner, S.H., Bigner, D.D., Papadopoulos, N., Markowitz, S., Willson, J.K., Kinzler, K.W., and Vogelstein, B. (1994). Deletion of p16 and p15 genes in brain tumors. *Cancer Res* 54, 6353-6358.
- Kaffman, A., Herskowitz, I., Tjian, R., and O'Shea, E.K. (1994). Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science* 263, 1153-1156.
- Kaffman, A., Rank, N.M., O'Neill, E.M., Huang, L.S., and O'Shea, E.K. (1998a). The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature* 396, 482-486.
- Kaffman, A., Rank, N.M., and O'Shea, E.K. (1998b). Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes Dev* 12, 2673-2683.
- Kerwin, C.L., and Wykoff, D.D. (2009). *Candida glabrata* PHO4 is necessary and sufficient for Pho2-independent transcription of phosphate starvation genes. *Genetics* 182, 471-479.
- Kim, D.U., Hayles, J., Kim, D., Wood, V., Park, H.O., Won, M., Yoo, H.S., Duhig, T., Nam, M., Palmer, G., *et al.* (2010). Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* 28, 617-623.

- Komeili, A., and O'Shea, E.K. (1999). Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science* 284, 977-980.
- Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. *Nature* 453, 246-250.
- Lee, Y.S., Huang, K., Quiocho, F.A., and O'Shea, E.K. (2008). Molecular basis of cyclin-CDK-CKI regulation by reversible binding of an inositol pyrophosphate. *Nat Chem Biol* 4, 25-32.
- Lee, Y.S., Mulugu, S., York, J.D., and O'Shea, E.K. (2007). Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science* 316, 109-112.
- Lemire, J.M., Willcocks, T., Halvorson, H.O., and Bostian, K.A. (1985). Regulation of repressible acid phosphatase gene transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* 5, 2131-2141.
- Lew, J., and Wang, J.H. (1995). Neuronal cdc2-like kinase. *Trends Biochem Sci* 20, 33-37.
- Maclsaac, K.D., Wang, T., Gordon, D.B., Gifford, D.K., Stormo, G.D., and Fraenkel, E. (2006). An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* 7, 113.
- MacPherson, S., Larochele, M., and Turcotte, B. (2006). A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiol Mol Biol Rev* 70, 583-604.
- Maerkl, S.J., and Quake, S.R. (2007). A systems approach to measuring the binding energy landscapes of transcription factors. *Science* 315, 233-237.
- Matsuoka, S., Edwards, M.C., Bai, C., Parker, S., Zhang, P., Baldini, A., Harper, J.W., and Elledge, S.J. (1995). p57KIP2, a structurally distinct member of the p21CIP1 Cdk inhibitor family, is a candidate tumor suppressor gene. *Genes Dev* 9, 650-662.
- Maundrell, K., Nurse, P., Schonholzer, F., and Schweingruber, M.E. (1985). Cloning and characterization of two genes restoring acid phosphatase activity in pho1- mutants of *Schizosaccharomyces pombe*. *Gene* 39, 223-230.
- Measday, V., Moore, L., Ogas, J., Tyers, M., and Andrews, B. (1994). The PCL2 (ORFD)-PHO85 cyclin-dependent kinase complex: a cell cycle regulator in yeast. *Science* 266, 1391-1395.
- Measday, V., Moore, L., Retnakaran, R., Lee, J., Donoviel, M., Neiman, A.M., and Andrews, B. (1997). A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. *Mol Cell Biol* 17, 1212-1223.

Meimoun, A., Holtzman, T., Weissman, Z., McBride, H.J., Stillman, D.J., Fink, G.R., and Kornitzer, D. (2000). Degradation of the transcription factor Gcn4 requires the kinase Pho85 and the SCF(CDC4) ubiquitin-ligase complex. *Mol Biol Cell* *11*, 915-927.

Metzenberg, R.L., and Chia, W. (1979). Genetic control of phosphorus assimilation in *Neurospora crassa*: dose-dependent dominance and recessiveness in constitutive mutants. *Genetics* *93*, 625-643.

Miller, M.E., and Cross, F.R. (2001). Cyclin specificity: how many wheels do you need on a unicycle? *J Cell Sci* *114*, 1811-1820.

Monahan, B.J., Villen, J., Marguerat, S., Bahler, J., Gygi, S.P., and Winston, F. (2008). Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nat Struct Mol Biol* *15*, 873-880.

Morgan, D.O. (1996). The dynamics of cyclin dependent kinase structure. *Curr Opin Cell Biol* *8*, 767-772.

Morgan, D.O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu Rev Cell Dev Biol* *13*, 261-291.

Nieuwenhuis, B.J., and Borst-Pauwels, G.W. (1984). Derepression of the high-affinity phosphate uptake in the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta* *770*, 40-46.

Nigg, E.A. (1995). Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *Bioessays* *17*, 471-480.

Nishizawa, M., Kanaya, Y., and Toh, E.A. (1999a). Mouse cyclin-dependent kinase (Cdk) 5 is a functional homologue of a yeast Cdk, pho85 kinase. *J Biol Chem* *274*, 33859-33862.

Nishizawa, M., Kawasumi, M., Fujino, M., and Toh-e, A. (1998). Phosphorylation of sic1, a cyclin-dependent kinase (Cdk) inhibitor, by Cdk including Pho85 kinase is required for its prompt degradation. *Mol Biol Cell* *9*, 2393-2405.

Nishizawa, M., Suzuki, K., Fujino, M., Oguchi, T., and Toh-e, A. (1999b). The Pho85 kinase, a member of the yeast cyclin-dependent kinase (Cdk) family, has a regulation mechanism different from Cdks functioning throughout the cell cycle. *Genes Cells* *4*, 627-642.

O'Neill, E.M., Kaffman, A., Jolly, E.R., and O'Shea, E.K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* *271*, 209-212.

Ogawa, N., DeRisi, J., and Brown, P.O. (2000). New components of a system for phosphate accumulation and polyphosphate metabolism in *Saccharomyces cerevisiae* revealed by genomic expression analysis. *Mol Biol Cell* *11*, 4309-4321.

- Ogawa, N., and Oshima, Y. (1990). Functional domains of a positive regulatory protein, PHO4, for transcriptional control of the phosphatase regulon in *Saccharomyces cerevisiae*. *Mol Cell Biol* 10, 2224-2236.
- Ogawa, N., Saitoh, H., Miura, K., Magbanua, J.P., Bun-ya, M., Harashima, S., and Oshima, Y. (1995). Structure and distribution of specific cis-elements for transcriptional regulation of PHO84 in *Saccharomyces cerevisiae*. *Mol Gen Genet* 249, 406-416.
- Orkwis, B.R., Davies, D.L., Kerwin, C.L., Sanglard, D., and Wykoff, D.D. (2010). Novel acid phosphatase in *Candida glabrata* suggests selective pressure and niche specialization in the phosphate signal transduction pathway. *Genetics* 186, 885-895.
- Parker, S.B., Eichele, G., Zhang, P., Rawls, A., Sands, A.T., Bradley, A., Olson, E.N., Harper, J.W., and Elledge, S.J. (1995). p53-independent expression of p21Cip1 in muscle and other terminally differentiating cells. *Science* 267, 1024-1027.
- Pei, Y., Du, H., Singer, J., Stamour, C., Granitto, S., Shuman, S., and Fisher, R.P. (2006). Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIH-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 in vivo. *Mol Cell Biol* 26, 777-788.
- Peleg, Y., Addison, R., Aramayo, R., and Metzenberg, R.L. (1996). Translocation of *Neurospora crassa* transcription factor NUC-1 into the nucleus is induced by phosphorus limitation. *Fungal Genet Biol* 20, 185-191.
- Peter, M., and Herskowitz, I. (1994). Joining the complex: cyclin-dependent kinase inhibitory proteins and the cell cycle. *Cell* 79, 181-184.
- Pinter, R.Y., Rokhlenko, O., Yeager-Lotem, E., and Ziv-Ukelson, M. (2005). Alignment of metabolic pathways. *Bioinformatics* 21, 3401-3408.
- Radzio-Andzelm, E., Lew, J., and Taylor, S. (1995). Bound to activate: conformational consequences of cyclin binding to CDK2. *Structure* 3, 1135-1141.
- Reddy, V.S., Singh, A.K., and Rajasekharan, R. (2008). The *Saccharomyces cerevisiae* PHM8 gene encodes a soluble magnesium-dependent lysophosphatidic acid phosphatase. *J Biol Chem* 283, 8846-8854.
- Reinders, A., Burckert, N., Boller, T., Wiemken, A., and De Virgilio, C. (1998). *Saccharomyces cerevisiae* cAMP-dependent protein kinase controls entry into stationary phase through the Rim15p protein kinase. *Genes Dev* 12, 2943-2955.
- Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J., and Pavletich, N.P. (1996a). Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature* 382, 325-331.

Russo, A.A., Jeffrey, P.D., and Pavletich, N.P. (1996b). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* 3, 696-700.

Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D., and Pavletich, N.P. (1998). Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* 395, 237-243.

Safrany, S.T., Ingram, S.W., Cartwright, J.L., Falck, J.R., McLennan, A.G., Barnes, L.D., and Shears, S.B. (1999). The diadenosine hexaphosphate hydrolases from *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* are homologues of the human diphosphoinositol polyphosphate phosphohydrolase. Overlapping substrate specificities in a MutT-type protein. *J Biol Chem* 274, 21735-21740.

Schneider, K.R., Smith, R.L., and O'Shea, E.K. (1994). Phosphate-regulated inactivation of the kinase PHO80-PHO85 by the CDK inhibitor PHO81. *Science* 266, 122-126.

Schulman, B.A., Lindstrom, D.L., and Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *Proc Natl Acad Sci U S A* 95, 10453-10458.

Schwaninger, R., Dumermuth, E., and Schweingruber, M.E. (1990). Effects of seven different mutations in the *pho1* gene on enzymatic activity, glycosylation and secretion of acid phosphatase in *Schizosaccharomyces pombe*. *Mol Gen Genet* 221, 403-410.

Schweingruber, M.E., Edenharter, E., Zurlinden, A., and Stockmaier, K.M. (1992). Regulation of *pho1*-encoded acid phosphatase of *Schizosaccharomyces pombe* by adenine and phosphate. *Curr Genet* 22, 289-292.

Serrano, M., Hannon, G.J., and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 366, 704-707.

Sethuraman, A., Rao, N.N., and Kornberg, A. (2001). The endopolyphosphatase gene: essential in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 98, 8542-8547.

Shao, D., Creasy, C.L., and Bergman, L.W. (1998). A cysteine residue in helixII of the bHLH domain is essential for homodimerization of the yeast transcription factor Pho4p. *Nucleic Acids Res* 26, 710-714.

Sherr, C.J. (1993). Mammalian G1 cyclins. *Cell* 73, 1059-1065.

Shimizu, T., Toumoto, A., Ihara, K., Shimizu, M., Kyogoku, Y., Ogawa, N., Oshima, Y., and Hakoshima, T. (1997). Crystal structure of PHO4 bHLH domain-DNA complex: flanking base recognition. *EMBO J* 16, 4689-4697.

Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnicka-Worms, H., and Cantley, L.C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol* 4, 973-982.

- Springer, M., Wykoff, D.D., Miller, N., and O'Shea, E.K. (2003). Partially phosphorylated Pho4 activates transcription of a subset of phosphate-responsive genes. *PLoS Biol* 1, E28.
- Srinivasan, J., Koszelak, M., Mendelow, M., Kwon, Y.G., and Lawrence, D.S. (1995). The design of peptide-based substrates for the cdc2 protein kinase. *Biochem J* 309 (Pt 3), 927-931.
- Svaren, J., Schmitz, J., and Horz, W. (1994). The transactivation domain of Pho4 is required for nucleosome disruption at the PHO5 promoter. *EMBO J* 13, 4856-4862.
- Takeda, D.Y., Wohlschlegel, J.A., and Dutta, A. (2001). A bipartite substrate recognition motif for cyclin-dependent kinases. *J Biol Chem* 276, 1993-1997.
- Tan, Y.S., Morcos, P.A., and Cannon, J.F. (2003). Pho85 phosphorylates the Glc7 protein phosphatase regulator Glc8 in vivo. *J Biol Chem* 278, 147-153.
- Tanaka, K., and Okayama, H. (2000). A pcl-like cyclin activates the Res2p-Cdc10p cell cycle "start" transcriptional factor complex in fission yeast. *Mol Biol Cell* 11, 2845-2862.
- Timblin, B.K., Tatchell, K., and Bergman, L.W. (1996). Deletion of the gene encoding the cyclin-dependent protein kinase Pho85 alters glycogen metabolism in *Saccharomyces cerevisiae*. *Genetics* 143, 57-66.
- To, E.A., Ueda, Y., Kakimoto, S.I., and Oshima, Y. (1973). Isolation and characterization of acid phosphatase mutants in *Saccharomyces cerevisiae*. *J Bacteriol* 113, 727-738.
- Toh-e, A., Tanaka, K., Uesono, Y., and Wickner, R.B. (1988). PHO85, a negative regulator of the PHO system, is a homolog of the protein kinase gene, CDC28, of *Saccharomyces cerevisiae*. *Mol Gen Genet* 214, 162-164.
- Toyoshima, H., and Hunter, T. (1994). p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. *Cell* 78, 67-74.
- Ueda, Y., To, E.A., and Oshima, Y. (1975). Isolation and characterization of recessive, constitutive mutations for repressible acid phosphatase synthesis in *Saccharomyces cerevisiae*. *J Bacteriol* 122, 911-922.
- Venter, U., Svaren, J., Schmitz, J., Schmid, A., and Horz, W. (1994). A nucleosome precludes binding of the transcription factor Pho4 in vivo to a critical target site in the PHO5 promoter. *EMBO J* 13, 4848-4855.
- Vogel, K., Horz, W., and Hinnen, A. (1989). The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Mol Cell Biol* 9, 2050-2057.

Wang, Z., Wilson, W.A., Fujino, M.A., and Roach, P.J. (2001). The yeast cyclins Pc16p and Pc17p are involved in the control of glycogen storage by the cyclin-dependent protein kinase Pho85p. *FEBS Lett* 506, 277-280.

Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature* 449, 54-61.

Wilson, W.A., Mahrenholz, A.M., and Roach, P.J. (1999). Substrate targeting of the yeast cyclin-dependent kinase Pho85p by the cyclin Pcl10p. *Mol Cell Biol* 19, 7020-7030.

Wilson, W.A., Wang, Z., and Roach, P.J. (2002). Analysis of respiratory mutants reveals new aspects of the control of glycogen accumulation by the cyclin-dependent protein kinase Pho85p. *FEBS Lett* 515, 104-108.

Wilson, W.A., Wang, Z., and Roach, P.J. (2005). Regulation of yeast glycogen phosphorylase by the cyclin-dependent protein kinase Pho85p. *Biochem Biophys Res Commun* 329, 161-167.

Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature* 430, 85-88.

Wykoff, D.D., and O'Shea, E.K. (2001). Phosphate transport and sensing in *Saccharomyces cerevisiae*. *Genetics* 159, 1491-1499.

Wykoff, D.D., Rizvi, A.H., Raser, J.M., Margolin, B., and O'Shea, E.K. (2007). Positive feedback regulates switching of phosphate transporters in *S. cerevisiae*. *Mol Cell* 27, 1005-1013.

Xiong, Y., Hannon, G.J., Zhang, H., Casso, D., Kobayashi, R., and Beach, D. (1993). p21 is a universal inhibitor of cyclin kinases. *Nature* 366, 701-704.

Yang, X., Safrany, S.T., and Shears, S.B. (1999). Site-directed mutagenesis of diphosphoinositol polyphosphate phosphohydrolase, a dual specificity NUDT enzyme that attacks diadenosine polyphosphates and diphosphoinositol polyphosphates. *J Biol Chem* 274, 35434-35440.

Zhou, X., and O'Shea, E.K. (2011). Integrated approaches reveal determinants of genome-wide binding and function of the transcription factor Pho4. *Mol Cell* 42, 826-836.

Zhu, C., Byers, K.J., McCord, R.P., Shi, Z., Berger, M.F., Newburger, D.E., Saulrieta, K., Smith, Z., Shah, M.V., Radhakrishnan, M., *et al.* (2009). High-resolution DNA-binding specificity analysis of yeast transcription factors. *Genome Res* 19, 556-566.

CHAPTER 2

Structure of the Pho85-Pho80 CDK-cyclin Complex of the Phosphate-responsive Signal Transduction Pathway*

Kexin Huang, Ian Carter-O'Connell, Wei Zhang, Gordon A. Leonard, Erin K. O'Shea,
and Florante A. Quioco

*This chapter is reprinted (with the exception of this page, footnotes, and changes to figure numbering to comply with FAS requirements) from the following:

Kexin, H., Ferrin-O'Connell, I., Zhang, W., Leonard, GA., O'Shea, EK, and Quioco, FA. (2007). Structure of the Pho85-Pho80 CDK-cyclin Complex of the Phosphate-responsive Signal Transduction Pathway. *Molecular Cell*. 28(4):614-623.

Reprinted by permission of the publisher and all co-authors.

AUTHOR CONTRIBUTIONS

The manuscript was written by KH, ICO, EKO and FAQ. Crystallography and molecular modeling was performed by KH with assistance from WZ and GAL. Potential sites for Pho4 interaction with Pho85-Pho80 were identified by KH and FAQ with assistance from ICO and EKO. Mutagenesis and kinetic characterization was performed by ICO. EKO and FAQ provided supervision throughout and FAQ is the corresponding author.

Structure of the Pho85-Pho80 CDK-cyclin Complex of the Phosphate-responsive Signal Transduction Pathway

Kexin Huang,¹ Ian Ferrin-O'Connell,² Wei Zhang,¹ Gordon A. Leonard,³ Erin K. O'Shea,²
and Florante A. Quijcho¹

¹Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030.

²Howard Hughes Medical Institute, Faculty of Arts and Sciences Center for Systems Biology, Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138.

³Macromolecular Crystallography Group, European Synchrotron Radiation Facility, BP 220, 38043 Grenoble Cedex, France.

ABSTRACT

The ability to sense and respond appropriately to environmental changes is a primary requirement of all living organisms. In response to phosphate limitation, *Saccharomyces cerevisiae* induces transcription of a set of genes involved in the regulation of phosphate acquisition from the ambient environment. A signal transduction pathway (the PHO pathway) mediates this response, with Pho85-Pho80 playing a vital role. Here we report the x-ray structure of Pho85-Pho80, the first for CDK-cyclin complex functioning in transcriptional regulation in response to environmental changes. The structure revealed a specific salt link between a Pho85 arginine and a Pho80 aspartate that makes phosphorylation of the Pho85 activation loop dispensable and that maintains a Pho80 loop conformation for possible substrate recognition. It further showed two new sites on the Pho80 cyclin for high affinity binding of the transcription factor substrate (Pho4) and the CDK inhibitor (Pho81) that are markedly distant to each other and the active site.

INTRODUCTION

Inorganic phosphate is an essential nutrient for all organisms, as it is required for the biosynthesis of diverse cellular components including nucleic acids, proteins, lipids, sugars and phospho-metabolites. The budding yeast *Saccharomyces cerevisiae* phosphate-responsive signaling system (known as the PHO pathway) senses and responds to changes in the concentration of inorganic phosphate in the medium [(Toh-e et al., 1973; Ueda et al., 1975); reviewed in (Carroll and O'Shea, 2002)]. Through this pathway, many PHO genes are repressed in high-phosphate conditions and induced in conditions of phosphate limitation. Central to the PHO pathway is a CDK-cyclin complex, Pho85-Pho80, whose activity is regulated in response to extracellular phosphate availability (Kaffman et al., 1994; Schneider et al., 1994; Toh-e et al., 1988). Pho81, a CDK inhibitor (CKI), binds to Pho85-Pho80 when cells are grown in both high- and no-phosphate conditions, but inhibits the kinase activity only during phosphate limitation (Schneider et al., 1994). The Pho85-Pho80-Pho81 complex regulates the location and activity of Pho4 (Kaffman et al., 1994), a transcription factor required for expression of phosphate-responsive genes, including PHO5, which encodes a secreted phosphatase. In high-phosphate medium, Pho85-Pho80 phosphorylates and inactivates Pho4. In medium devoid of phosphate, Pho81 inhibits Pho85-Pho80, leading to accumulation of unphosphorylated Pho4 in the nucleus and activation of PHO5 transcription.

Pho85, through its association with nine other Pho85 cyclins (called Pcls) (Measday et al., 1997), is one of the most versatile CDKs. Pcls target Pho85 to different substrates and thus other cellular functions (Carroll and O'Shea, 2002; Toh-e and

Nishizawa, 2001), but the structural basis for substrate targeting is unclear. Of the diverse cellular functions regulated by Pho85, the PHO pathway is by far the best studied.

Despite significant similarity between Pho85 and the cell cycle CDKs, especially Cdc28/CDK2 (Toh-e et al., 1988), Pho85 possesses several prominent distinct features. Whereas phosphorylation of a conserved threonine or serine residue on the kinase subunit activation loop is required for full activation of CDK-cyclin complexes functioning in cell cycle [reviewed in Morgan, 1996; Russo et al., 1996b]), it is dispensable for Pho85-Pho80 kinase function (Nishizawa et al., 1999). The residue at the +3 position of the consensus sequence (S/TPXK/R, where S/T is the phosphorylatable residue and X is any residue) of the substrates of most cell cycle CDK-cyclin complexes differs radically from that (SPXI/L) of the five phosphorylation sites on the Pho4 substrate of Pho85-Pho80 (O'Neill et al., 1996). Moreover, tight interaction between Pho80 and a site distal to the phosphorylation sites in Pho4 enhances catalytic efficiency by orders of magnitude and enables semi-processive phosphorylation (Byrne et al., 2004; Jeffery et al., 2001). The inhibitory domain of the Pho81 CKI differs from those of the two major types of mammalian CKIs, the INK4s and Cip/Kips, of cell cycle regulation (Huang et al., 2001). In addition, unlike CKIs of the cell cycle CDK-cyclin complexes which either target the kinase solely or both kinase and cyclin [reviewed in (Endicott et al., 1999)], Pho81 interacts with Pho85-Pho80 primarily through association with the Pho80 subunit (Schneider et al., 1994). Interestingly, Pho81 has the unusual property of forming a stable complex with Pho85-Pho80 under both high- and low-phosphate concentrations, but only inhibiting under low phosphate conditions (Schneider et al., 1994). Recently, it

has been reported that kinase inhibition by the constitutively associated Pho81 requires myo-D-heptakisphosphate (IP7) (Lee et al., 2007). To gain new insights into the molecular nature of the distinct features of the Pho85 in conjunction with Pho80 cyclin, as well as the role of the complex in the PHO pathway, we carried out the determination of the crystal structure of the Pho85-Pho80 complex.

RESULTS AND DISCUSSION

Structure of the Pho85-Pho80 Complex

Crystals of the complex of full-length Pho85 (305 residues) and Pho80 (293 residues) were obtained by the hanging drop method and optimized by the addition of strontium to the precipitant and incubation at 10°C (see Supplemental Data). The structure, with two Pho85-Pho80 complexes in the asymmetric unit and an unusually high water content (~80%), was determined by single wavelength anomalous dispersion (SAD) and refined to 2.9 Å resolution (Table 2.1). The structure was used subsequently to determine, at an identical resolution, the structure of the isomorphous co-crystal of Pho85-Pho80 and the nonhydrolyzable ATP analog, ATP-γ-S (Table 2.1). The co-crystal structure is the focus of this report.

Table 2.1. Data Collection and Refinement Statistics

	SeMet SAD ^a	Native ^b	Native with bound ATP- γ -S ^b
Diffraction Data			
Space group	P3 ₁ 21	P3 ₁ 21	P3 ₁ 21
Cell dimensions	a=b=147.1 Å, c=212.4 Å $\alpha=\beta=90^\circ$; $\gamma=120^\circ$	a=b=147.8 Å, c=212.8 Å $\alpha=\beta=90^\circ$; $\gamma=120^\circ$	a=b=146.6 Å, c=212.8 Å $\alpha=\beta=90^\circ$; $\gamma=120^\circ$
Wavelength	0.97926 peak	0.97948	1.0332
Resolution (Å)	30 – 3.0	15 – 2.9	12 – 2.9
Total reflections	32,066	52,156	53,149
R _{sym} ^{c, d} (%)	12.2 (53.5)	9.6 (29.0)	9.9 (24.8)
I/ σ I	9.6 (2.3)	12.2 (2.4)	11.0 (2.4)
Completeness ^d (%)	99.9 (99.9)	88.4 (71.2)	91.1 (75.6)
Redundancy ^d	4.6 (4.6)	3.2 (2.2)	3.1 (2.0)
Refinement			
Resolution (Å)		15 – 2.9	12 – 2.9
Total reflections		52,156	53,149
R _{crys} ^e /R _{free}		0.281/0.315	0.287/0.325
Rmsd			
Bond length (Å)		0.008	0.007
Bond angles (°)		1.318	1.32
^a Data collected at beamline ID29, ESRF (1° oscillation) and processed with MOSFLM (Leslie, 1992).			
^b Data collected at beamline ID19, APS (1° oscillation) and processed with HKL2000 (Otwinowski and Minor, 1997).			
^c R _{sym} = $\sum_h \sum_l I(h) - \langle I(h) \rangle / \sum_h \sum_l I(h)$.			
^d Numbers in parenthesis are for outer shell data.			
^e R _{crys} = $\sum (F_o - k F_c) / \sum F_o $. R _{free} was calculated using a random 5% of the reflection data that was omitted in the refinement.			

Pho85 has a typical CDK fold, consisting of two distinct lobes – a smaller N-terminal (or N) lobe composed mainly of a five-stranded antiparallel β sheet and the PSTAIRE helix and a much larger C-terminal (or C) lobe rich in α helices (Figure 2.1A). The structure is not unexpected given amino acid sequence identity as high as ~55% between Pho85 and CDK2, the mammalian cell cycle regulator (Morgan, 1997), and CDK5, an important signal transducer in neuronal development (Smith et al., 2001) and the closest known functional homolog of yeast Pho85 (Huang et al., 1999; Nishizawa et al., 1999). The ATP analog ATP- γ -S is bound in the cleft between the N and C lobes of Pho85 (Figure 2.1A) and held in place by nearly identical residues that bind the AMPPNP in the structure of the CDK2-cyclin A-peptide substrate-AMPPNP-Mg²⁺ complex (Brown et al., 1999). The cleft further harbors the catalytic center, which

contains the conserved catalytic triad of residues in CDKs (Lys33, Glu51 and Asp145 in CDK2 and Lys36, Glu53 and Asp151 in Pho85).

Pho80 has a single globular domain structure composed of eight helices, identified as α NT for the first helix, α 1 to α 5 for the core helix bundle and α CT1 and α CT2 for the last two helices that cap one end of the helix bundle and bring the C-terminal loop close to the N-terminal loop (Figures 2.1A and B). Pho80 contains several unusually long meandering loops, including those at both termini and between helices α NT and α 1, and α CT1 and α CT2 (Figure 2.1A and Figure S2.1 in Supplemental Data). Furthermore, unlike the cyclins functioning in the cell cycle [e.g., the structure of the 259-residue fragment of vertebrate cyclin A bound to CDK2 (Jeffrey et al., 1995)], which contain tandem duplications of five-helix bundles (also known as cyclin boxes), the Pho80 structure shows only one cyclin box, which corresponds to the core helices α 1 to α 5 (Figures 2.1A and B). This structural result solidifies previous sequence comparison indicating one cyclin box for Pho80 and all other nine Pcls (Moffat et al., 2000) (see also Supplementary Figure S2.1). p25, the 147-residue fragment of the p35 activator of CDK5, is to our knowledge the only other cyclin whose crystal structure (Tarricone et al., 2001) shows a single cyclin box (minus the missing α 4), as well as the presence of α NT helix and two additional C-terminal helices at locations quite close to those in Pho80 (Figure 2.1B).

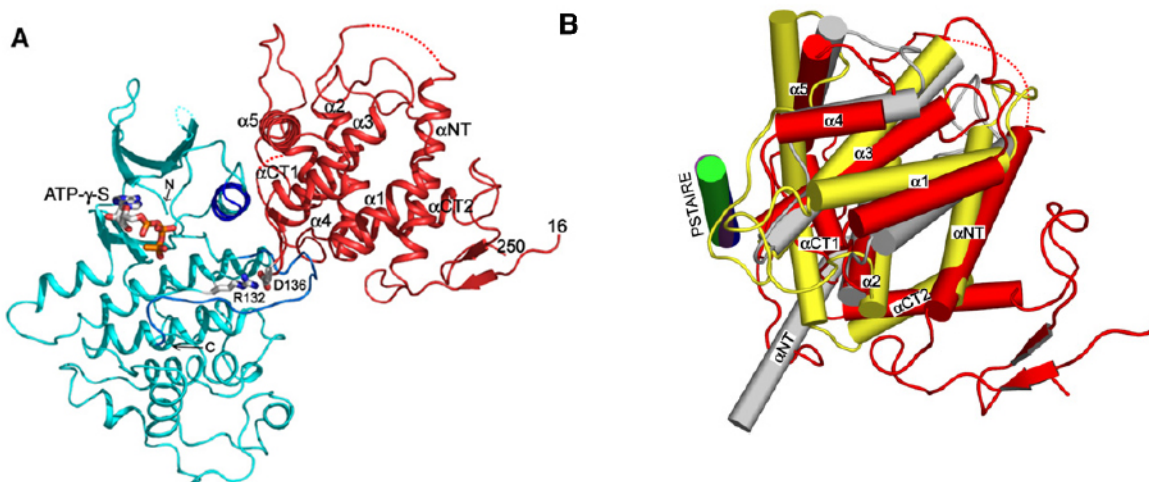


Figure 2.1. Structure of Pho85-Pho80

(A) Ribbon backbone trace of the crystal structure of the Pho85-Pho80 complex with the bound ATP analog, ATP- γ -S. Pho85 is shown in cyan, and its PSTAIRE helix and activation loop in blue and marine, respectively. Pho80 is depicted in red. The ATP- γ -S is represented in ball-and-stick model. The amino- and carboxy-terminal ends of Pho85 are marked by N and C, respectively, and those of Pho80 are denoted by residue numbers 16 and 250, respectively. The first 15 and last 43 residues in Pho80 are not visible (or disordered) in the electron density map, indicating their flexibilities. The long loops in Pho80 include the visible N-terminal 17 and C-terminal 22 residue loops, the 25 residue loop between α NT and α 1, and the 33 residue loop between α CT1 and α CT2 helices (see Figure S2.1). The two dotted lines in the Pho80 trace represent disordered segments (see Figure S2.1). R132 on Pho85 and D136 on Pho80 form a salt link (details shown in Figure 2.2). (B) Relative positions of Pho80 (red), cyclin A (gray), and p25 (yellow) following superposition of their cognate kinase subunits as they occur in the structures of the Pho85-Pho80-ATP- γ -S complex, the phosphorylated CDK2-cyclin A complex (Brown et al., 1999), and CDK5-p25 (Tarricone et al., 2001). The orientation of Pho80 is identical to that shown in (A). The depiction of the α helices by cylinders emphasizes the differences between the relative orientations of the cyclins. p25 lacks the equivalent of the two-turn α 4 helices in Pho80 and cyclin A, and in its place is a turn that includes the P247 residue. The PSTAIRE (Pho85 and CDK2) or PSSALRE (p25) helices are shown in blue (Pho85), green (CDK2), and purple (CDK5). Although the PSTAIRE/PSSALRE helices are completely superimposable, the contacts they make with their cognate cyclins vary (e.g., see Figures S2.2 and S2.3). In the truncated cyclin A structure, only the first helix (α NT) and the first cyclin box are shown. The succeeding two helices (α 10 and α 20) in cyclin A (not shown) constitute the first two helices of the second cyclin box and adopt positions different from those of α CT1 and α CT2, respectively, in Pho80 or p25. The α NT helix of cyclin A makes significant interaction with the C-terminal lobe of CDK2 (Russo et al., 1996b). In contrast, the α NT helices of Pho80 and p25, which approximately adopt similar positions, do not participate in kinase binding because they lie on the opposite side of the cyclin structures to the kinase-cyclin interface or the location of α NT in cyclin A.

However, superposition of the Pho80 and p25 cyclin structures indicates only 84 overlapped α -carbons with rmsd of 2.0 Å, of which a significant number (59 α -carbons) are confined to α 3, α 4 and α 5 helices that contact the region adjacent to the active site cleft of the cognate kinase subunits. As in the case of many other cyclins, Pho80 and

p25, as well as the Pcls (Supplementary Figure S2.1), show no significant sequence similarity.

Superposition of the CDK subunits as they occur in the structure of the fully activated, phosphorylated CDK2-cyclin A complex (Brown et al., 1999) and those of the Pho85-Pho80-ATP- γ -S and CDK5-p25 (Tarricone et al., 2001) complexes, indicates no gross changes in the kinase conformation. Moreover, the superposition indicates that similar regions, mainly the PSTAIRE (Pho85 and CDK2)/PSSALRE (CDK5) helix and its preceding loop and the activation (or T-) loop, are engaged in major interactions with the regulatory partners. The superposition also clearly shows the differences in the relative dispositions of the equivalent helices and in the lengths and conformations of the inter-helix connecting loops in the cyclin subunits (Figure 2.1B). Consequently, the involvements of the helices and loops of the cyclins at the interface with their cognate kinases vary (for example, see Figures S2.2 and S2.3 in Supplemental Data). These variations, indicative of the plasticity of the CDK-cyclin interfaces, are underscored by comparing the three different complex structures in the following framework. In the Pho85-Pho80 complex, the cyclin helices (mostly the set of helices α 3, α 4 and α 5) are involved in slightly fewer contacts (~45%) at the interface than the loop regions (~55%) located mainly between α 3 and α 4, α 5 and α CT1 and α CT1 and α CT2. In contrast, in the CDK5-p25 complex, the cyclin makes significantly more contacts (~75%) via the same set of helices and α 6 than via principally one loop that follows α 3. The CDK2-cyclin A interface portrays the extreme case; the distribution is ~85% helices, contributed mostly by α 3, α 5 and α NT and ~15% loops, provided almost entirely by the one between α 5 and α 1' helices, the first helix of the second cyclin box.

Discovery of a Specific Mechanism for Circumventing Activation Loop

Phosphorylation

Despite the difference in the juxtapositions of Pho80 and p25 with their kinase partners, the two complexes share one key functional feature that is distinct from the CDK-cyclin complexes involved in cell-cycle regulation. Whereas phosphorylation of the T/S residue on the T loop (e.g., T160 of CDK2) by the CDK activating enzyme is obligatory for full activation of the cell-cycle CDK-cyclin complexes, it is not required for Pho85-Pho80 and CDK5-p35 despite the presence of an equivalent serine residue on their T loops (S166 on Pho85 and S159 on CDK5) (Espinoza et al., 1998; Nishizawa et al., 1999; Poon et al., 1997). Phosphorylation of the T loop T160 residue of CDK2-cyclin A causes a large conformational rearrangement that results in the opening of the substrate site and proper formation of the entire active-site region (reviewed in Pavletich, 1999 and Russo et al., 1996b). Moreover, the presence of the phosphoryl adduct dictates the strong preference for the positively charged residue at the +3 position of the consensus sequence (SPXK/R) of substrates of cell-cycle CDK-cyclin complexes. The crystal structure of phosphorylated CDK2-cyclin A in the presence of a peptide substrate containing the consensus sequence SPRK, AMPPNP, and Mg^{2+} illustrates that the structural basis for this substrate specificity is a salt link formed between the phosphoryl adduct and the +3 positively charged residue (Brown et al., 1999). The dispensability of the T loop phosphorylation in the Pho85-Pho80 complex and the strong preference for large aliphatic side chains (I/L) at the +3 position of the consensus sequence of the five phosphorylation sites on the Pho4 substrate indicate

that Pho80 binding is sufficient for full activation of the kinase activity and that a different strategy is used for the recognition of the residue at the +3 position.

Our Pho85-Pho80 structure analysis has led to the discovery of a charge-coupling interaction between the kinase and the cyclin subunits that makes activation loop phosphorylation dispensable. The interaction is between the guanidinium group of R132 on the C lobe of Pho85 and the carboxylate group of D136 on the 6 residue α 3- α 4 loop (named D loop) of Pho80 (Figures 2.1A and 2.2).

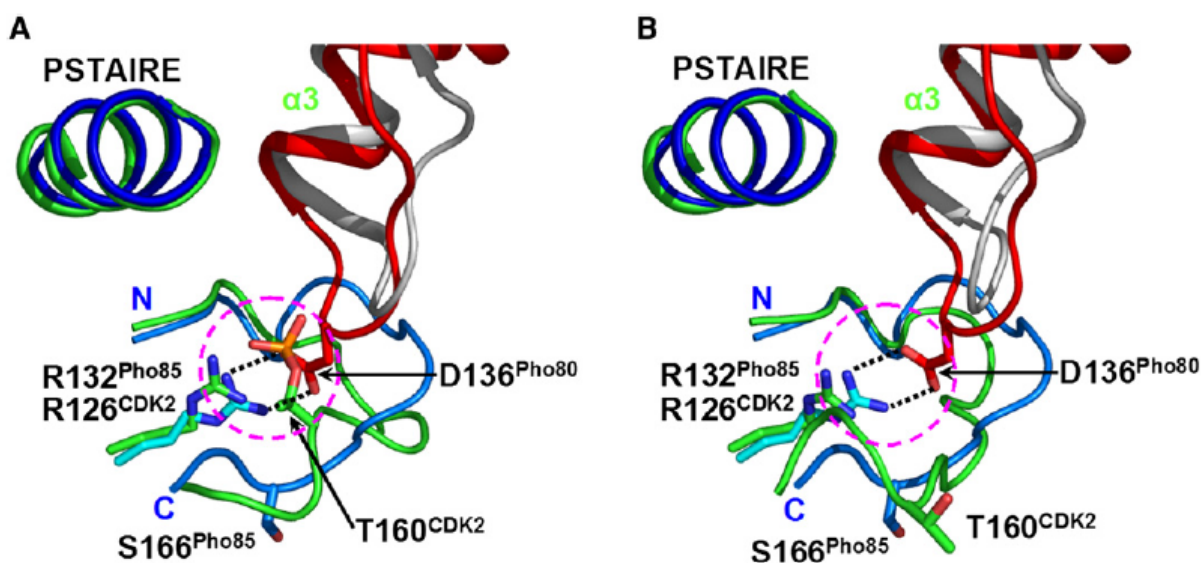


Figure 2.2. A Salt Link between Pho85 R132 and Pho80 D136 that Makes T Loop Phosphorylation Dispensable

(A) Relative positions of the loop between α 3 and α 4 in Pho80 (red) and cyclin A (gray) and the PSTAIRE helices and segments of the activation loops in Pho85 (residues 155–168) (marine) and phosphorylated (T160) CDK2 (residues 149–163) (green). The overlay is based on the superposition of the structures of the kinase subunits as they occur in the structures of the Pho85-Pho80-ATP- γ -S complex, and the phosphorylated CDK2-cyclin A complex (Brown et al., 1999). The α 3- α 4 loop in Pho80 deploys D136, hence the “D loop” name. The activation loop T160 (with the phosphoryl adduct) of CDK2 corresponds to S166 in Pho85. Pho85 R132 and CDK2 R126 are superimposable. CDK5 R125 adopts an almost identical position as those of Pho85 R132 and CDK2 R126 (data not shown). Pho80 D136 has no counterpart in cyclin A and p25. A more global view of the D loop and the salt link is shown in Figure 2.1A. (B) Similar superposition as shown in (A), except that the CDK2 structure is the unphosphorylated form (Jeffrey et al., 1995). The position of R126 of CDK2 is essentially unchanged from that in the phosphorylated CDK2 (A). Note the significant change in the conformation of the segment of the T loop of CDK2 following phosphorylation (compare with that in [A]).

The charge-coupling interaction matches closely that occurring in the phosphorylated CDK2-cyclin A complex, between the guanidinium group of CDK2 R126, the equivalent

of Pho85 R132, and the phosphoryl adduct to the T loop T160 residue (Figure 2.2A). Thus, the R132-D136 salt link makes T loop phosphorylation of Pho85 no longer necessary for full kinase activation. The fact that the R126 in the phosphorylated CDK2 structure remains in an identical position in the unphosphorylated kinase subunit (Figure 2.2B) indicates its major role in dictating the ultimate position of the phosphoryl group required for full kinase activation and substrate recognition. In contrast, if the T loop S166 of Pho85 were phosphorylatable, the presence of Pho80 D136 would prevent the phosphoryl group from assuming a position similar to that of the phosphoryl group on T160 of CDK2. Support for the importance of the R132-D136 salt link is provided by the observation of loss of Pho80 function following very conservative mutation of D136 to N (Madden et al., 1990). Moreover, the complete conservation of Pho80 D136 in all of the other nine Pcls (see Figure S2.1) highlights the arginine-aspartate salt link as a common feature of a large family of CDK-cyclin complexes essential in the regulation of diverse cellular functions.

Potential Involvement of the D Loop in Substrate Recognition

The CDK5-p25 complex crystal structure offered the first detailed view of a phosphorylation-independent, fully active CDK-cyclin complex (Tarricone et al., 2001). Moreover, a modeling study, followed by mutational analysis, has indicated the importance of the carboxylate side chain of E240, located close to the beginning of the loop that follows the α 3 helix of p25, in the recognition of the positively charged residue at the +3 position of the substrate consensus sequence SPXK/R (Tarricone et al., 2001). Interestingly, F138, the sole hydrophobic residue on the D (or α 3- α 4) loop in Pho80 (Figure 2.3A and Figure S2.1), is in a position to mimic the role of the p25 E240,

in recognizing the I/L at the +3 position of the consensus sequence (SPXI/L) of the five phosphorylation sites on the Pho4 substrate of Pho85-Pho80 (O'Neill et al., 1996).

Figure 2.3. Three Functional Sites of the Pho80 Cyclin Subunit

(A) Ribbon trace of Pho80. The regions of the three functional sites are identified as follows (see also text): the bound ATP- γ -S (stick representation) marking the location of the kinase active-site region; the cluster of residues C30, L38, R41, M42, and G229 and segments of residues 17–24 and 243–246 forming a large cavity for possible tight interaction with the site on Pho4 distal to the phosphorylatable sites; and the salt link between R121 and E154 in a region for docking the inhibitor, Pho81. The peptide near the ATP- γ -S corresponds to a model of the the SPRL consensus sequence of the five Pho4 phosphorylation sites (O'Neill et al., 1996). Placement of the peptide was derived by superimposing the atomic coordinates of only the Pho85 and CDK2 subunits in the Pho85-Pho80-ATP- γ -S structure and the structure of the phosphorylated CDK2-cyclin A with a bound peptide containing the SPRK consensus sequence and AMPPNP-Mg²⁺ (Brown et al., 1999). Changing the K at the +3 position to L to conform to the Pho4 consensus sequence, SPXI/L, indicates that the L side chain would be within van der Waals distance to the F138 located on the α 3- α 4 (or D) loop in Pho80. (B) Electrostatic surface potential surface of Pho80 contoured at -10 kT (red) and +10 kT (blue) calculated with GRASP (Nicholls et al., 1991). Pho80 is in identical orientation as in (A). The region involved in binding of the CKI Pho81 is enclosed by red dashed lines. The region involved in docking of the site on Pho4 distal to the phosphorylation sites is enclosed in magenta dashed lines. (C) Hydrophobic patch on cyclin A for docking the RXL motif of substrates or CKIs. The bound ligand, with backbone trace in green, shows only the RRL sequence (the equivalent of the RXL motif) of the substrate recruitment peptide RRLFGE of p107 (Brown et al., 1999). The residues M210, I213, and E220 on α 1 and R250 on α 3 define the specificity of the patch in a group of cell-cycle cyclins for the RXL motif (Endicott et al., 1999). (D) Identical to (C), but with the surface of cyclin A removed for clarity, and the α 1 and α 3 helices and their preceding loops of Pho80 superimposed. The first two turns of α 1 of cyclin A, which provides M210 and I213, are missing in the corresponding α 1 of Pho80.

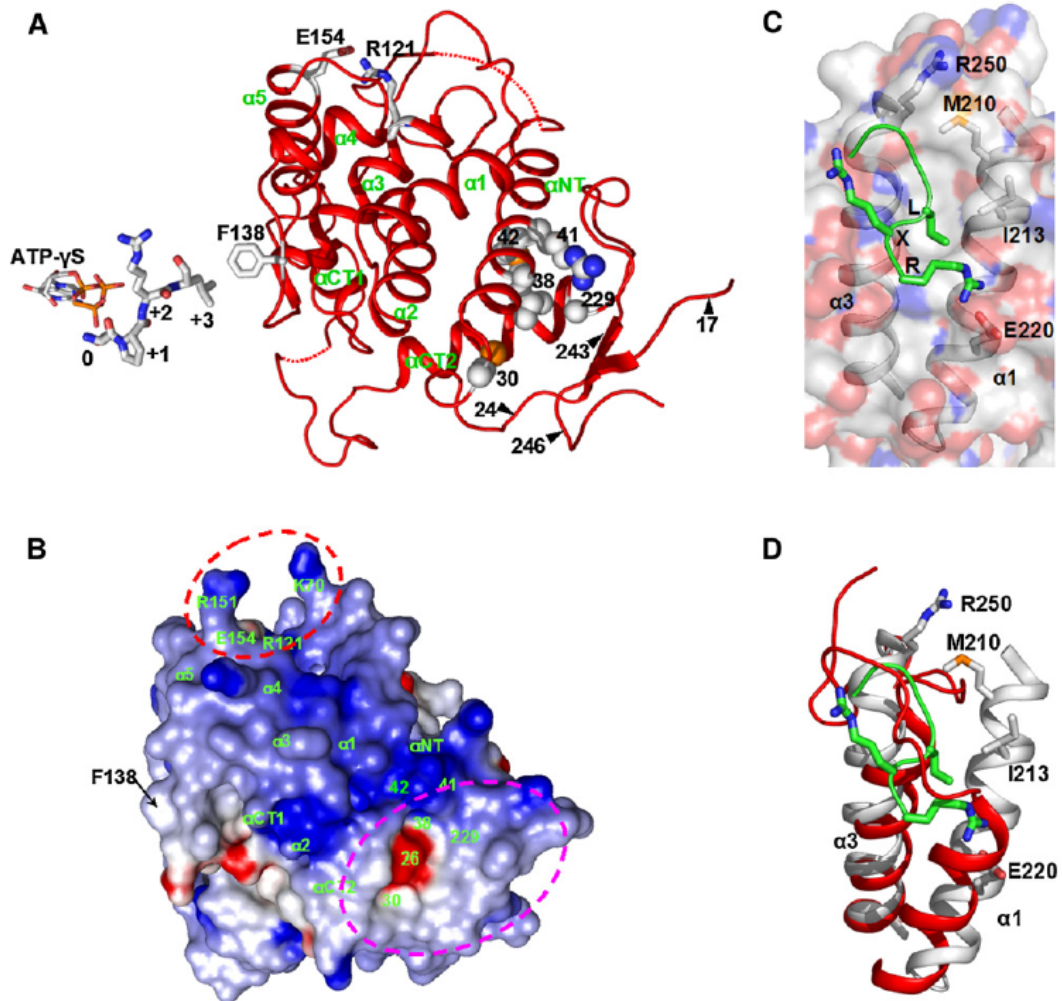


Figure 2.3. Continued

To assess the role of Pho80 F138 in substrate phosphorylation, we analyzed the effects of substitutions of this residue (Table 2.2). The Pho80 F138A mutation caused a 3.5- and 7.8-fold decrease in k_{cat}/K_M for phosphorylation of the natural substrate, Pho4, and the specific 24 residue peptide substrate (called SPVI), respectively. These results, together with an even greater (15.7-fold) reduction in the phosphorylation of Pho4 caused by the mutation F138E, suggest a major role for Pho80 F138 in attaining full kinase activity. The approximately three orders of magnitude greater catalytic efficiency of wild-type Pho85-Pho80 toward Pho4 compared to SPVI has been previously observed (Jeffery et al., 2001) and is attributed to the interaction between a site in Pho4 distal to the phosphorylation sites and a site in the Pho80 cyclin (Byrne et al., 2004; discussed further below). The 9.8-fold decrease in activity toward SPVA, a nonspecific peptide substrate, is consistent with the strong preference for the large aliphatic side chain at +3 position of the consensus sequence of the Pho4 phosphorylation sites. These observations suggest that F138 of Pho80 and position +3 of the substrate are important for efficient phosphorylation. If F138 interacts favorably with I in the +3 position of the substrate and the only effect of F138A is to prevent this interaction, we expect that the F138A mutation will decrease phosphorylation of the SPVI substrate, but that SPVA will be phosphorylated with efficiency comparable to the wild-type Pho80. Unexpectedly, the decrease in phosphorylation resulting from the F138A mutation is similar for both SPVI and SPVA substrates (14.9- versus 11.1-fold), suggesting that mutation of F138 may have additional effects on the structure of the kinase which cause a general reduction in catalytic efficiency.

Table 2.2 Probing the Role of F138 in Pho80 in Substrate Recognition

Substrate	k_{cat}/K_M ($M^{-1} s^{-1}$) ^a		
	Pho85-Pho80	Pho85-Pho80 F138A	Pho85-Pho80 F138E
Pho4	$6.6 \pm 1.1 \times 10^6$	$1.9 \pm 0.5 \times 10^6$	$4.2 \pm 1.2 \times 10^5$
SPV ^b	3353 ± 2650	431 ± 224	—
SPVA ^c	225 ± 78	39 ± 1	—

^a Kinase assays were done in triplicate (see the Supplemental Data).

^b SPVI is an identification of the peptide substrate used in the kinase assay. The sequence of the peptide substrate used is SAEGVV VASE**SPVI**APHGSTHARSY, which corresponds to the segment (residues 213–237) of the serine phosphorylation site 6 (or SP6) in Pho4 (Jeffery et al., 2001). The consensus sequence (SPVI) is in bold.

^c SPVA, identical to SP6 peptide substrate except that the I residue at the +3 position of the consensus sequence is replaced by A.

The close similarity in the mechanism of substrate recognition by Pho85-Pho80 and CDK5-p25, being governed by F138 in Pho80 and E240 in p25, is in accord with the finding that the mammalian CDK5 is the closest functional homolog of the yeast Pho85. The mechanism is consistent with the findings that Pho85 can bind to p35/p25 and phosphorylate CDK5 substrate or, vice versa, that CDK5 can bind to some of the Pho80/Pcls and phosphorylate Pho85 substrates (Huang et al., 1999; Nishizawa et al., 1999).

Whether the residues in the other Pcls corresponding to F138 of Pho80 or E240 of p25, which are located in a similar loop, also play a role in recognizing the residue at the +3 position of the phosphorylation sites of substrates of other Pho85-Pcl complexes remains uncertain. Interestingly, the substrates that have been identified for several other complexes have different residues at the +3 positions (Table S2.1). However, sequence alignment of Pho80 and Pcls indicates that the residues aligned with F138 residue vary, but not in a manner that would be complementary to the different +3 residues of the substrates (Figure S2.1 and Table S2.1).

A Unique Site on Pho80 for Tight Binding of Pho4 Substrate

The structure of Pho85-Pho80 has further provided insights into two other major roles of the regulatory subunit in the PHO signaling system, owing to its tight

interactions with a site on Pho4 distal to the phosphorylation sites and with the CKI, Pho81. In addition to its potential role in interacting with the substrate in the active-site region, Pho80 plays an important role in substrate selection through binding with the Pho4 distal site, thereby enhancing the catalytic efficiency by orders of magnitude (e.g., see Table 2.2) and enabling semiprocessive phosphorylation of multiple serine residues (Byrne et al., 2004; Jeffery et al., 2001). The Pho4 distal site apparently comprises two long segments (1–30 and 156–200) (Byrne et al., 2004; Jayaraman et al., 1994). To identify the site on Pho80 for tight interaction with the Pho4 distal site, we have examined the wealth of data from mutagenesis studies of Pho80, focusing initially on those single amino acid changes that produce stable protein mutants but are unable to phosphorylate and inactivate Pho4 and mapping them onto the crystal structure. Point mutations of M42 to I, V, or L in Pho80 were found to suppress constitutively active alleles of PHO4, implicating this region, located in the α NT helix (Figure 2.3A), in Pho4 interaction (Okada and Toh-e, 1992). Interestingly, the three M42 mutations are able to suppress the constitutive phenotype of Pho4 L171 and 163–171 deletions (Okada and Toh-e, 1992), which are located in one of the two segments of the Pho4 sequence that comprise the distal, high-affinity site. Nine single amino acid missense mutations of Pho80, excluding the initiator M1 residue, have also been shown to cause the failure to repress PHO5 transcription at high phosphate levels (Madden et al., 1990).

Four of these mutations cluster adjacent to each other in the Pho80 structure: C30Y, L38F, and R41Q, which reside on the α NT helix and its preceding loop, and G229D, which resides close to the C-terminal end of α CT2 (Figure 2.3A). These four residues, together with M42, form a solvent-exposed, extended surface (Figure 2.3B)

remote from the active center or the Pho85-Pho80 interface. The other five mutations involving residues 130, 136, 148, 149, and 172 do not belong to another cluster or occupy positions near the active center, with the exception of the D loop D136 (Figure 2.1A; discussed below). The extended cluster makes up a significant portion of an oblong shallow cavity punctuated by a small central hole (Figure 2.3B) that is further bounded by the combined regions of the ordered N- and C-terminal loops (Figure 2.3A). Support for the participation of the two terminal loops in Pho4 binding is provided by the finding (Madden et al., 1990) that deletion of the first 16 or the last 47 (247–293) residue segments (which interestingly are completely invisible in the electron density [Figure 2.1A]) caused no diminution of the Pho80 function, but further deletion of residues 17–24 or 243–246 (which are far from the Pho85-Pho80 interface) resulted in nonfunctional Pho80. The cavity and its close vicinity exhibit three different surfaces: an intense positive electrostatic surface contributed mainly by portions of helices α NT and α CT2, a negative surface associated with a single carboxylate confined on the central hole, and a hydrophobic surface located primarily on the N- and C-terminal loop regions (Figure 2.3B). The role of hydrophobic interaction is consistent with the observation that five of the seven single amino acid substitutions in the two segments that cause Pho4 to be constitutively active are hydrophobic residues (Byrne et al., 2004). Hydrophobic interaction, combined with electrostatic interactions, may be sufficient to keep Pho4 in place in order to achieve multiple phosphorylations in a semiprocessive manner.

To test the hypothesis that the residues C30, L38F, R41, and G229 are part of a distal binding site for Pho4, we generated the four Pho80 mutants known to be defective in phosphorylation of Pho4 in vivo (C30Y, L38F, R41Q, and G229D) and assayed their

enzymatic activity. Each of the Pho80 mutants was coexpressed and purified in a stable 1:1 complex with Pho85 (Supplemental Data). Consistent with the proposed role of the residues in distal Pho4 binding, each of the mutations caused a 2.7- to 5.7-fold increase in K_M values (Table 2.3). As reported previously for Pho4 mutants defective in interaction with Pho80-Pho85 (Byrne et al., 2004), this decrease in K_M suggests that these residues may play a role in increasing the local concentration of Pho4. Surprisingly, the alterations in the putative distal site also caused a 1.8- to 8.1-fold decrease in k_{cat} . This finding has two possible explanations: either the mutations introduced have caused a subtle change in the Pho80 structure, perturbing the Pho85-Pho80 kinase active center, or the mutations are not only involved in increasing the local concentration of Pho4 but play some other undetermined role in the mechanism of Pho4 phosphorylation.

Table 2.3. Role of Distal Site Cluster on Pho80 in Pho4 Substrate Recognition

Pho80	Pho4 ^a			SPV ^{a,b}
	K_M (nM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (M ⁻¹ s ⁻¹)	k_{cat}/K_M (M ⁻¹ s ⁻¹)
Native	349.2 ± 56.1	11.4 ± 0.5	6.62 ± 1.12 × 10 ⁶	3353 ± 2650
C30Y	960.2 ± 233.8	2.4 ± 0.2	4.95 ± 1.64 × 10 ⁵	2167 ± 166
L38F	1263.0 ± 119.5	6.5 ± 0.3	1.05 ± 0.194 × 10 ⁶	1663 ± 859
R41Q	1999.4 ± 174.9	1.4 ± 0.1	1.42 ± 0.052 × 10 ⁵	1856 ± 594
G229D	1562.7 ± 80.8	6.0 ± 0.1	7.66 ± 0.357 × 10 ⁵	1525 ± 415

^a Kinase activity assays were performed in triplicate.

^b SPVI represents the peptide substrate that corresponds to the segment (residues 213–237) of the serine phosphorylation site 6 (or SP6) in Pho4 (Jeffery et al., 2001). See Table 2.2, footnote b.

To determine whether the distal site mutations somehow caused alteration of the active site, we assayed the kinase activity of the cyclin mutants toward the SPVI peptide substrate, which lacks a site for high-affinity interaction with Pho80. If the active site of the kinase is perturbed, then we expect to see a similar decrease in kinase activity with the SPVI peptide as we saw for full-length Pho4. If, however, the loss in kinase activity is due solely to the lack of interaction between Pho4 and the putative distal site, then

the Pho80 mutants should phosphorylate SPVI peptide in a manner similar to wild-type Pho80-Pho85. In each case, k_{cat}/K_M of the distal site mutants for the peptide substrate SPVI changed by no more than 2.2-fold (Table 2.3; compared to 6.3- to 46-fold change in k_{cat}/K_M for full-length Pho4). While this change in activity does not rule out a slight perturbation of the active site, it suggests that the primary effect of these mutations is to disrupt distal interactions with Pho4. The structural and biochemical data presented here suggest that this region is a site for distal, high-affinity Pho4 interaction.

Many other CDK-cyclin substrates are believed to also contain distal binding sites, and an RXL motif has been identified for recruiting substrates, as well as the Cip/Kip type CKIs, to the cell-cycle cyclins (Adams et al., 1996; Schulman et al., 1998; reviewed in Endicott et al., 1999). Structure analysis has revealed a small hydrophobic patch on cyclin A for interacting with the RXL motif (Brown et al., 1999; Russo et al., 1996a) (Figure 2.3C). The hydrophobic patch, which is conserved among the A, B, D, and E cyclins (Endicott et al., 1999), is located on the opposite side of the cyclin structure to the catalytic center and composed mostly of the N-terminal regions of helices $\alpha 1$ and $\alpha 3$ (Figure 2.3C). Although RXL motifs are present in the Pho4 distal binding site and the 80 residue segment of the Pho81 CKI that is sufficient for kinase inhibition, they are not required for substrate and inhibitor binding to Pho85-Pho80 (Huang et al., 2001 and the Supplemental Data). A clear explanation for this finding has emerged from the structure of Pho85-Pho80. Neither the hydrophobic patch on cyclin A nor the set of residues (M210, I213, and E220 on $\alpha 1$ and R250 on $\alpha 3$) that defines the specificity of the patch for the RXL motif (Endicott et al., 1999) (Figure 2.3C) is present in Pho80 (Figure 2.3D). In fact, the area in question in Pho80 bears no resemblance to

that harboring the hydrophobic patch on cyclin A. As shown in the superimposed structures in Figure 2.3D, the area in Pho80 is overlaid mostly by segments of the loops preceding $\alpha 1$ and $\alpha 3$ of cyclin A and, moreover, the first two turns of the $\alpha 1$ helix in cyclin A, which provide M210 and I213, are nonexistent in Pho80. Interestingly, the hydrophobic patch on the cell-cycle cyclins (Figure 2.3C) appears to be much smaller than the region in Pho80 for tight binding of Pho4 (Figures 2.3A and 2.3B). A larger region in Pho80 likely also reflects the need to accommodate and bind tightly a more extensive Pho4 distal site formed from the folding of two long discontinuous polypeptide segments.

Location of a Site on Pho80 for Tight Binding of Pho81 CKI

The activity of Pho85-Pho80 is further regulated by a CKI, Pho81 (Ogawa et al., 1995; Schneider et al., 1994). Pho81 is thus far the only well-established CKI of the Pho85-Pho80/Pcl family. It has the unusual property of forming a stable complex with Pho85-Pho80 under both high and low-phosphate concentrations, but only inhibiting under low-phosphate conditions (Schneider et al., 1994). Functional studies have demonstrated that Pho81 exerts its inhibitory activity by tightly binding mainly to Pho80 (Huang et al., 2001; Schneider et al., 1994) and have identified a segment of 80 residues that is necessary and sufficient for CKI function (Huang et al., 2001). Genetic screen studies have also indicated that the region of R121 and E154 in Pho80 is involved in binding Pho81 (Schneider et al., 1994). R121 and E154 reside near the N termini of helices $\alpha 3$ and $\alpha 5$, respectively, and form a salt link at the base of a solvent-exposed U-shaped surface, which is ~ 35 Å to the catalytic center (Figures 2.3A and 2.3B). The U-shaped surface and its close vicinity deploy positively charged residues,

including K70 near the C terminus of the α NT- α 1 loop and R151 at the N terminus of α 5, which form the two sides of the U surface (Figures 2.3A and 2.3B). Binding of the inhibitory segment may be accomplished partly by docking onto the U-shaped surface and, concomitantly, by making complementary electrostatic interactions with the nearby positively charged surface. Pho81 binding may also involve hydrophobic interaction because hydrophobic residues constitute half of the 80 residue primary sequence of the inhibitory domain of Pho81.

Recently, Lee et al. demonstrated that *myo*-D-heptakisphosphate (IP₇) regulates Pho81 inhibition of Pho85-Pho80 in vivo and in vitro (Lee et al., 2007). The positively charged nature of the putative Pho81 binding site and surrounding area (Figure 2.3B) could reflect a role in binding IP₇.

CONCLUSIONS

Because the Pho85-Pho80 complex structure is to our knowledge the first for a member of the superfamily of Pho85-Pho80/Pcl CDK-cyclin complexes, it sets the stage for future comparison with structures of other complexes that regulate other cellular functions. The structure has shed new light on the remarkable multiple key functional roles that Pho80 plays in the PHO pathway by harboring sites for conferring substrate specificity, for tight interaction with the distal site of the Pho4 substrate, and for high-affinity binding of the inhibitory domain of the CKI Pho81. The binding sites for the Pho4 and Pho81 have heretofore not been seen in other CDK-cyclin complexes, and moreover, neither site contains the equivalent of the hydrophobic patch on cell-cycle cyclins for interacting with the RXL motif of substrates or inhibitors. Pho80, in partnership with Pho85, further plays an important role in a prototypic mechanism for

circumventing T loop phosphorylation for full activation of the CDK-cyclin complex. By deploying both D136 and F138, the $\alpha 3$ - $\alpha 4$ loop or D loop of the Pho80 cyclin box has emerged as playing important dual roles in the activity of Pho85-Pho80. The charge coupling of Pho80 D136 with the Pho85 R132 makes activation loop phosphorylation no longer required for full kinase activation of the CDK-cyclin complex. The Pho80 F138 is potentially involved in the recognition of the I/L residue at position +3 of the consensus sequence of the five phosphorylatable sites on the Pho4 substrate. Further studies based on the heterodimer structure are required to further deepen our understanding of Pho85-Pho80 function, including the mode of binding of a substrate that would include the distal site, the strategy by which phosphorylation of multiple sites in Pho4 occurs in a semiprocessive manner, and the mechanism by which the Pho81 CKI, while remaining bound to Pho85-Pho80, modulates kinase activity as a function of phosphate concentration. Finally, structural studies of other Pho85-Pcl complexes are also necessary to unravel the mechanisms by which these complexes recognize different substrates.

EXPERIMENTAL PROCEDURES

Bacterial Expression and Purification of Pho85-Pho80

Coexpression of His6-Pho85 and Pho80 in *E. coli* essentially follows the published procedure (Jeffery et al., 2001). The methods used in the purification of Pho85-Pho80 are described in the Supplemental Data.

Crystallization, Diffraction Data Collection, and Structure Determination

Details regarding crystallization, diffraction data collection, and structure determination by single wavelength anomalous dispersion technique can be found in

the Supplemental Data. Statistics of the data collection and refinement are shown in Table 2.1.

Probing the Roles of the F138 Residue Close to the Active Site and the Distal Site on Pho80 in Substrate Recognition

Details regarding site-directed mutagenesis and kinase activity assay to assess the role of the Pho80 F138 and the potential site on Pho80 for interacting with the site on the Pho4 substrate distal to the phosphorylation sites are described in the Supplemental Data.

ACKNOWLEDGMENTS

We thank Dr. Stephan Ginell and staff members for their assistance in data collection at beamline ID19, APS, which is supported by the Office of Research, U.S. Department of Energy. The work was supported by NIH grants GM051377 to E.K.O. and GM068826 to F.A.Q., and by the Howard Hughes Medical Institute to E.K.O. and the Welch Foundation (Q-0581) to F.A.Q.

REFERENCES

- Adams, P.D., Sellers, W.R., Sharma, S.K., Wu, A.D., Nalin, C.M., and Kaelin, W.G., Jr. (1996). Identification of a cyclin-cdk2 recognition motif present in substrates and p21-like cyclin- dependent kinase inhibitors. *Mol. Cell. Biol.* *16*, 6623–6633.
- Brown, N.R., Noble, M.E., Endicott, J.A., and Johnson, L.N. (1999). The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat. Cell Biol.* *1*, 438–443.
- Byrne, M., Miller, N., Springer, M., and O’Shea, E.K. (2004). A distal, high-affinity binding site on the cyclin-CDK substrate Pho4 is important for its phosphorylation and regulation. *J. Mol. Biol.* *335*, 57–70.
- Carroll, A.S., and O’Shea, E.K. (2002). Pho85 and signaling environmental conditions. *Trends Biochem. Sci.* *27*, 87–93.

Endicott, J.A., Noble, M.E., and Tucker, J.A. (1999). Cyclin-dependent kinases: inhibition and substrate recognition. *Curr. Opin. Struct. Biol.* **9**, 738–744.

Espinoza, F.H., Farrell, A., Nourse, J.L., Chamberlin, H.M., Gileadi, O., and Morgan, D.O. (1998). Cak1 is required for Kin28 phosphorylation and activation in vivo. *Mol. Cell. Biol.* **18**, 6365–6373.

Huang, D., Patrick, G., Moffat, J., Tsai, L.H., and Andrews, B. (1999). Mammalian Cdk5 is a functional homologue of the budding yeast Pho85 cyclin-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **96**, 14445–14450.

Huang, S., Jeffery, D.A., Anthony, M.D., and O'Shea, E.K. (2001). Functional analysis of the cyclin-dependent kinase inhibitor Pho81 identifies a novel inhibitory domain. *Mol. Cell. Biol.* **21**, 6695–6705.

Jayaraman, P.S., Hirst, K., and Goding, C.R. (1994). The activation domain of a basic helix-loop-helix protein is masked by repressor interaction with domains distinct from that required for transcription regulation. *EMBO J.* **13**, 2192–2199.

Jeffery, D.A., Springer, M., King, D.S., and O'Shea, E.K. (2001). Multisite phosphorylation of Pho4 by the cyclin-CDK Pho80-Pho85 is semiprocessive with site preference. *J. Mol. Biol.* **306**, 997–1010.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N.P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 313–320.

Kaffman, A., Herskowitz, I., Tjian, R., and O'Shea, E.K. (1994). Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science* **263**, 1153–1156.

Lee, Y.-S., Mulugu, S., York, J.D., and O'Shea, E.K. (2007). Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science* **316**, 109–112.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EACBM Newsletter on Protein Crystallography* **26**.

Madden, S.L., Johnson, D.L., and Bergman, L.W. (1990). Molecular and expression analysis of the negative regulators involved in the transcriptional regulation of acid phosphatase production in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10**, 5950–5957.

Measday, V., Moore, L., Retnakaran, R., Lee, J., Donoviel, M., Neiman, A.M., and Andrews, B. (1997). A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. *Mol. Cell. Biol.* **17**, 1212–1223.

- Moffat, J., Huang, D., and Andrews, B. (2000). Functions of Pho85 cyclin-dependent kinases in budding yeast. *Prog. Cell Cycle Res.* *4*, 97–106.
- Morgan, D.O. (1996). The dynamics of cyclin dependent kinase structure. *Curr. Opin. Cell Biol.* *8*, 767–772.
- Morgan, D.O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu. Rev. Cell Dev. Biol.* *13*, 261–291.
- Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* *11*, 281–296.
- Nishizawa, M., Suzuki, K., Fujino, M., Oguchi, T., and Toh-e, A. (1999). The Pho85 kinase, a member of the yeast cyclin-dependent kinase (Cdk) family, has a regulation mechanism different from Cdks functioning throughout the cell cycle. *Genes Cells* *4*, 627–642.
- Ogawa, N., Noguchi, K., Sawai, H., Yamashita, Y., Yompakdee, C., and Oshima, Y. (1995). Functional domains of Pho81p, an inhibitor of Pho85p protein kinase, in the transduction pathway of Pi signals in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* *15*, 997–1004.
- Okada, H., and Toh-e, A. (1992). A novel mutation occurring in the PHO80 gene suppresses the PHO4c mutations of *Saccharomyces cerevisiae*. *Curr. Genet.* *21*, 95–99.
- O'Neill, E.M., Kaffman, A., Jolly, E.R., and O'Shea, E.K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* *271*, 209–212.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* *276*, 307–326.
- Pavletich, N.P. (1999). Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J. Mol. Biol.* *287*, 821–828.
- Poon, R.Y., Lew, J., and Hunter, T. (1997). Identification of functional domains in the neuronal Cdk5 activator protein. *J. Biol. Chem.* *272*, 5703–5708.
- Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J., and Pavletich, N.P. (1996a). Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature* *382*, 325–331.
- Russo, A.A., Jeffrey, P.D., and Pavletich, N.P. (1996b). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Biol.* *3*, 696–700.

Schneider, K.R., Smith, R.L., and O'Shea, E.K. (1994). Phosphate-regulated inactivation of the kinase PHO80-PHO85 by the CDK inhibitor PHO81. *Science* 266, 122–126.

Schulman, B.A., Lindstrom, D.L., and Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *Proc. Natl. Acad. Sci. USA* 95, 10453–10458.

Smith, D.S., Greer, P.L., and Tsai, L.H. (2001). Cdk5 on the brain. *Cell Growth Differ.* 12, 277–283.

Tarricone, C., Dhavan, R., Peng, J., Areces, L.B., Tsai, L.H., and Musacchio, A. (2001). Structure and regulation of the CDK5-p25(nck5a) complex. *Mol. Cell* 8, 657–669.

Toh-e, A., and Nishizawa, M. (2001). Structure and function of cyclindependent Pho85 kinase of *Saccharomyces cerevisiae*. *J. Gen. Appl. Microbiol.* 47, 107–117.

Toh-e, A., Ueda, Y., Kakimoto, S.I., and Oshima, Y. (1973). Isolation and characterization of acid phosphatase mutants in *Saccharomyces cerevisiae*. *J. Bacteriol.* 113, 727–738.

Toh-e, A., Tanaka, K., Uesono, Y., and Wickner, R.B. (1988). PHO85, a negative regulator of the PHO system, is a homologue of the protein kinase gene, CDC28, of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 214, 162–164.

Ueda, Y., Toh-e, A., and Oshima, Y. (1975). Isolation and characterization of recessive, constitutive mutations for repressible acid phosphatase synthesis in *Saccharomyces cerevisiae*. *J. Bacteriol.* 122, 911–922.

ACCESSION NUMBERS

Two sets of atomic coordinates have been deposited with ID codes 2PK9 and 2PMI.

CHAPTER 3

Genome Wide Characterization of the Phosphate Starvation Response in

Schizosaccharomyces pombe

Ian Carter-O'Connell, Michael Peel, Dennis D. Wykoff, and Erin K. O'Shea

AUTHOR CONTRIBUTIONS

The manuscript was written by ICO. Microarrays, ChIP-Seq, ChIP-qPCR, Pho7-TAP purification, and RNA-Seq were performed by ICO. DDW provided *S. pombe* strains. MP performed the Pho7 UAS and Csk1 URS analysis with the *pho1*⁺pr-YFP constructs. Computational analysis was performed by ICO using modified methods from previously published sources (Vijayan et al., 2011; Vijayan et al., 2009; Zhou and O'Shea, 2011). EKO and DDW provided supervision throughout and EKO is the corresponding author.

Genome Wide Characterization of the Phosphate Starvation Response in *Schizosaccharomyces pombe*

Ian Carter-O'Connell^{1,2}, Michael Peel⁴, Dennis D. Wykoff⁴, and Erin K. O'Shea^{1,2,3}

¹Howard Hughes Medical Institute,

²Department of Molecular and Cellular Biology,

³Department of Chemistry and Chemical Biology,

Harvard University, Faculty of Arts and Sciences, Center for Systems Biology,

Northwest Labs, 52 Oxford Street, Cambridge, MA 02138.

⁴Department of Biology,

Villanova University, 800 Lancaster Ave., Villanova, PA 19085.

INTRODUCTION

Inorganic phosphate (Pi) is an essential nutrient required for proper signal transduction, energy metabolism, and biochemistry within all living organisms. Maintaining a constant, stable concentration of internal inorganic phosphate is a major challenge for biological systems. External concentrations of inorganic phosphate fluctuate unpredictably, forcing microorganisms to develop a number of strategies to sense external phosphate levels (Abel, 2011; Bergwitz and Juppner, 2011; Wykoff and O'Shea, 2001), communicate this information to the nucleus (Doerner, 2008; Hulett, 1996), and induce transcription to respond to phosphate flux (Ogawa et al., 2000; Springer et al., 2003; Zhou and O'Shea, 2011). The phosphate signal transduction (PHO) pathway in the budding yeast, *Saccharomyces cerevisiae*, is the most thoroughly studied example of phosphate homeostasis in eukaryotes (Lenburg and O'Shea, 1996; Mouillon and Persson, 2006; Oshima, 1997).

The transcription factors Pho4 and Pho2 play a key role in the phosphate starvation response in *S. cerevisiae*. When cells are grown in conditions where inorganic phosphate is plentiful, Pho4 is multiply phosphorylated by the cyclin-dependent kinase-cyclin (CDK-cyclin) complex, Pho85-Pho80 (Kaffman et al., 1994). Phosphorylation excludes Pho4 from the nucleus (Kaffman et al., 1998a; O'Neill et al., 1996), prevents Pho4 from interacting with Pho2 (Komeili and O'Shea, 1999), and inactivates the PHO regulon. During phosphate starvation, the CDK inhibitor Pho81 binds to the secondary metabolite *myo*-D-inositol heptakisphosphate (IP₇) and inhibits the Pho85-Pho80 complex (Lee et al., 2008; Lee et al., 2007). Inhibition of Pho85-Pho80 allows Pho4 to become dephosphorylated, enter the nucleus (Kaffman et al., 1998b), co-operate with Pho2 (Zhou and O'Shea, 2011), and induce a set of genes responsible for harvesting inorganic phosphate from the environment (Ogawa et al., 2000). Pho4 regulation is measured via the activity of the secreted acid phosphatase, *PHO5*, which is one of the most highly induced members of the PHO response (Schmidt et al., 1963; Vogel et al., 1989). The other genes that comprise the PHO regulon have been well characterized and the precise sites within the genome where Pho4 binds during phosphate starvation are known (Maerkl and Quake, 2007; Zhou and O'Shea, 2011). Pho4 regulation occurs in response to changes in external phosphate levels and Pho4 activity is not known to be regulated by other stress responses.

In this study we ask the following: is the signaling of the PHO response in *S. cerevisiae* conserved in other ascomycota? *Schizosaccharomyces pombe* presents an interesting opportunity for answering this question because: (1) *S. pombe* shares a number of regulatory systems that appear in higher eukaryotes – RNAi-mediated

silencing (Bailis and Forsburg, 2002; Verdel and Moazed, 2005), extensive splicing (Wilhelm et al., 2008), more complex heterochromatin (Ekwall, 2004) – that are largely absent in *S. cerevisiae*; (2) *S. pombe* did not undergo the whole-genome duplication event, which is thought to contribute to specialization (Turunen et al., 2009) and; (3) the orthologs for the PHO pathway either do not exist (*PHO81*, *PHO2*, *PHO4*) or are not involved in the PHO response (*PHO80*, *PHO85*) in *S. pombe* (Tanaka and Okayama, 2000). However, due to the limited information available regarding the PHO pathway regulatory elements from *S. pombe* we were unable to pursue this question until recently.

Using a deletion library in *S. pombe*, the Wykoff lab recently identified a set of genes responsible for controlling the expression of *pho1*⁺, the secreted acid phosphatase orthologous to *PHO5* (Henry et al., 2011). They discovered that during phosphate starvation, Pho7 – a putative transcription factor containing a Zn₂Cys₆ binuclear cluster (MacPherson et al., 2006) – activates *pho1*⁺ expression. Csk1 – a CDK-activating kinase-activating kinase (CAKAK) (Hermand et al., 1998) – represses *pho1*⁺ expression in high-Pi conditions. Epigenetic analysis indicates that Pho7 acts downstream of Csk1.

In this study, we explore how these factors affect transcriptional output by characterizing the PHO response in *S. pombe*. We analyze the PHO response as a function of phosphate, Pho7, and Csk1 availability using DNA microarrays. We delineate a core PHO response comprising of the genes *pho1*⁺, *spbc8e4.01c*⁺ (an *S. cerevisiae* *PHO84* ortholog), and *spbc1271.09*⁺ (an *S. cerevisiae* *GIT1* ortholog), whose regulation in response to phosphate starvation is conserved between *S. cerevisiae* and

S. pombe. Interestingly, while these three genes share a functionally analogous regulatory pathway (i.e. activation through a transcription factor that is normally repressed by a kinase) we find that the mechanism for regulation differs widely between species. Our analysis of the Pho7-regulated transcriptional output – coupled with a global profile for Pho7 enrichment to promoters of stress responsive genes – also leads us to the conclusion that, unlike Pho4, Pho7 plays a role in multiple stress response pathways besides PHO. Further, high-throughput sequencing of the *S. pombe* transcriptome (RNA-Seq) reveals a contingent of non-coding and anti-sense transcription events regulated via *pho7*⁺. Finally, our RNA-Seq data provides evidence for *pho7*⁺ bi-functionality, as it activates certain genes and represses others. We conclude from this analysis that while there is a core PHO response shared between these two ascomycetes, the systems logic and specialization of PHO components varies widely.

RESULTS

Identification of the PHO responsive genes in *S. pombe*

The kinetics and abundance of transcriptional output vary widely between different environmental stress response pathways (Chen et al., 2003; Gasch et al., 2000; Hao and O'Shea, 2011). As gene regulatory networks respond to persistent stress, induction of downstream effectors may cause the false identification of genes that are not directly regulated by the stress being studied. To outline the specific PHO response in *S. pombe* for subsequent analysis, and to avoid indirect activation, we initiated a time-dependent, genome-wide expression analysis of wild-type *S. pombe* cells in medium lacking inorganic phosphate (no-Pi conditions). RNA samples from cells

starved for Pi were collected before and after 30, 60, 120, and 240 minutes of starvation and were analyzed by DNA microarray hybridization.

Through 60 minutes post-starvation we do not observe much transcriptional induction (Figure 3.1A), although by 120 minutes we note a greater than 2-fold-change (F.C.) in 62 genes (genes were required to maintain a greater than 2-F.C. at 240 minutes post-starvation, red lines in Figure 3.1A). This class contains the secreted acid-phosphatase, *pho1*⁺, the inorganic phosphate transporter, *spbc8e4.01c*⁺, and the organic phosphate transporter, *spbc1271.09c*⁺. They are three of the top induced genes through the entire time-course. As *pho1*⁺ and *spbc8e4.01c*⁺ induction has been previously observed in response to Pi starvation (Henry et al., 2011; Maundrell et al., 1985), we believe that this set accurately reflects the genes that respond quickly to changes in external Pi.

Does the apparent conservation of orthologous Pi harvesting and transport proteins between *S. pombe* and *S. cerevisiae* reflect a broad conservation in the PHO regulon? Comparing the candidate gene set with the previously defined PHO response genes in *S. cerevisiae* (Zhou and O'Shea, 2011) reveals that only 16% of the candidate genes in *S. pombe* are induced by phosphate starvation in *S. cerevisiae*. A much larger percentage (61%) of genes are specific to *S. pombe* Pi-regulation (Figure 3.1B). Besides *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09c*⁺, most of the conserved genes are thought to play a broad role in general stress response (i.e. *hsp104*⁺, *ssa1*⁺, *hsp16*⁺) (Chen et al., 2003).

Expanding to 240 minutes post-starvation captured 97 additional genes that are induced greater than the defined threshold (discussed in materials and methods)

(Figure 3.1A, blue lines). Only 5 of the additional genes are orthologous to genes in the *S. cerevisiae* PHO regulon (Figure 3.1C), with a large percentage implicated in a more generalized stress response. Interestingly, within this class of slow responding genes are multiple orthologs of the phosphate metabolism (PHM) class of enzymes responsible for maintaining vacuolar stores of polyphosphate and harvesting additional sources of organic phosphate upon Pi limitation (Ogawa et al., 2000). They include the glycerophosphoryl diester phosphodiesterase, *spapb1e7.05*⁺ (ortholog of *GDE1*) (Fisher et al., 2005), the lysophosphatidic acid phosphatase, *spac24b11.05*⁺ (ortholog to *PHM8*) (Reddy et al., 2008), and two members of the vacuolar transport chaperone complex, *spcc1322.14c*⁺ (ortholog to *VTC4*), and *spac14c4.11*⁺ (ortholog to *VTC2/3*) (Cohen et al., 1999; Huang et al., 2002; Thomas and O'Shea, 2005). We focused our attention on the fast responding genes for future analysis for two reasons: (1) to avoid indirect effects caused by persistent stress in cells and (2) because a significant portion of the slow responding genes are not regulated by *pho7*⁺ as determined by time-course experiments in a *pho7Δ* strain (data not shown).

Figure 3.1. Temporal Dynamics of the Phosphate Starvation Response in *S. pombe*.

(A) Line plot depicting the induction profile for genes displaying a fast (red) or slow (blue) response to phosphate starvation as measured by microarray analysis. Thresholds used to delineate the response time are described in the text. Induction was normalized to the initial sample pre-starvation ($t=0$). *spb1271.09c*⁺ (■) and *spb1322.14c*⁺ (●) are depicted as representatives of the fast and slow response, respectively. (B) Pie chart showing the conservation of the *S. cerevisiae* PHO regulon in *S. pombe*. Genes within the fast response were split based on orthology to genes in *S. cerevisiae* regulated by phosphate starvation (PHO Ortholog), not regulated by phosphate starvation (*S. pombe* only), or lacking an ortholog (No Orthologs). (C) As in B, except expanding to all genes that are induced 240 minutes post-starvation.

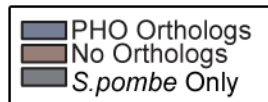
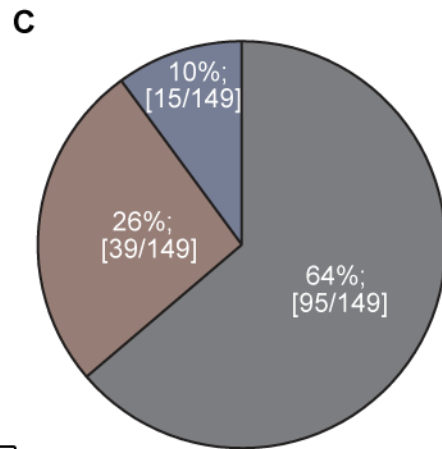
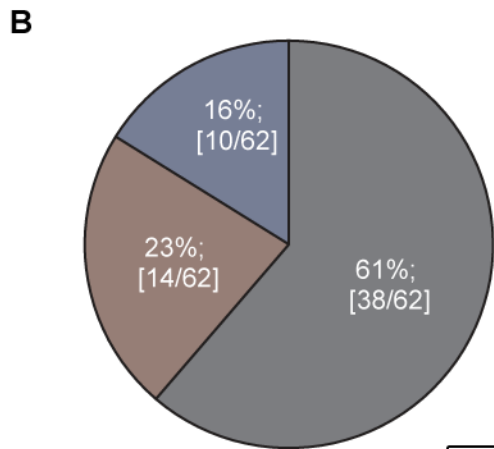
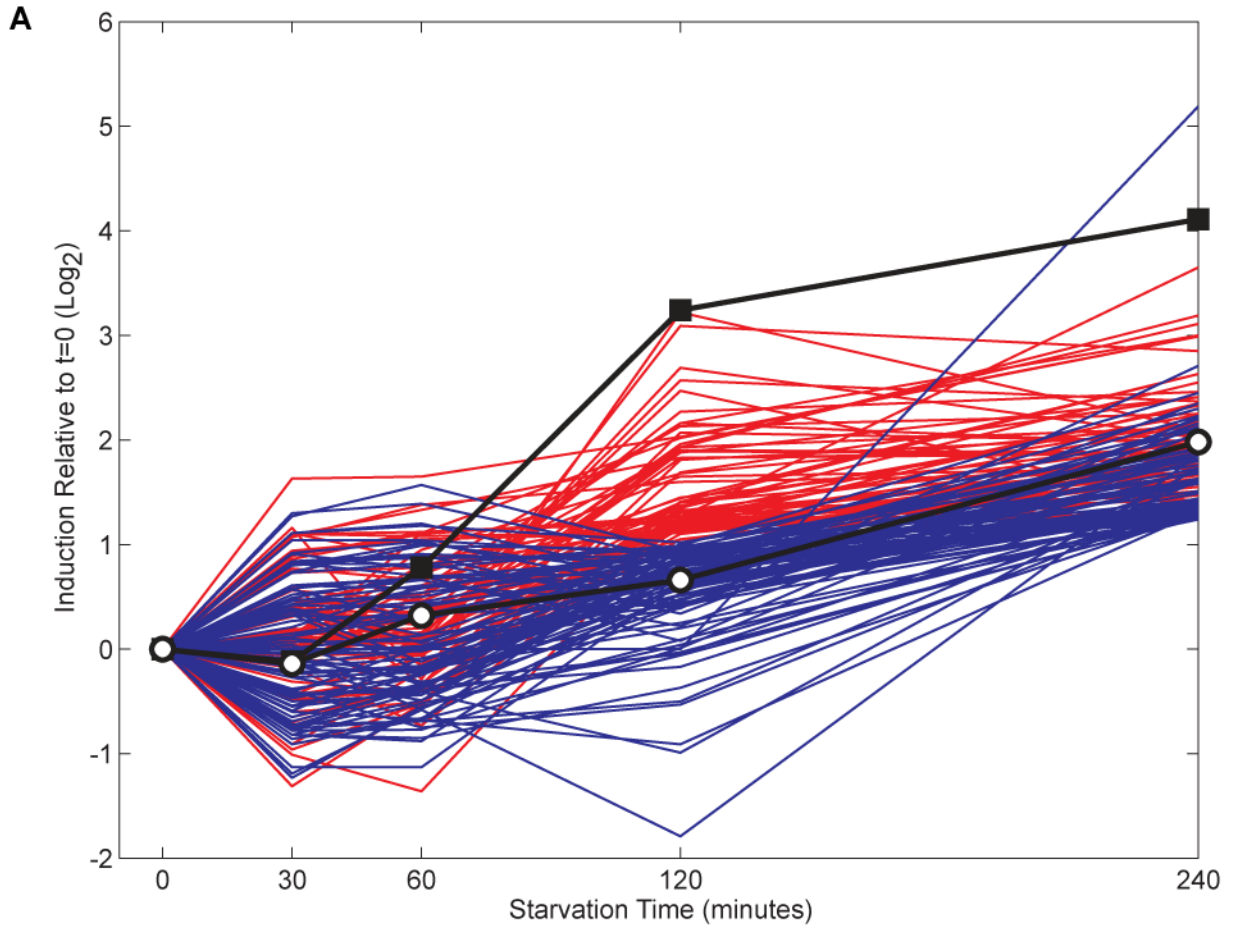


Figure 3.1. Continued

pho7⁺* and *csk1⁺* Regulate a Core Subset of the PHO Response in *S. pombe

Previous work indicates that *pho7⁺* and *csk1⁺* are important regulators of *pho1⁺* expression; we expected that they would also play a significant role in regulating additional components of the PHO response (Henry et al., 2011). To test our hypothesis we probed the transcriptional profiles of wild-type, *pho7Δ*, *csk1Δ*, and *pho7Δcsk1Δ* strains in high-Pi and no-Pi conditions at 120 minutes post-starvation using DNA hybridization microarrays. Each set of experiments was independently repeated and regulated genes were defined as having a ≥ 1.8 fold-change in expression between the two conditions of interest and a p-value ≤ 0.01 (see materials and methods).

We identified a set of 29 genes whose induction was responsive to Pi loss at 120 minutes post-starvation (Figure 3.2A, column A). If the induction of these genes is primarily dependent on the activity of *pho7⁺*, then their transcript abundance should decrease remarkably in a *pho7Δ* strain when compared to a wild-type background (both starved of Pi). 27.6% (8/29) of these genes have such a *pho7⁺* component (Figure 3.2A, column B, and Figure 3.2B). Additionally, a set of 20 genes has a Pi-independent, *pho7⁺* component (discussed below). If the repressor (*csk1⁺*) is responsible for preventing this induction by *pho7⁺* in high-Pi conditions, then we should see an increase in transcript abundance in a *csk1Δ* strain compared to the *csk1⁺* background in high-Pi conditions. Three genes display this response (Figure 3.2A, column C, and Figure 3.2B). A complete listing of all genes regulated by Pi, *pho7⁺*, and/or *csk1⁺*, along with their orthologs in *S. cerevisiae* can be found in Table 3.1. Finally, comparisons between each of the individual *pho7Δ* and *csk1Δ* strains with the *pho7Δcsk1Δ* double delete confirm the previously-described epistatic relationship between *pho7⁺* and *csk1⁺* (i.e. a loss of

pho7⁺ in a *csk1Δ* switches constitutive expression to uninducible expression) (Figure S3.1).

When phosphate is limiting, *S. cerevisiae* Pho4, along with Pho2, induces transcription of genes required for phosphate acquisition (Zhou and O'Shea, 2011). The secreted acid phosphatase, *PHO5*, is highly induced during phosphate starvation by the Pho4-Pho2 complex (Vogel et al., 1989; Yoshida et al., 1989). In high-Pi conditions the CDK-cyclin complex, Pho85-Pho80, represses Pho4 (Kaffman et al., 1994; Okada and Toh-e, 1992; Yoshida et al., 1989). The orthologs to Pho4, Pho2, Pho81, and Pho80 are not found in *S. pombe*, raising the question: does a functionally analogous signaling pathway involving *pho7⁺* and *csk1⁺* regulate the PHO response in *S. pombe*?

Comparing the Pho4, Pi-regulated genes in *S. cerevisiae* with the *pho7⁺*, Pi-regulated genes in *S. pombe* reveals an overlap of only three orthologs (Figure 3.2C). For *pho1⁺*, *spbcb8e4.01c⁺*, and *spbcb1271.09c⁺* a similar system of transcription factor activation, with repression by a kinase in high-Pi conditions, occurs. Unlike Pho85-Pho80 regulation of Pho4 in *S. cerevisiae*, most of the *pho7⁺* mediated response is independent of *csk1⁺* regulation.

To prevent the possible exclusion of additional Pi, *pho7⁺*, and *csk1⁺*-regulated genes in our analysis due to overly restrictive thresholds, we also examined gene expression under a less stringent threshold ($\geq 2\sigma + \text{median-F.C.}$) between microarray conditions (Figure 3.2D). The expansion of the candidate gene set reveals no new orthologs of Pho4-regulated PHO genes (data not shown). Besides *spbpb2b2.05⁺*, the only genes regulated by Pi, *pho7⁺*, and *csk1⁺* are the three previously identified genes (*spbpb2b2.05⁺* is very near the cutoff for having a *csk1⁺* component). Therefore, the

primary role for *pho7*⁺ and *csk1*⁺ in the PHO pathway is regulating this core set of phosphate harvesting and transport genes.

Figure 3.2. Pho7 and Csk1 Regulate a Core Set of PHO Responsive Genes. (A) Heat map showing the fold change (\log_2 scale) of Pi-regulated genes in a wild-type background (column A), a *pho7* Δ strain in no-Pi conditions (column B), and a *csk1* Δ strain in high-Pi conditions (column C), as measured by microarray analysis. Sub-columns indicate biological replicates. To the right is a heat map depicting the p-values generated for each gene within a given set of array replicates. Thresholds used for gene selection are described in the text. (B) Venn diagram showing the overlap between genes regulated by Pi starvation, *pho7*⁺, and/or *csk1*⁺. For a full list of Pi, *pho7*⁺, and *csk1*⁺ dependent genes please see Table 3.1. (C) Pie chart showing the percentage of Pi- and *pho7*⁺-dependent genes with Pi- and Pho4-dependent orthologs in *S. cerevisiae* (Zhou and O’Shea, 2011). Genes are divided as in Figure 3.1B, except orthologs also have a Pho4 component (Pho4-regulated Orthologs). (D) Same as B without a p-value minimum and a threshold for induction set to $\geq 2\sigma + \text{median}$ for each array set.

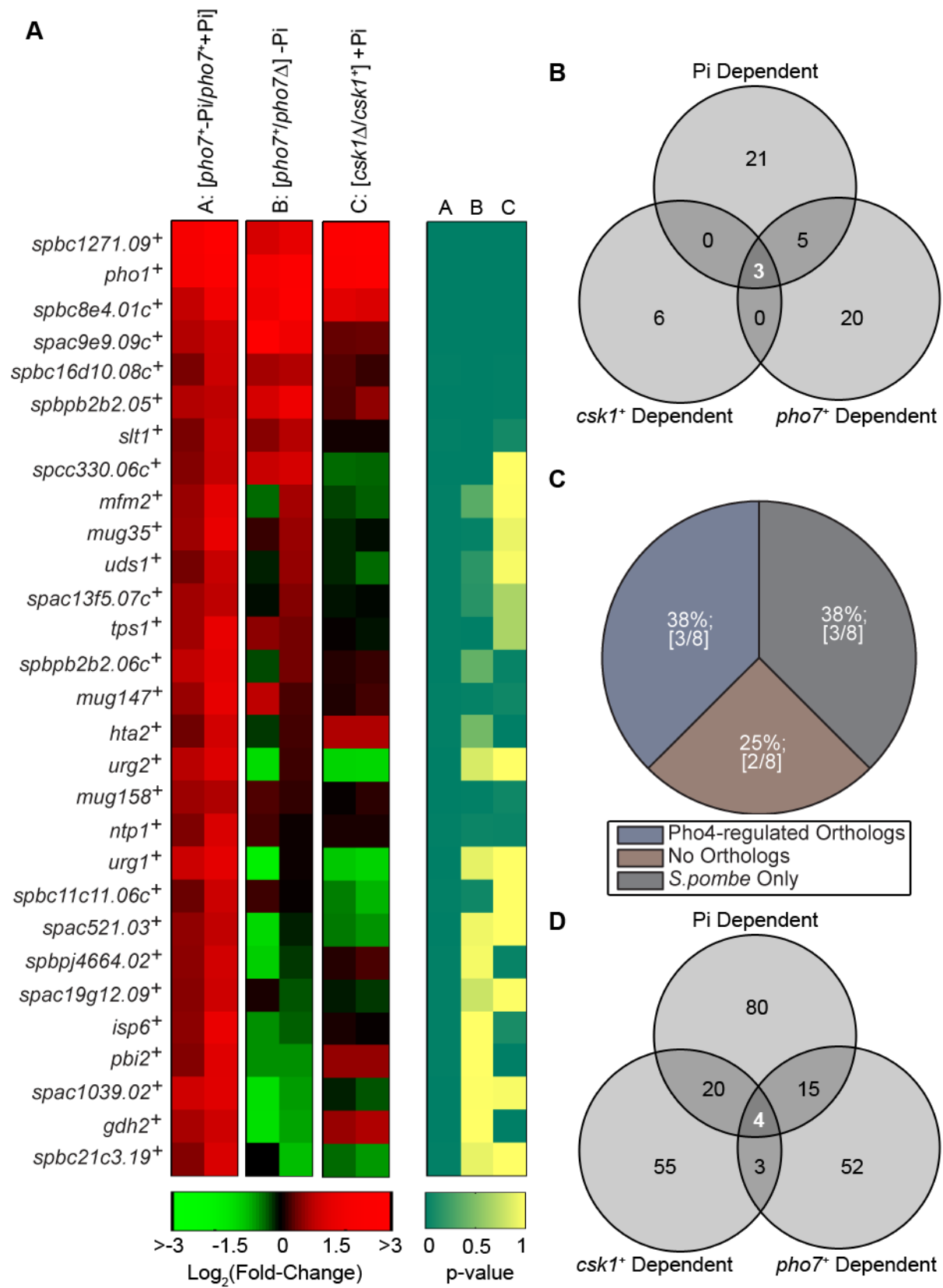


Figure 3.2. Continued

Table 3.1. Microarray results for genes regulated by Pi starvation, *pho7+*, and/or *csk1+*

ORF	NAME	Description ^a	S. cerevisiae Ortholog(s) ^b	[<i>pho7+</i> -P] <i>pho7+</i> +P]	p-value	[<i>pho7+</i> -P] <i>pho7Δ</i> -P]	p-value	[<i>csk1Δ</i> -P] <i>csk1+</i> +P]	p-value	PHO Ortholog ^c	Pho4-regulated Ortholog ^c
spac1002.17c	<i>urg2+</i>	uracil phosphoribosyltransferase (predicted)	YHR128W	1.21	0.00	-0.61	0.84	-1.07	1.00	--	--
spac1002.19	<i>urg1+</i>	GTP cyclohydrolase II (predicted)	YHR202W	0.00	0.00	-0.99	0.90	-0.99	1.00	--	--
spac1039.02+	<i>spac1039.02+</i>	phosphoprotein phosphatase (predicted)	NONE	1.38	0.00	-1.23	0.99	-1.23	0.99	--	--
spac1365.07c	<i>spac1365.07c+</i>	ZFPARP type zinc finger protein	NONE	0.94	0.01	0.30	0.16	-0.05	0.67	--	--
spac1367.02c	<i>ssa1+</i>	heat shock protein Ssa1 (predicted)	YBR169C, YLR046C, YER103W, YL024C, YLR034W, YAL005C, YBL075C, YPL106C, YLR030W, YMR006C, YMR008C, YLO111W	0.75	0.02	0.97	0.00	0.38	0.01	++	++
spac1786.02+	<i>spac1786.02+</i>	phospholipase (predicted)	YDR225W, YBL003C	-0.46	1.00	1.07	0.00	0.05	0.14	--	--
spac1786.13	<i>sh1+</i>	sequence orphan SH1	NONE	0.86	0.01	0.87	0.00	0.06	0.08	--	--
spac1912.06c	<i>hse2+</i>	histone H2A beta	YDR225W, YBL003C	0.90	0.01	0.02	0.45	0.74	0.00	--	--
spac1912.09	<i>spac1912.09+</i>	NADH/NADPH dependent indole-3-acetaldehyde reductase AKR3C2	YER145C, YBR207W	0.93	0.01	-0.15	0.79	-0.15	0.99	--	--
spac17.07c	<i>fyf1+</i>	iron permease Fyf1	NONE	0.41	0.12	0.96	0.00	-0.41	1.00	--	--
spac22h12.01c	<i>mgp35+</i>	sequence orphan	NONE	1.25	0.01	0.92	0.02	-0.10	0.92	--	--
spac2364.06c+	<i>spac2364.06c+</i>	methyltransferase (predicted)	NONE	0.86	0.01	1.14	0.01	0.13	0.18	--	--
spac24h6.13+	<i>spac24h6.13+</i>	meiotic recombination protein Rec10	YOL064W, YMR266W	-0.27	0.71	1.65	0.00	0.12	0.03	--	--
spac2767.09c+	<i>rect10+</i>	meiotic recombination protein Rec10	YLR263W	0.27	0.01	0.24	0.00	0.86	0.01	--	--
spac2767.09c+	<i>spac2767.09c+</i>	BuI2 family protein	NONE	0.72	0.05	0.99	0.01	-0.87	0.98	--	--
spac2h10.01	<i>spac2h10.01+</i>	transcription factor, z-fungal binuclear cluster type (predicted)	NONE	0.19	0.19	1.18	0.01	0.04	0.19	--	--
spac328.03	<i>lps1+</i>	alpha-alphatrehalose-phosphate synthase [UDP-forming]	YMR261C, YMR126C, YML100W, YML120C, YMR145C, YL065W, YEL060C, YOR003W	1.25	0.01	0.67	0.00	-0.02	0.67	--	--
spac3a1.107	<i>lps1+</i>	NADH dehydrogenase (predicted)	NONE	0.05	0.07	1.72	0.00	-0.15	0.95	--	--
spac4a8.04	<i>isp6+</i>	vacuolar serine protease Isp6	YKRO09C	1.20	0.01	-0.62	1.00	0.06	0.10	--	--
spac4b3.08+	<i>spac4b3.08+</i>	3-hydroxyacyl-CoA dehydrogenase (predicted)	YKRO09C	0.63	0.01	1.90	0.00	-0.84	0.98	--	--
spac513.03	<i>rim2+</i>	M factor precursor Rim2	NONE	1.16	0.01	NONE	0.35	-0.31	0.99	--	--
spac521.03+	<i>spac521.03+</i>	short chain dehydrogenase (predicted)	YMR226C	0.91	0.00	0.81	0.94	-0.52	1.00	--	--
spac599.09c+	<i>spac599.09c+</i>	aldehyde dehydrogenase (predicted)	YMR226C	1.09	0.00	2.86	0.00	0.39	0.00	--	--
spac618a9.04c+	<i>spac618a9.04c+</i>	sequence orphan	NONE	-0.41	1.00	-0.34	0.99	0.88	0.01	--	--
spac624d3.07c+	<i>spac624d3.07c+</i>	sequence orphan	NONE	0.80	0.00	0.96	0.01	-0.18	1.00	--	--
spac11c11.06c	<i>spac11c11.06c+</i>	sequence orphan	NONE	0.85	0.01	0.17	0.06	-0.63	1.00	--	--
spac11c11.12	<i>spac11c11.12+</i>	sequence orphan	NONE	0.57	0.02	-0.39	0.97	0.00	0.00	--	--
spbc1271.09	spbc1271.09+	glycophosphodiester transporter (predicted)	YCR098C	2.81	0.00	1.59	0.00	2.58	0.00	++	++
spbc1347.11	<i>sro1+</i>	S. pombe specific 51m protein family	NONE	0.40	0.01	1.19	0.01	-0.04	0.63	--	--
spbc1348.02+	<i>spbc1348.02+</i>	stress responsive orphan 1	NONE	0.27	0.00	0.21	0.06	0.87	0.00	--	--
spbc1348.13	<i>spbc1348.13+</i>	argininonine biosynthesis protein (predicted)	NONE	0.41	0.02	-0.40	0.95	1.06	0.00	--	--
spbc1604.01	<i>mgp159+</i>	heat shock protein Hsp104 (predicted)	YLR026W	0.85	0.00	0.30	0.01	0.09	0.06	--	--
spbc16d10.08c+	<i>spbc16d10.08c+</i>	ornithine transaminase Car2	YLR026W	0.89	0.01	0.97	0.00	0.26	0.01	++	++
spbc21c3.08c	<i>car2+</i>	SEDS family protein Rtc3 (predicted)	YLR026W	-0.26	1.00	1.17	0.01	-0.06	0.89	--	--
spbc21c3.19	<i>car2+</i>	SEDS family protein Rtc3 (predicted)	YHR087W	1.02	0.01	-0.56	0.91	-0.50	1.00	--	--
spbc27.04	<i>ucsf1+</i>	sulfate adenylyltransferase	NONE	0.85	0.01	0.31	0.17	-0.27	0.97	--	--
spbc27.08c	<i>svr1+</i>	heat shock protein Hsp16	YJR010W	-0.20	0.99	1.38	0.00	-0.20	0.85	--	--
spbc3a7.02c	<i>hsp16+</i>	heat shock protein Hsp16	YHR072W, YDR171W	0.45	0.11	2.32	0.00	-0.20	0.99	++	++
spbc56f2.06	<i>mgp147+</i>	sequence orphan	NONE	1.20	0.01	0.20	0.02	0.18	0.06	--	--
spbc660.07	<i>rlp1+</i>	alpha-alphatrehalase Nlp1	YDR001C, YBR001C	1.00	0.01	0.20	0.05	0.09	0.04	--	--
spbc664.01c	spbc664.01c+	inorganic phosphate transporter (predicted)	YML123C	1.54	0.00	2.75	0.00	1.18	0.00	++	++
spbc947.04+	<i>spbc947.04+</i>	cell surface glycoprotein (predicted), DIPS5 family	NONE	-0.48	1.00	1.54	0.00	0.33	0.00	--	--
spbc947.04+	<i>spbc947.04+</i>	FRG1 family protein, involved in mRNA processing (predicted)	NONE	0.05	0.30	0.85	0.00	0.14	0.04	--	--
spbc93a10.12	<i>spbc93a10.12</i>	acid phosphatase Pho1	YBR083C, YL024C, YBR082C, YHR215W, YAR071W	2.56	0.00	2.72	0.00	2.35	0.00	++	++
spbb2b25.05	<i>spbb2b25.05+</i>	peptidase family C26 protein	NONE	1.00	0.00	1.72	0.00	0.42	0.01	--	--
spbb2b25.06c+	<i>spbb2b25.06c+</i>	phosphoprotein phosphatase (predicted)	YHR202W	1.32	0.00	0.12	0.39	0.16	0.03	--	--
spbp1664.02	<i>spbp1664.02+</i>	cell surface glycoprotein (predicted)	NONE	0.97	0.01	-0.79	0.96	0.20	0.03	--	--
spcc1020.10	<i>oca2+</i>	serine/threonine protein kinase Oca2	YML163C, YDL214C, YDL025C, YOR267C	0.32	0.00	1.15	0.00	0.07	0.23	--	--
spcc132.04c	<i>gdi2+</i>	NAD-dependent glutamate dehydrogenase Gdi2 (predicted)	YDL215C	1.05	0.00	-1.26	1.00	0.68	0.00	--	--
spcc135.10	<i>ctr4+</i>	copper transporter complex subunit Ctr4	YLR411W	0.36	0.13	1.02	0.01	-0.58	1.00	--	--
spcc330.06c	<i>spcc330.06c+</i>	thioredoxin peroxidase (predicted)	YLR019W	0.86	0.00	1.34	1.00	-0.39	1.00	++	++
spcc338.12	<i>pbz2+</i>	protease B inhibitor Pbz2 (predicted)	YML015W	1.05	0.01	-0.77	1.00	0.68	0.00	--	--
spcc622.09	<i>rlp1+</i>	histone H2B Hlt1	YBL022W, YDR224C	0.84	0.04	-0.05	0.85	0.94	0.00	--	--
spcc645.14c	<i>svr1+</i>	chaperone activator Svr1 (predicted)	YOR272W	0.39	0.00	0.85	0.00	0.37	0.01	--	--
spcc650.07c	<i>zsf1+</i>	DNAJ domain protein, involved in translation initiation Pst1	YML027C	0.73	0.01	0.87	0.00	0.08	0.09	--	--

^aDescriptions for each gene product are provided from the S. pombe GeneDB: <http://rod.genedb.org/genedb/pombe/index.jsp> (Hertz-Fowler, et al., 2004)

^bOrthologs were identified using the Fungal Orthologs Repository: <http://www.broadinstitute.org/regulator/groups> (Wapinski, et al., 2006)

^cGenes passing the induction thresholds described in the text were compared to previously published results (Zhou and O'Shea, 2011) to determine whether they were orthologous to genes regulated by phosphate starvation (PHO) or by the transcription factor Pho4 (Pho4-regulated) in S. cerevisiae

Pi Starvation Causes Pho7 Enrichment at the PHO Core Promoters

Pho7 is classified as a putative transcription factor because it possesses a Zn_2Cys_6 binuclear cluster (ZC). The ZC domain serves as the DNA binding domain for a number of transcription factors (Punta et al., 2012; Sonnhammer et al., 1997). To test if Pho7 binds to *pho7*⁺-regulated promoters, we used a combination of chromatin-immunoprecipitation experiments followed by real-time quantitative polymerase chain reaction (ChIP-qPCR) or high-throughput sequencing (ChIP-Seq, see materials and methods).

We began our ChIP-qPCR analysis with a SWI/SNF component, Snf22, which was previously shown to bind the *pho1*⁺ promoter because we did not know if Pho7 would bind to the promoters of the candidate genes from our microarray analysis (Monahan et al., 2008). Cells containing an epitope tagged version of Snf22 (Snf22-TAP) were grown in high-Pi or no-Pi media for 2 hours and purified DNA was assayed by qPCR following the ChIP procedure. Snf22 binds to the *pho1*⁺ promoter (Figure 3.3A) and is enriched ~3.5-fold during Pi starvation. This increased enrichment of Snf22 upon Pi starvation was not observed in the previous report because they were only concerned with Snf22 enrichment in high-Pi (YES) media. Given that Snf22 acts by remodeling chromatin to allow efficient transcript production (Hirota et al., 2008), we believe this change in enrichment corresponds to the increased read-through of the *pho1*⁺ gene during Pi starvation.

Next we analyzed the *pho1*⁺ promoter by ChIP-qPCR to determine whether (1) Pho7 binds upon Pi starvation and (2) where binding occurs. Like Snf22, Pho7 displays a low level enrichment at the *pho1*⁺ promoter even in high-Pi conditions (Figure 3.3B).

Pho7 enrichment increased by ~3-fold during Pi starvation and appears preferentially bound in the -200 to -300 region of the PHO1 promoter. Pho7 and Snf22 share a similar pattern of enrichment at the PHO1 promoter, arguing that Pho7 is involved in transcriptional regulation (data not shown).

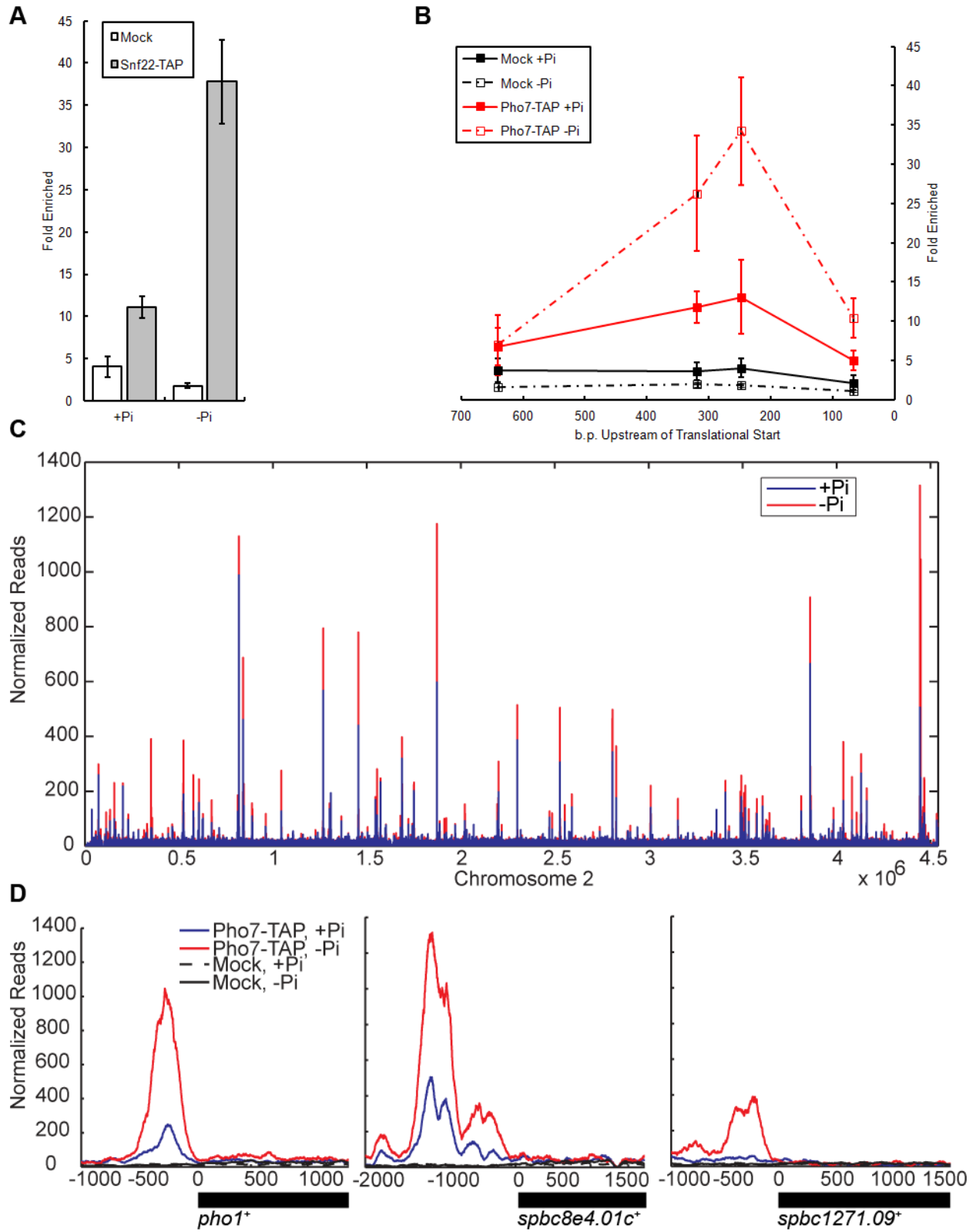
Armed with the preliminary evidence that Pho7 binds the promoter of *pho1*⁺ we used ChIP-Seq to assay global Pho7 binding events *in vivo* in response to Pi availability. Surprisingly, there is widespread Pho7 binding even in high-Pi conditions (1676 peaks out of 4208 total passed peak-ID thresholds, see materials and methods) (Figure 3.3C). Globally, Pho7 is enriched during Pi starvation at a number of sites (367 sites of 1476 total showed a ≥ 2 -F.C.), though the highest levels of enrichment are seen in the promoters of Pho7-regulated genes identified in the microarray analysis. This is in marked contrast to the global binding profile of Pho4 in *S. cerevisiae*. In that system, Pho4 is only recruited to the promoters of PHO regulated genes during phosphate starvation and, even then, only to relatively few locations (115) within the genome (Zhou and O'Shea, 2011).

Pho7 binding is significantly enriched in the promoters of *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺ in no-Pi conditions (Figure 3.3D). In the case of both *pho1*⁺ and *spbc8e4.01c*⁺, there is a significant enrichment of Pho7 even in high-Pi conditions when compared to mock treated samples. This explains the previously noted basal expression of *pho1*⁺ in *S. pombe* (Maundrell et al., 1985; Schweingruber et al., 1992), as it appears that Pho7 is bound and activating transcription even in the absence of stress (Figure S3.1, Column 1). The number of distinct Pho7 binding sites appears variable (with only one identified for *pho1*⁺, but multiples identified for *spbc8e4.01c*⁺ and

spbc1271.09⁺), which may be due to the presence of multiple DNA binding motifs within a segment of DNA or degeneracy in DNA motif preference (i.e. an ability to slide along the promoter). We do not find a clear correlation between the number of binding sites and transcriptional up-regulation. Identification of the upstream activating sequence for Pho7 is discussed below.

Figure 3.3. Pho7 is Recruited to the Promoter During Pi Starvation.

(A) Cells lacking the TAP epitope tag (Mock) or containing an epitope fused version of Snf22 (Snf22-TAP) were assayed for their enrichment of Snf22 at the *pho1*⁺ promoter via ChIP-qPCR, using previously described primer sets (Monahan et al., 2008). Cells were either incubated in high-Pi conditions (+Pi) or starved (-Pi) for 2 hours before cross-linking. Fold enrichment was determined by normalizing the *pho1*⁺ transcript counts to the Pi-independent *spo3*⁺ transcript as described in the text. Shown is the average of three biological replicates ± SE. (B) Primer sets were designed across the *pho1*⁺ promoter to determine local Pho7 binding. As in A, cells lacking the TAP epitope tag (Mock) or containing an epitope fused version of Pho7 (Pho7-TAP) were incubated for 2 hours in either high-Pi (+Pi) or no-Pi (-Pi) conditions. Fold enrichment was determined as previously described. Shown is the average of three biological replicates ± SE. (C) ChIP-Seq analysis of Pho7 binding in cells incubated in either high-Pi (blue) or no-Pi (red) conditions for 2 hours. Reads were normalized to total counts for each chromosome. Chromosome 2 is shown. (D) Magnified ChIP-Seq profiles for *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺. As in C, cells were incubated in high-Pi (blue) or no-Pi (red) conditions for 2 hours. Mock samples lacking the TAP epitope (black) were treated identically to experimental samples. Coordinates are relative to the initial ATG, with gene products identified by black bars. Reads were normalized as previously described.



Identification of the Pho7 Upstream Activating Sequence (UAS) in the *pho1*⁺

Promoter

Pho7 is a large (81.9 kDa) protein predicted to contain significant unstructured regions. It contains the previously mentioned ZC domain near its n-terminus and a number of predicted low-complexity regions (Figure 3.4A and Figure S3.2). The most thoroughly characterized member of the ZC domain family is the galactose metabolic regulator, Gal4 (see reviews: (Johnston, 1987; MacPherson et al., 2006). An alignment of the Pho7 ZC domain with a set of Gal4 ZC homologs reveals a high level of conservation in the cysteine residues responsible for zinc ion coordination and a modest amount of similarity in the surrounding sequence (Figure 3.4B). Based on this conservation, we initially attempted to use the structural information already known for ZC domains to determine the DNA binding motif for Pho7.

The crystal structure of Gal4 reveals a homodimer that contacts the functional groups of a CGG tandem repeat in DNA through two adjacent lysine residues (boxed in Figure 3.4B) (Marmorstein et al., 1992). The CGG repeats are separated by 11 semi-random nucleotides. The spacing between the tandem CGG repeats is essential for motif recognition and varies between ZC family members (Liang et al., 1996) (Figure 3.4C). Our motif analysis started by extracting the central 100 bp region of only the top 200 peaks as identified by Pho7 ChIP-Seq and using the MEME suite to identify enriched motifs *de novo* (Bailey et al., 2009). This initial analysis fails to provide a significant motif or set of motifs within the experimental pool. We lowered the gap penalty to allow identification of tandem repeats separated by an unknown linker length due to the previous observations regarding the Gal4 like ZC domains. This also fails to

produce a motif. Varying the initial input data set (larger or smaller lengths of DNA centered about the peak, all peaks from the high- or no-Pi conditions, only peaks for genes known to be regulated by Pho7), floating the number of required motifs within each sequence, and limiting the background model to sequences with no Pho7 binding (among many other perturbations), all fail to produce a motif. The most frequently found motif is presented as a MEME bit plot in Figure 3.4D. While the enrichment value for this motif over random noise would suggest a binding site for Pho7 (E-value=2.6e-10), even in our best results the motif only appears in fraction of the total sequences (31 out of 200). We do not know if Pho7 binds DNA as a dimer (let alone a homo or heterodimer) and the obvious differences between the DNA-contacting lysines in the Gal4 ZCs and the aspartate and asparagine residues found in the Pho7 orthogroup (Figure 3.4B), limits our ability to feed our analysis with enough seed information to detect real motifs. Advances in motif discovery based on new techniques with high-throughput sequencing and ChIP-Seq algorithms will facilitate this analysis (Bailey, 2011; Rhee and Pugh, 2011). We are hopeful that if we can obtain more structural information about Pho7 we could use these new techniques to positively identify the region of DNA targeted by Pho7.

Additionally, we pursued an *in vivo* strategy for identifying Pho7-promoter interactions that involved generating promoter fragments by reverse-transcription followed by quantitative real-time polymerase chain reaction (RT-qPCR). Differing lengths of the *pho1*⁺ promoter driving the expression of yellow fluorescent protein (*yfp*⁺) were constructed in an expression plasmid that was then transformed into either a *pho7*⁺ or *pho7*Δ background (Figure 3.4E). Transformants were assayed for their

expression of either the endogenous copy of *pho1*⁺ or the introduced *yfp*⁺ in high-Pi or no-Pi conditions (see materials and methods). The 2kb segment of the *pho1*⁺ promoter fully recapitulates the Pi dependent regulation that we see for the endogenous copy of *pho1*⁺ (Figure 3.4F, first two columns). This induction during Pi starvation is entirely dependent on the presence of Pho7, as it is abolished in a *pho7*Δ background (Figure 3.4F, middle column). As we expected from the ChIP-qPCR and ChIP-Seq experiments, trimming the promoter to either 1kb or 500 bp still results in Pho7 dependent *yfp*⁺ expression (Figure 3.4F, last two columns). Interestingly, the loss of the promoter region between -2kb and -1kb results in *yfp*⁺ de-repression in high-Pi conditions (discussed below). Trimming the promoter to 240 bp (*pho1UAS*Δpr-*yfp*⁺) abolishes *yfp*⁺ expression, confirming Pho7 binding to the region between -500 and -240 bp (Figure S3.6). Constructs containing incremental 20 bp deletions near the center of the ChIP-Seq peak for Pho7 binding were then transformed into *pho7*⁺ cells. The loss of *yfp*⁺ expression in cells confirms that the region -245 to -225 is necessary for Pho7 binding (Figure S3.6). These residues comprise an upstream activating sequence (UAS) for *pho1*⁺ expression.

Figure 3.4. Pho7 Alignment to Other Zn₂Cys₆ Binuclear Cluster Proteins and the Identification of the Pho7 Upstream Activating Sequence in the *pho1*⁺ Promoter.

(A) The predicted domain architecture of Pho7 based on electronic inference from the Pfam database (Sonnhammer et al., 1997). Low complexity regions are colored grey and the Zn₂Cys₆ binuclear cluster domain (ZC, residues 290-328) is colored black. (B) Alignment of amino acid sequences from 4 homologous ZC domains in other well characterized DNA binding transcription factors, along with the two Pho7 orthologs (Wapinski et al., 2006). Sequence identity is shaded in blue (darker for higher similarity). The bound box displays lysines 17 and 18 in Gal4 known to contact DNA and the equivalents in the other family members. (C) The DNA binding motifs for the ZC family. Each previously characterized member binds as a homodimer to a CGG tandem separated by a variably spaced linker (Linag et al., 1996). (D) Most frequently observed motif score from the Pho7 ChIP-Seq analysis as determined by MEME (Bailey et al., 2009). (E) Schematic for the Pho7 UAS discovery in the *pho1*⁺ promoter. Differing lengths of the *pho1*⁺ promoter (2kb to >200 bp) were fused to *yfp*⁺ on an exogenous plasmid. Additional constructs were designed with varying lengths of the putative Pho7 binding site deleted (500-245:225-1). (F) Constructs were transformed into *pho7*⁺ or *pho7*Δ backgrounds and the abundance of *pho1*⁺ produced from the endogenous locus or *yfp*⁺ generated from each promoter in high-Pi (+Pi) or no-Pi conditions (-Pi) was analyzed by RT-qPCR. All data were normalized to the expression of *act1*⁺. Shown is the average of three biological replicates ± SE.

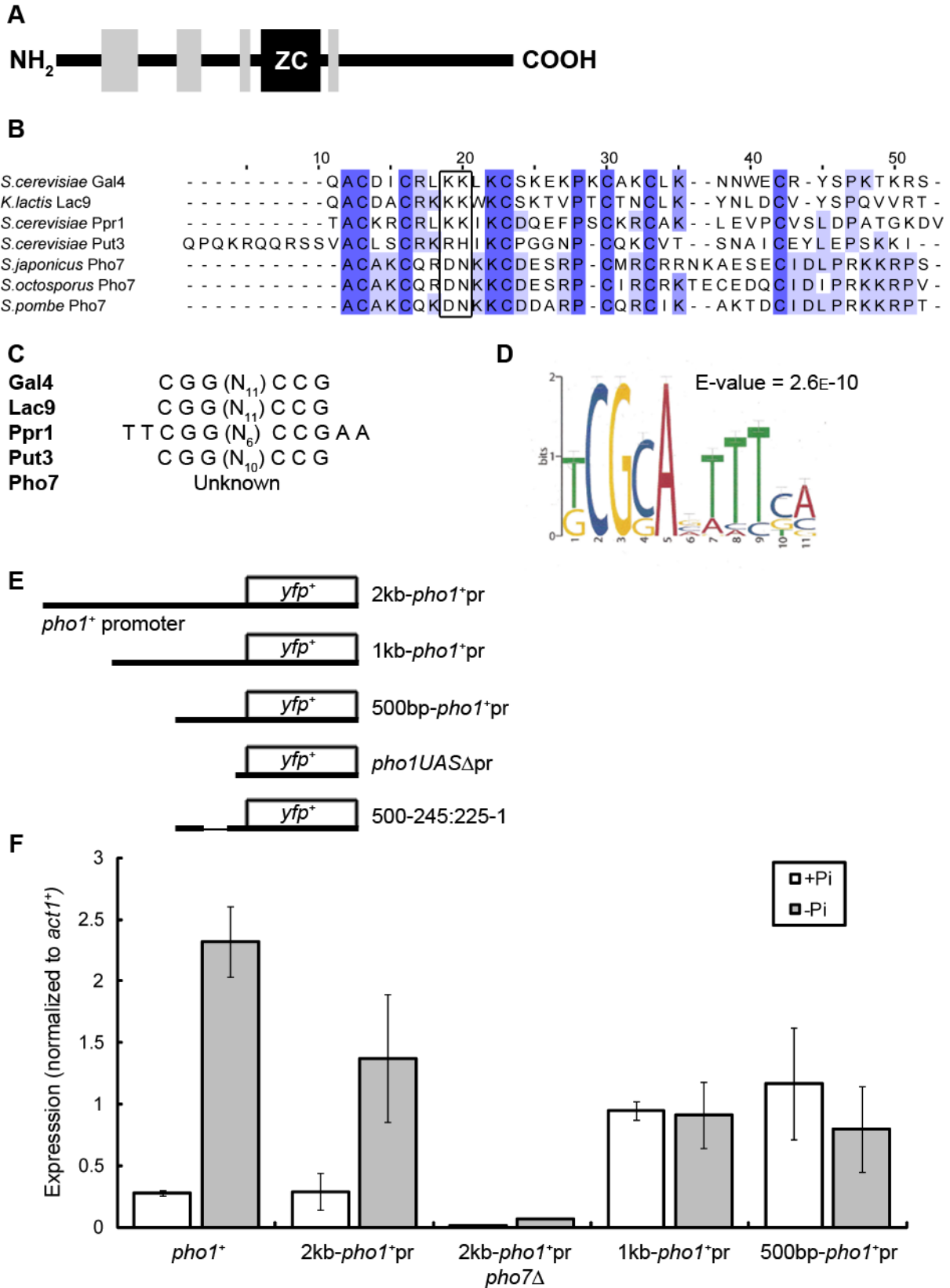


Figure 3.4. Continued

Attempts at Demonstrating a Direct Link Between Csk1 Kinase Activity and Pho7 Repression

Our microarray analysis confirms Csk1 repression of the core PHO genes (*pho1⁺*, *spbc8e4.01c⁺*, and *spbc1271.09⁺*) in high-Pi medium. We initially thought that repression of Pho7 activity might occur through direct phosphorylation of Pho7 by the Csk1 kinase. Recently, a genome-wide characterization of the phosphorylation state of the *S. pombe* proteome using phosphoform enriched mass spectrometry (MS) revealed two potential modification sites in the Pho7 protein (this study was conducted in YE media with sufficient amounts of external Pi to be considered high-Pi) (Wilson-Grady et al., 2008). We aligned these sites with (1) the other members of the Pho7 orthogroup and (2) sequence motifs known to be modified by Csk1 and Csk1 homologs (Hermand et al., 1998; Pei et al., 2006; Tsakraklides and Solomon, 2002). One of the sites (residues 459-467) is conserved in the orthogroup and has high similarity to other Csk1 sites (Figure S3.3). If Csk1 targets T463 *in vivo*, mutation of this residue to alanine (a phospho-delete mimic) should produce a Pho7 allele with decreased sensitivity to Csk1 regulation. This mutation does not produce a difference in *pho1⁺* expression in high-Pi conditions when compared to the *pho7⁺* allele (data not shown). This result is not sufficient to rule out a direct phosphotransfer between Csk1 and Pho7, as it remains possible that Pho7 is modified on a site not recognized in the MS analysis performed (due to a poorly purified phosphopeptide, a phosphopeptide that does not fly well in the MS chamber, etc.). Moreover, additional post-translational modifications might compensate for the loss of phosphorylation at residue 463.

To test whether Pho7 is post-translationally modified in response to Pi starvation we purified Pho7-TAP for post-translational modification mass spectrometry (PT-MS) from cells grown in either high-Pi or no-Pi conditions. Using a modified protocol for capturing sufficient amounts of endogenously TAP tagged protein (Gould et al., 2004) we affinity purified bulk preparations of Pho7-TAP (see supplemental data). We note that the purification of Pho7-TAP is contaminated with a high amount of background protein and a high prevalence of apparent Pho7-TAP degradation product (Figure S3.4B). We sought to address these issues by removing contamination with step-wise ammonium sulfate or polyethylimine precipitation, but we are unable to successfully purify a large enough pool of pure, full-length Pho7-TAP for MS analysis. This purification is further complicated by the low abundance of endogenous Pho7. Therefore, we set these efforts aside while we worked on alternative methods to determine the Csk1 repression mechanism.

Csk1 Represses Pho7 Through an Upstream Repressor Sequence (URS)

We decided to study what effect the loss of Csk1 would have on Pho7 enrichment in high-Pi conditions because of our observation that a decrease in Pi causes enrichment of Pho7 at the core PHO promoters (Figure 3.3D). If binding of Pho7 to the PHO promoters is necessary to drive increased transcriptional output, and Csk1 represses transcription by preventing Pho7 binding, then a loss of Csk1 in high-Pi conditions should mimic the no-Pi binding profile in wild-type cells. Pho7 binding at the *pho1⁺* promoter in a *csk1Δ* background is mildly increased compared to a *csk1⁺* background (Figure 3.5A, both in high-Pi). This increase in binding is less than the

observed enrichment of Pho7 in *csk1*⁺ cells grown in no-Pi conditions (1.6-fold versus 3-fold).

Intrigued, we examined the global effect of Csk1 loss on the binding profile of Pho7 using ChIP-Seq with *csk1Δ* cells grown in high-Pi conditions. Unlike the enrichment during Pi starvation, the deletion of Csk1 does not result in a global increase in Pho7 binding in high-Pi (compare Figure 3.3C with Figure 3.5B). At the core PHO responsive genes we observe either no change (*spbc8e4.01c*⁺) or a slight increase in Pho7 binding in the *csk1Δ* (*pho1*⁺ and *spbc1271.09*⁺), which is still well below the enrichment seen during Pi starvation (Figure 3.5C). We draw two conclusions from this data: (1) the level of Pho7 bound in high-Pi conditions would be sufficient to induce high levels of transcription save for the repressive action of Csk1; and (2) Csk1 does not repress Pho7 activity by preventing Pho7 from binding the promoter of responsive genes.

Our repression defective 1kb-*pho1*⁺pr-*yfp*⁺ plasmid (Figure 3.4F) provides us with an opportunity to ascertain whether Csk1 is repressing Pho7 through an interaction with the promoter. First, we want to know if our full *pho1*⁺ promoted *yfp*⁺ system (2kb-*pho1*⁺pr) also shows de-repression in high-Pi conditions in a *csk1Δ* strain. We find that, as with endogenous *pho1*⁺, a loss of Csk1 de-represses the expression of *yfp*⁺ from the *pho1*⁺ promoter (data not shown). We imagine that if Csk1 is directly modulating the activity of Pho7 independent of the proximity to the *pho1*⁺ promoter (i.e. phosphorylation in the nucleus or cytoplasm) then, so long as Pho7 binds a segment of our *pho1*⁺pr-*yfp*⁺ deletion constructs, expression should be modulated in a Csk1-dependent manner based on Pi availability. This is not the case, as Csk1 no longer represses the 1kb-

pho1⁺pr-*yfp*⁺ plasmid in high-Pi and there is no difference in *yfp*⁺ expression in a *csk1*Δ in high-Pi conditions. Thus, we conclude that Csk1 interaction with the *pho1*⁺ promoter at the upstream repressor sequence (URS) is necessary for Pho7 repression.

Our results lead us to the following model for Pho7 and Csk1 regulation. In high-Pi conditions a minimal amount of Pho7 is bound to the UAS in the *pho1*⁺ promoter. Pho7 in this state drives a basal expression of *pho1*⁺. Csk1, through an interaction with the URS in the *pho1*⁺ promoter, prevents Pho7 from maximally inducing *pho1*⁺ transcription. During Pi starvation, Csk1 repression is released and additional Pho7 is recruited to the *pho1*⁺ promoter, driving maximal expression (Figure 3.5D).

Figure 3.5. Csk1 Represses Pho7 Activity Through An Upstream Repressor Sequence.

(A) Cells lacking a TAP epitope tag (Mock) or containing a TAP fusion of Pho7 (Pho7-TAP) were incubated in high-Pi conditions in either a *csk1*⁺ or *csk1*Δ background and Pho7 binding to the *pho1*⁺ promoter was measured by ChIP-qPCR. Fold enrichment was calculated as previously detailed. Shown is the average of three biological replicates ± SE. (B) ChIP-Seq profile for chromosome 2 comparing Pho7 binding in either a *csk1*⁺ (blue) or *csk1*Δ (cyan) background in high-Pi conditions. Reads were normalized to the total number of reads for each chromosome. (C) Magnified ChIP-Seq profiles for *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺. As in B, cells from either a *csk1*⁺ (blue) or *csk1*Δ (cyan) background were incubated in high-Pi conditions for 2 hours. Mock samples lacking the TAP epitope (black) were treated identically to experimental samples. Coordinates are relative to the initial ATG, with gene products identified by black bars. Reads were normalized as previously described. (D) Model detailing the role of Csk1 in repressing Pho7 function at the *pho1*⁺ promoter. In high-Pi conditions Pho7 is not recruited to the UAS and Csk1 interacts at the URS to prevent full Pho7 activity. When the external concentration of Pi drops, the URS is repressed and additional Pho7 is recruited to the promoter to induce maximal induction.



Pho7 Regulates Expression in Response to Additional Stressors Besides Pi

Limitation

We initially avoided additional conditions known to cause a stress response in fission yeast because: (1) our previous work suggested that Pho7 was not involved in inducing gene expression in response to nitrogen limitation or carbon switching (Henry et al., 2011); and (2) the functionally analogous PHO system in *S. cerevisiae* is thought to be primarily dedicated to Pi homeostasis (Lenburg and O'Shea, 1996; Oshima, 1997). However, during our expression analysis we noticed a pool of genes that are repressed in a *pho7Δ* background and are not induced during phosphate starvation (Figure 3.2B). Additionally, we find that a Pho7 is bound to the promoter of a number of these genes in the ChIP-Seq analysis (Figure S3.5). These twin observations raise the following question: is Pho7 dedicated solely to the phosphate starvation pathway, like Pho4, or does it play a broader role in stress response?

To answer this question, we collected all the genes from our microarray analysis that display a $\geq 2\sigma + \text{median-F.C.}$ in expression between the wild-type or *pho7Δ* background (in either +Pi or -Pi conditions, removing *pho1⁺*, *spbc8e4.01c⁺*, and *spbc1271.09⁺*) and asked whether their promoters contained a significant peak of bound Pho7 within 800 bp of the start codon (Figure 3.6A). This produces a set of 63 genes that meet our requirements. This gene set was then processed through the Gene Ontology Tools: Term Enrichment algorithm (Carbon et al., 2009), allowing us to look for biological functions over-represented in our gene set. In order for a GO term to be considered enriched we require that at least three unique gene products be included in the term family and the enrichment must have a p-value ≤ 0.01 . 32 of the original 63

genes meet these requirements and are classified by the highest parent term available (Figure 3.6B). 15 of these 32 genes are involved in transmembrane transport, including 3 additional products involved in phosphate ion transport. Looking at the transmembrane category in further depth we find genes involved in iron, copper, and zinc transport (Figure 3.6C). A complete listing of GO Terms and the enriched gene set can be found in Table 3.2.

From the 32 identified genes we decided that *gpd1*⁺, *hpk2*⁺, *fiol*⁺, and *ctr4*⁺ expression would serve as proxies for Pho7 mediated transcriptional induction in various stress conditions. Gpd1 is a glycerol-3-phosphate dehydrogenase that is essential for survival during osmotic stress (Ohmiya et al., 1995). As the cell experiences increased osmotic pressure Gpd1 provides osmoprotection by increasing the intracellular pool of glycerol. Accordingly, its expression level is tied to the MAP-kinase pathway in *S. pombe* and it is induced upon an increase in osmolarity (Aiba et al., 1995). Hpk2 is a hexokinase that plays a role in regulating alternative carbon utilization when glucose sources are limited (Petit et al., 1996). It is maximally induced in response to a switch from glucose to glycerol as a carbon source. Fiol, in conjunction with Fip1, comprises the oxidase-permease iron transport system responsible for harvesting iron in depleted conditions (Askwith and Kaplan, 1997). During iron repletion *fiol*⁺ is repressed by the activity of Fep1, an iron sensing transcription factor (Pelletier et al., 2002). During iron starvation it is de-repressed and induced ~70-fold. Finally, Ctr4 is a high-affinity copper transporter that is induced in copper depletion conditions by Cuf1, a copper sensing transcription factor (Labbe et al., 1999). We designed RT-qPCR primer sets for each of these genes and measured their expression as a function of

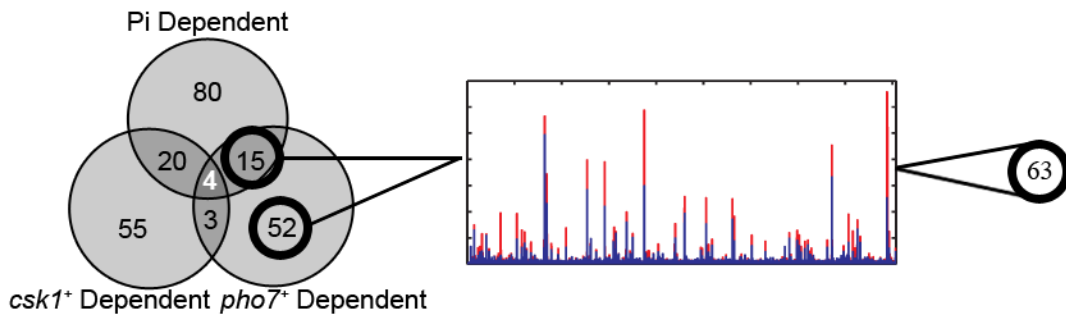
osmotic, iron, copper, and carbon utilization stress in both a *pho7⁺* and *pho7Δ* background (see materials and methods).

As previously demonstrated (Henry et al., 2011) the loss of Pho7 completely abrogates the induction of *pho1⁺* in no-Pi media (Figure 3.6D, first panel). For each of the additional stresses that tested, the loss of Pho7 causes a significant decrease (p-value ≤ 0.05) in the maximal induction of the target gene (Figure 3.6D). The Pho7 dependence of these genes varies, with some (*fiol⁺*, *ctr4⁺*) showing a relatively minor Pho7 component, while others (*gpd1⁺*, *hvk2⁺*) appear fully dependent on Pho7 to reach an optimal stress response. None of the Pho7-regulated responses were as drastic as that seen for Pi starvation, which may indicate that Pho7 plays a subtle role in coordinating expression at Pi-independent loci. Given this subtle response, the only recent availability of a deletion collection for *S. pombe* (Kim et al., 2010), and the fact that Pho7 function was unverified until this year, it is perhaps unsurprising that this role for Pho7 has not been previously observed.

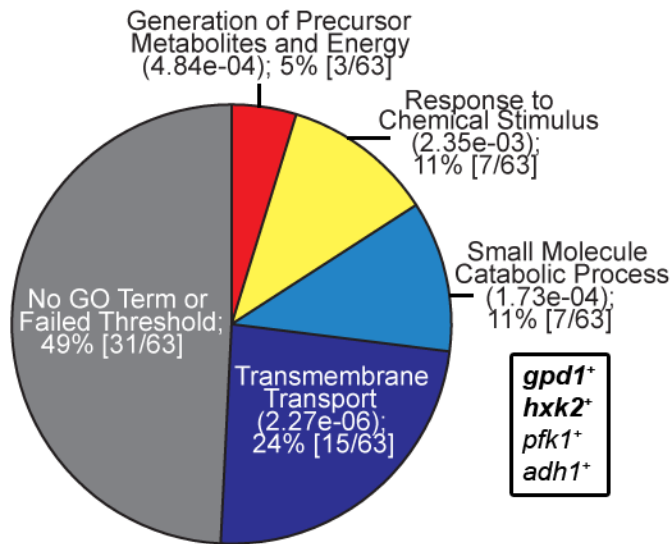
Figure 3.6. *pho7*⁺ Functions in Additional Stress Responses.

(A) Schematic depicting the initial selection of *pho7*⁺-regulated genes analyzed for possible involvement in other stress response networks. Genes that had a Pi- and *pho7*⁺-component based on microarray analysis were cross-referenced with ChIP-Seq to limit analysis to genes with *pho7*⁺ enrichment within 800 bp of the initial ATG. This produced a set of 63 unique genes. (B) Pie chart showing the GO Term enrichment for the 63 candidate genes from A as compared to a background model of all *S. pombe* transcripts. Due to multiple parent-child relationships and categorizations, each gene is shown as belonging to its highest parent class and child relationships have been compressed. For a complete summary of the GO output see Table 3.2. The boxed inset highlights a number of genes that were found in the small molecule catabolic process class, with those in bold serving as the initial targets for expression analysis. P-values for the enrichment against the background model are given; the computation is described in the text. (C) The transmembrane transport category from B was expanded to identify specific ions whose transport is regulated by *pho7*⁺. Candidate genes known to regulate iron (*fio1*⁺) and copper transport (*ctr4*⁺) were selected for expression analysis in iron and copper starvation conditions. (D) Expression of the stress response candidate genes identified in B and C in replete and stress conditions. *pho7*⁺ cells were incubated in either high-Pi (white) or stress (phosphate starvation: no-Pi, iron starvation: 250 uM DIP, copper starvation: 100 uM BCS, osmotic shift: 1.2M NaCl, carbon shift: 2% glycerol / 1% ethanol) (light gray) conditions, and the expression of the appropriate gene was measured by RT-qPCR. *pho7* Δ cells were also stressed (dark grey) to identify the level of *pho7*⁺-dependent regulation. Expression was normalized to *act1*⁺. Shown is the average of three biological replicates \pm SE.

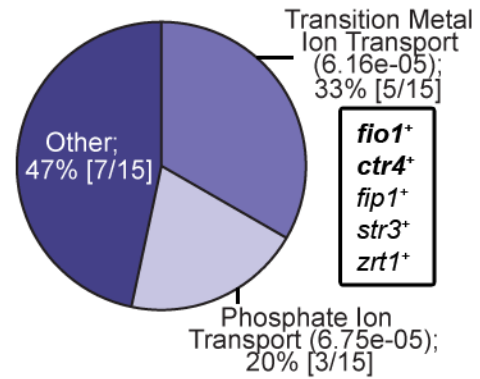
A



B



C



D

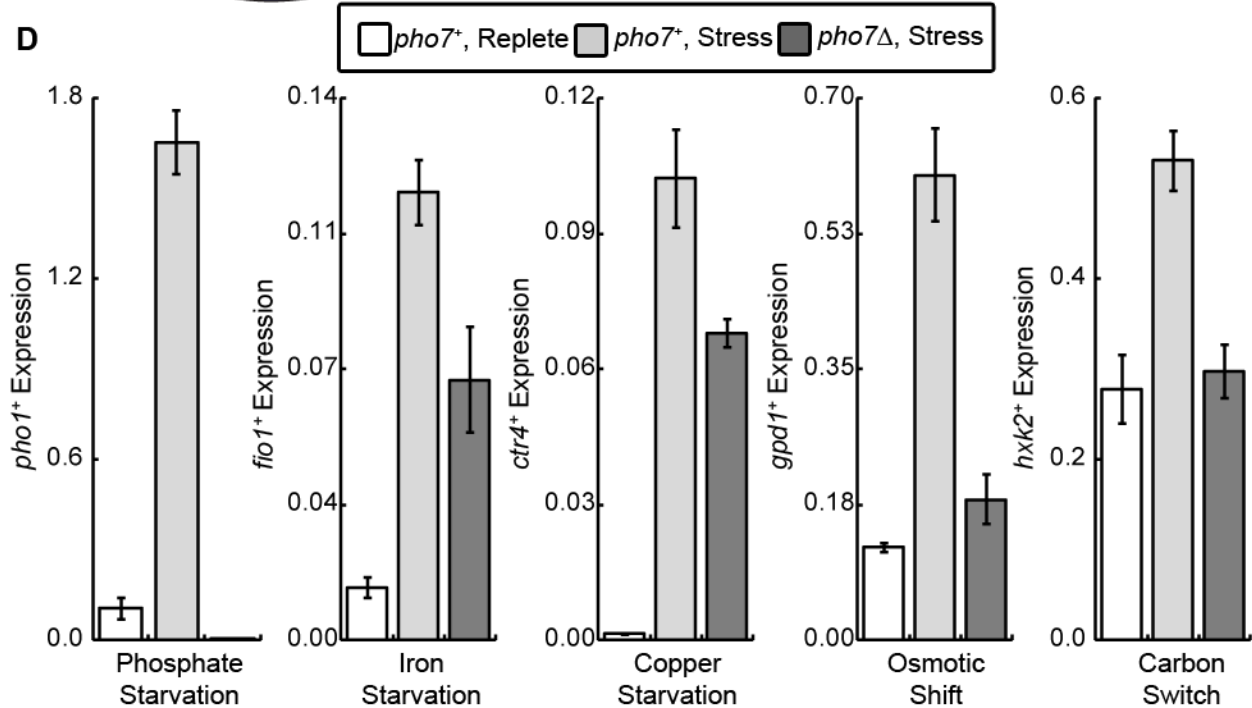


Figure 3.6. Continued

Table 3.2. Continued

GO Accession Number ^a	Description	P-value ^b	Sample Frequency ^c	Background Frequency ^d	Genes
GO:0006116	NADH oxidation	6.78E-05	3/63 (4.8%)	7/4914 (0.1%)	<i>spbc947.15c+</i> , <i>gpd1+</i> , <i>spac3at11.07+</i>
GO:0042592	homeostatic process	7.93E-03	7/63 (11.1%)	181/4914 (3.7%)	<i>zrf1+</i> , <i>sro1+</i> , <i>fio1+</i> , <i>str3+</i> , <i>gpd1+</i> , <i>fip1+</i> , <i>spcc330.06c+</i>
GO:0019725	cellular homeostasis	3.60E-03	6/63 (9.5%)	117/4914 (2.4%)	<i>sro1+</i> , <i>fio1+</i> , <i>str3+</i> , <i>gpd1+</i> , <i>fip1+</i> , <i>spcc330.06c+</i>
GO:0006879	cellular iron ion homeostasis	3.35E-03	3/63 (4.8%)	24/4914 (0.5%)	<i>fio1+</i> , <i>str3+</i> , <i>fip1+</i>
GO:0033212	iron assimilation	1.60E-04	3/63 (4.8%)	9/4914 (0.2%)	<i>fio1+</i> , <i>str3+</i> , <i>fip1+</i>
GO:0044282	small molecule catabolic process	1.73E-04	7/63 (11.1%)	94/4914 (1.9%)	<i>car2+</i> , <i>spac9e9.09c+</i> , <i>gpd1+</i> , <i>hxk2+</i> , <i>spac3g9.11c+</i> , <i>aruf1+</i> , <i>pfk1+</i>
GO:0046164	alcohol catabolic process	2.90E-03	4/63 (6.3%)	47/4914 (1.0%)	<i>spac9e9.09c+</i> , <i>gpd1+</i> , <i>hxk2+</i> , <i>pfk1+</i>
GO:0006091	generation of precursor metabolites and energy	4.84E-04	7/63 (11.1%)	111/4914 (2.3%)	<i>spbc947.15c+</i> , <i>cyc1+</i> , <i>gpd1+</i> , <i>hxk2+</i> , <i>spac3g9.11c+</i> , <i>spac3at11.07+</i> , <i>pfk1+</i>
GO:0070887	cellular response to chemical stimulus	1.05E-03	8/63 (12.7%)	164/4914 (3.3%)	<i>rgs1+</i> , <i>hsp16+</i> , <i>zrf1+</i> , <i>sro1+</i> , <i>gpa2+</i> , <i>hsp90+</i> , <i>spcc330.06c+</i> , <i>hsr1+</i>
GO:0042221	response to chemical stimulus	2.35E-03	8/63 (12.7%)	186/4914 (3.8%)	<i>rgs1+</i> , <i>hsp16+</i> , <i>zrf1+</i> , <i>sro1+</i> , <i>gpa2+</i> , <i>hsp90+</i> , <i>spcc330.06c+</i> , <i>hsr1+</i>
GO:0072524	pyridine-containing compound metabolic process	1.06E-03	4/63 (6.3%)	36/4914 (0.7%)	<i>spac9e9.09c+</i> , <i>spbc947.15c+</i> , <i>gpd1+</i> , <i>spac3at11.07+</i>
GO:0006733	oxidoreduction coenzyme metabolic process	2.90E-03	4/63 (6.3%)	47/4914 (1.0%)	<i>spac9e9.09c+</i> , <i>spbc947.15c+</i> , <i>gpd1+</i> , <i>spac3at11.07+</i>
GO:0055072	iron ion homeostasis	3.35E-03	3/63 (4.8%)	24/4914 (0.5%)	<i>fio1+</i> , <i>str3+</i> , <i>fip1+</i>

^aGO Term accession numbers and descriptions are provided by the Gene Ontology Consortium (Ashburner, et al. 2000). The AmiGO toolkit (Ireland, et al. 2009) was utilized to view parent-child relationships, which have been maintained in the above table. All data presented is from GO Database Release 2011-12-10.

^bP-values were calculated using the GO:TermFinder software (Boyle, et al. 2004). Only those GO Terms containing ≥ 3 genes and having a p-value ≤ 0.01 were included for further analysis.

^cSample genes are defined as (1) being repressed ≥ 1.5 -fold by the loss of *pho7+* in either phosphate replete or depleted conditions and (2) possessing one or more Pho7 binding sites within 800 bp of the translation start site in either condition. *pho1+*, *spbc8e4.01c+*, and *spbc1271.09+* were removed from this analysis.

^dThe background data-set was comprised of all ORFs present in the microarrays minus genes that are not annotated in the GO database.

It remains possible that the stress response effects we observe are artifacts limited to only the few genes we could study using RT-qPCR. Using microarray analysis with RNA collected from either *pho7*⁺ or *pho7*Δ cells grown in sated (non-stressed) or stress conditions, we examined the stress response mediated by *pho7*⁺ in the above conditions. We find that of 496 genes induced in the stress conditions studied, 124 genes have a *pho7*⁺ dependence (12 genes are induced in multiple stress conditions) (Figure 3.7A and Table 3.3). For the iron starvation response a subset of 15-18% of the total response relies on *pho7*⁺ (compare Figure 3.7B and C). This is similar in magnitude to the *pho7*⁺-dependent, Pi starvation response. This set of iron-responsive, *pho7*⁺-regulated genes has a significant enrichment for the biological process of iron assimilation (*fio1*⁺, *fip1*⁺, *str3*⁺, *sib2*⁺, p-value = 7.52e-06), akin to what we see for phosphate in no-Pi conditions. Moreover, we also show a significant enrichment (p-value = 4.9e-03) in genes responsible for cysteine homeostasis and transport, suggesting either an overlap in sulfur and iron metabolism regulation and/or an additional role for *pho7*⁺ in amino acid metabolism.

The copper starvation response is the least effected by the loss of *pho7*⁺ (only 8 genes had a *pho7*⁺ component) (Figure 3.7D). We do not find any GO term enrichment for this set of genes that would hint at a role for *pho7*⁺ maintaining a specific aspect of the copper homeostatic balance (i.e. copper transport). This is not surprising, as our ChIP-Seq analysis identifies only a minor Pho7 binding site in the *ctr4*⁺ promoter (compare the *ctr4*⁺ peak with that for *fio1*⁺/*fip1*⁺ in Figure S3.5) and the decrease in *ctr4*⁺ expression measured by RT-qPCR is relatively minor in a *pho7*Δ strain (p-value = 0.04) (Figure 3.6D). In fact, *ctr4*⁺ itself did not meet our requirements for a *pho7*⁺-regulated,

copper responsive gene in the microarray analysis. *pho7⁺* plays a role in mediating the PHO and iron starvation response, and does little to regulate the copper starvation response.

When we measure *pho7⁺*-regulated expression during an increase in osmotic pressure we find that of the 182 induced genes, 50% (90) are dependent on *pho7⁺* for maximal expression. The overlap between osmotic and *pho7⁺* responsive genes displays a significant enrichment (p-value = 1.35e-32) for genes involved in the general stress response. The candidate gene for this set, *gpd1⁺*, has a promoter bound by Pho7 to a high degree (Figure S3.5) corroborating our belief that *pho7⁺* is involved in regulating osmotolerance. One caveat remains for the osmolarity results: when we initiated this study, we wanted to keep the time variables the same for each condition to provide a more accurate comparison to our Pi starvation results. This was not a problem for the iron, copper, or carbon switching stress conditions – as the temporal dynamics appear very similar to the PHO response (Bellemare et al., 2001; Labbe et al., 1999; Pelletier et al., 2002) – but we do know that the osmotic stress response is thought to occur much more rapidly (Chen et al., 2003). Future experiments should include a temporal analysis for each of these *pho7⁺*-regulated stress responses to confirm that each set of genes studied are directly influenced by the activity of *pho7⁺* in response to the primary stress. Finally, we are planning microarray analysis with carbon switching to verify the *hvk2⁺* result. We are hopeful that this will reveal yet another stress response coordinated, at least in part, by *pho7⁺*.

Figure 3.7. Microarray Analysis of the Pho7-dependent Stress Response.

(A) Heat map showing the fold induction (\log_2 scale) of *pho7*⁺-regulated genes in different stress conditions. To assist in clustering, the range of induction in each condition was normalized between a min-max of -1 to 1. Genes that had a $\geq 2\sigma$ + median increase in expression between replete and stress conditions and $\geq 2\sigma$ + median loss in expression in a *pho7* Δ background are clustered using a k-means clustering algorithm (Eisen et al., 1998). Highlighted are the Pi, Fe, and Cu specific clusters. It should be noted that some genes (i.e. *spbc8e4.01c*⁺, *str3*⁺) are regulated in response to multiple stressors in a *pho7*⁺-dependent manner. (B-E) Venn diagrams showing the overlap between genes regulated by each tested stress and those with a *pho7*⁺ dependency in that stress (B: Pi starvation, C: iron starvation, D: copper starvation, E: osmotic shift).

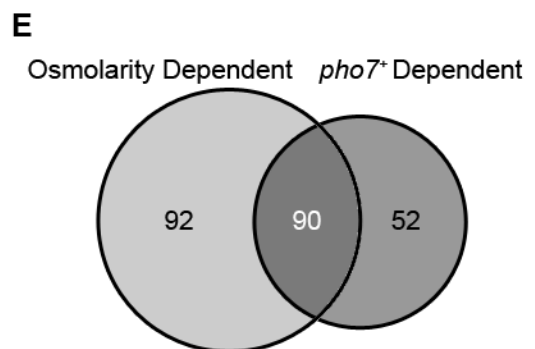
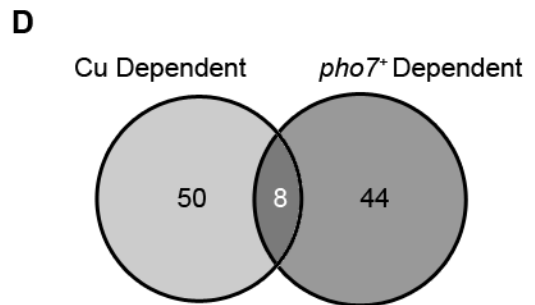
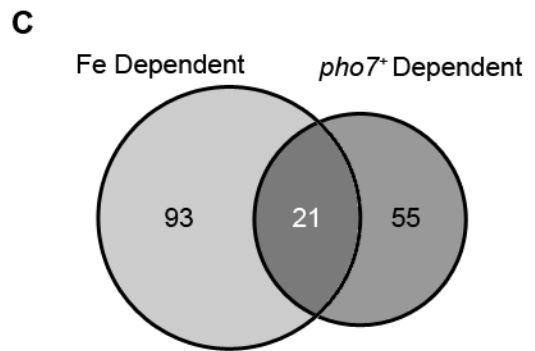
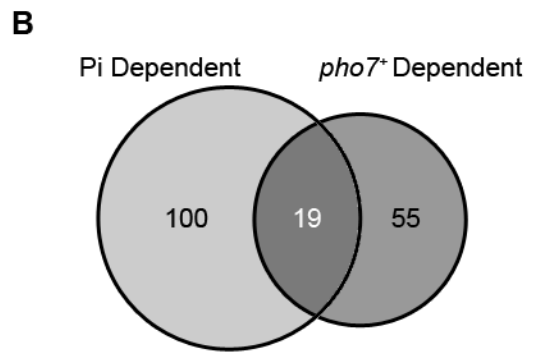
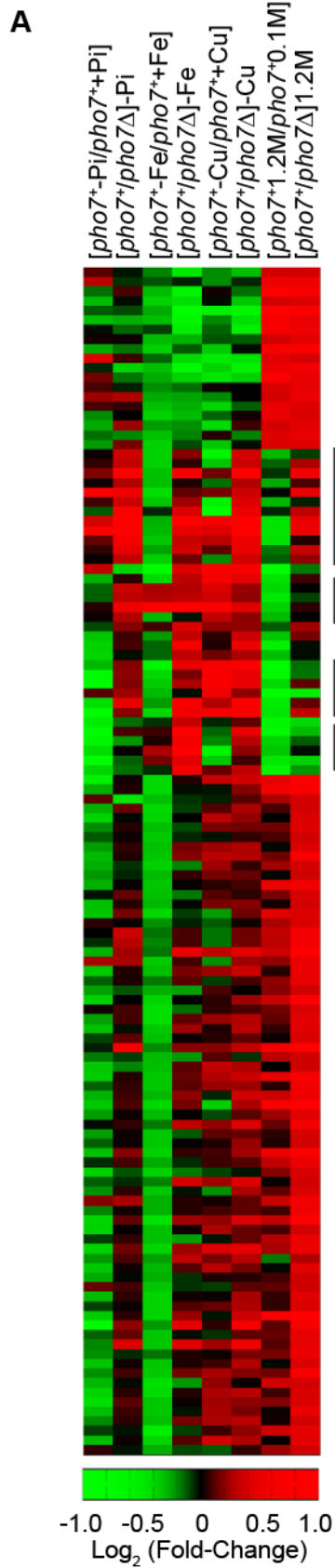


Figure 3.7. Continued

Table 3.3. Microarray results for genes regulated by phosphate, iron, copper, and/or osmotic stress in a Pho7 dependent manner

ORF NAME ^a	Description ^b	S. cerevisiae Ortholog(s) ^c	[-PI+PI] ^d	[pho7 ⁻ /pho7 ^Δ Δ]-PI	[Fe+HFe]	[pho7 ⁻ /pho7 ^Δ Δ]-Fe	[Cu+Cu]	[pho7 ⁻ /pho7 ^Δ Δ]-Cu	[1.2M0.1M]	[pho7 ⁻ /pho7 ^Δ Δ]-1.2M
spac1039.02+	phosphoprotein phosphatase (predicted)	YHR202W	1.38	-1.23	-2.00	1.27	3.99	1.45	-3.67	-1.10
spac1403.01c+	conserved fungal protein	NONE	-0.50	0.01	-0.17	0.11	0.28	0.42	3.79	2.53
spac139.05+	succinate-semialdehyde dehydrogenase (predicted)	YBR006W	0.14	0.06	-0.16	-0.12	-2.17	-0.20	4.41	2.51
spac13c5.04+	amidotransferase (predicted)	YLR126C	0.97	-1.46	0.26	-0.29	-2.23	-0.48	2.71	1.10
spac13f5.03c+	mitochondrial glycerol dehydrogenase Gld1	NONE	0.13	0.35	-0.12	-0.30	-1.20	0.05	1.86	1.12
spac13g6.06c+	glycine cleavage complex subunit P (predicted)	YMR189W	0.15	0.09	-0.13	0.91	2.50	0.81	-0.91	-0.08
spac13g7.02c+	heat shock protein Ssa1 (predicted)	YBR169C, YJR045C, YER103W, YLL024C, YJL034W, YAL005C, YBL075C, YPL106C, YEL009W	0.75	0.97	0.02	0.55	-6.30	0.50	-0.09	-0.71
spac14c4.08+	meiotically upregulated gene Mlg5	NONE	-0.09	-0.09	-0.08	-0.05	0.06	-0.12	1.37	1.24
spac15a10.05c+	YjeF family pyridoxamine-phosphate oxidase (predicted)	YNL200C	0.29	0.08	0.26	-0.37	-0.81	-0.20	2.16	1.32
spac167.06c+	sequence orphan	NONE	0.46	0.07	-0.20	-0.20	-2.16	0.61	1.89	1.07
spac17g6.13+	sequence orphan SH1	NONE	0.86	0.87	0.00	0.16	-4.22	0.34	0.71	0.62
spac1805.09c+	methionyl-tRNA formyltransferase Fmt1 (predicted)	YBL073W	0.24	0.14	0.24	0.44	-0.40	0.08	1.34	1.26
spac186.02c+	hydroxyacid dehydrogenase (predicted)	NONE	0.09	0.04	0.17	0.54	0.00	0.12	2.14	2.52
spac19g12.09+	NADH/NADPH dependent indole-3-acetaldehyde reductase AKR3C2	YDL124W	0.93	-0.15	0.83	-1.92	-3.35	-1.92	3.23	1.08
spac1H7.07c+	iron permease Fip1	YER145C, YBR207W	0.41	0.96	2.77	0.83	-0.09	-0.09	-0.64	-0.40
spac1H7.08+	iron transport multicopper oxidase Fio1	YFL041W, YMR058W, YDR506C	0.41	0.83	2.92	0.91	0.52	-0.13	-1.06	-0.76
spac1H8.03c+	siderophore-iron transporter Sit3	YHL040C, YKR106W, YEL065W, YCL073C, YOL156C, YHL047C	0.82	1.52	6.12	3.10	3.32	0.64	-0.50	-0.19
spac1H8.05+	sequence orphan	NONE	0.15	-0.29	0.00	-0.63	-0.42	-0.62	2.62	1.10
spac20g4.02c+	formin Fus1	YIL159W, YNL271C	0.24	-0.21	0.30	0.15	-0.88	-0.37	1.33	1.04
spac22a12.17c+	short chain dehydrogenase (predicted)	NONE	0.12	-0.04	0.17	-0.03	-1.55	-0.54	4.05	1.89
spac22e12.03c+	ThuJ domain protein	NONE	0.65	0.32	0.21	-0.59	-1.20	-0.65	2.27	1.21
spac22f3.02+	transcription factor Atf31	NONE	0.03	0.08	0.35	-0.08	-0.19	-0.19	1.54	1.10
spac22f3.12c+	regulator of G-protein signaling Rgs1	YLR452C	0.74	0.75	0.37	-0.15	0.17	-0.17	-0.42	-0.40
spac22f8.05+	alpha.alpha.trehalose-phosphate synthase (predicted)	YMR261C, YBR126C, YML100W	1.31	-0.29	1.18	-0.17	-3.33	-1.03	4.06	1.77
spac22g7.11c+	conserved fungal protein	NONE	0.02	0.41	-0.14	-0.30	-3.51	-0.34	3.72	1.27
spac22h10.13+	metallothionein Zym1	YOR031W	0.39	0.23	0.32	-1.95	-1.74	-1.36	5.54	2.42
spac23c11.06c+	hydrolase (inferred_from_context)	YNL115C	0.09	-0.13	-0.04	-1.46	-2.12	-2.02	3.30	1.26
spac23g3.03+	ornithine NS monooxygenase (predicted)	NONE	0.49	0.49	2.06	1.24	-0.57	-0.14	0.76	0.55
spac23h3.15c+	sequence orphan	NONE	0.47	-0.36	0.50	-2.75	-5.64	-3.03	6.45	2.31
spac26f1.05+	sequence orphan	NONE	0.01	0.16	0.61	0.95	-1.54	0.36	-0.32	0.26
spac26f1.07+	glucose 1-dehydrogenase (NADP+) (predicted)	YDR368W, YOR120W	0.84	-0.04	0.73	-0.19	-2.79	-0.76	2.69	1.19
spac27d7.09c+	But2 family protein	NONE	0.72	0.99	0.93	0.26	-3.05	-0.33	2.06	2.06
spac27d7.10c+	But2 family protein	NONE	0.58	0.79	0.81	0.25	-3.05	-0.34	2.93	2.01
spac27d7.11c+	glutathione transporter Pgt1	YJL212C, YPR194C	0.13	0.62	0.93	1.52	0.22	0.73	3.14	3.45
spac29b12.10c+	alpha.alpha.trehalose-phosphate synthase [UDP-forming]	YMR261C, YBR126C, YML100W	-0.10	0.18	1.04	2.09	-1.66	0.93	-1.61	-0.50
spac328.04+	AAA family ATPase, unknown biological role	YER047C, YPL074W	1.25	0.67	0.11	0.23	-3.39	-0.06	2.44	1.61
spac32a11.02c+	conserved fungal protein	NONE	0.00	-0.11	0.25	0.48	-0.41	0.35	1.63	1.61
spac343.12+	conserved fungal protein	NONE	0.07	-0.85	0.27	-2.16	-3.23	-2.36	4.75	1.62
spac3c7.13c+	glucose-6-phosphate 1-dehydrogenase (predicted)	YNL241C	0.43	0.05	0.90	0.90	-2.48	-0.53	2.41	1.45
spac3h8.09c+	poly(A) binding protein Nab3 (predicted)	YPL190C	0.21	-0.39	0.45	-0.81	-1.27	-0.85	2.33	1.08
spac4d7.02c+	glycerophosphoryl diester phosphodiesterase (predicted)	YPL206C	-0.04	-0.08	-0.06	-0.04	-1.59	-0.07	3.11	2.65
spac4f10.17+	conserved fungal protein	YOL109W	0.77	0.19	-0.30	-0.54	-1.56	-0.50	2.16	1.24
spac4g9.22+	conserved fungal protein	NONE	0.21	0.19	-0.33	0.35	-0.49	0.12	2.05	1.45
spac4h3.04c+	UPF0103 family	YJR008W	0.06	0.10	0.15	0.66	-1.70	0.16	1.75	1.36
spac4h3.04c+	3-hydroxyacyl-CoA dehydrogenase (predicted)	YJR009C	0.52	0.30	0.21	0.00	-1.62	-0.41	1.49	1.19
spac513.02+	phosphoglycerate mutase family	NONE	0.63	1.90	0.78	0.56	-0.04	-0.25	4.41	3.44
spac5f10.06+	sequence orphan	NONE	0.35	-0.31	0.30	-0.74	-1.21	-1.09	2.86	1.04
spac637.03+	glutathione S-transferase Gst3	YOR036C	-0.11	-0.14	1.05	0.93	-0.33	0.03	-0.15	-0.48
spac688.04c+	isocitrate dehydrogenase Icdp1 (predicted)	YDL066W, YLR174W, YNL009W	0.67	-0.49	1.31	-1.74	-2.95	-2.74	4.71	1.79
spac6f10.08+	dubious	NONE	0.77	-0.02	0.30	-0.93	-2.73	-0.26	3.32	1.37
spac806.11+	HHE domain cation binding protein (predicted)	NONE	0.01	-0.60	0.29	0.34	0.20	0.06	1.40	1.06
spac869.06c+	HHE domain cation binding protein (predicted)	NONE	0.11	0.50	-1.21	1.51	2.53	1.84	-2.40	-1.00
spac869.06c+	HHE domain cation binding protein (predicted)	NONE	-0.02	0.07	0.15	0.39	-0.02	0.33	3.89	2.45

Table 3.3. Continued

CRF	NAME ²	Description ³	S. cerevisiae Ortholog(s) ⁴	[PI+PII]	[rho7+rho7Δ]	[Fe+HFe]	[rho7+rho7Δ]	[Cu+Cu]	[rho7+rho7Δ]	[rho7+rho7Δ]	[rho7+rho7Δ]	[rho7+rho7Δ]
spac609.07c+	mel1+	alpha-galactosidase, melibiose	NONE	0.43	0.13	0.13	0.31	-0.47	-0.07	1.66	1.96	
spac609.08+	pcn2+	protein-L-isoaspartate O-methyltransferase Pcn2 (predicted)	NONE	0.12	0.15	0.36	0.56	-2.32	0.41	3.16	1.77	
spac609.09+	spac609.09+	conserved fungal protein	NONE	0.00	0.21	-0.14	0.44	-4.41	0.05	4.92	3.59	
spac611.10+	spac611.10+	sorbitase reductase (predicted)	NONE	-0.73	0.50	-0.14	0.44	2.96	0.82	-2.05	-1.06	
spac977.12+	spac977.12+	L-asparaginase (predicted)	YLR155C, YLR160C, YLR157C, YLR158C	-1.04	0.46	0.08	0.81	2.93	0.85	-0.63	0.10	
spac969.04+	spac969.04+	bcap family homolog (predicted)	YDL072C, YKL065C, YMR040W, YPL061W, YER073W, YOR374W, YPR127W	0.23	-0.01	0.32	0.06	-2.68	-0.45	2.56	1.33	
spac969.09c+	spac969.09c+	aldehyde dehydrogenase (predicted)	NONE	1.09	2.86	0.25	2.08	-0.37	2.31	1.06	2.00	
spac969.11+	plr1+	pyridoxal reductase Plr1	NONE	0.75	0.72	0.25	0.09	-2.36	0.08	0.07	-0.01	
spacunk4.17+	spacunk4.17+	NAD binding dehydrogenase family protein	NONE	0.16	-0.55	0.04	-0.45	-0.33	-0.91	4.97	2.70	
spac8a3.04c+	hsp9+	heat shock protein Hsp9	NONE	1.60	1.16	1.02	1.90	-5.51	-2.40	3.67	1.15	
spapb17e12.09+	spapb17e12.09+	sequence orphan	NONE	0.68	0.13	0.20	0.97	-0.54	0.58	2.10	2.07	
spapb2b4.04c+	spapb2b4.04c+	vacuolar calcium transporter P-type ATPase P2 type, Pmc1	YGL006W	0.21	-0.02	0.16	-0.43	-1.10	-0.38	1.69	1.13	
spbc106.10+	pkc1+	cAMP-dependent protein kinase catalytic subunit Pkc1	YKL166C, YLR164C, YPL203W, YDR261C, YLR300W, YGR190W	0.43	-0.01	0.12	-0.01	-1.72	-0.10	1.60	1.09	
spbc1105.05+	exg1+	glucan 1,6-beta-glucosidase Exg1	NONE	0.93	0.59	0.30	-0.13	1.85	-0.17	-2.85	-2.62	
spbc1198.01+	spbc1198.01+	glutathione-dependent formaldehyde dehydrogenase (predicted)	NONE	0.05	0.47	-0.01	0.17	-1.33	-0.03	4.03	2.61	
spbc1271.09+	spbc1271.09+	glycero-phosphodiester transporter (predicted)	YCR096C	2.81	1.59	-0.25	0.35	-0.85	0.35	1.34	1.56	
spbc12c2.04+	spbc12c2.04+	NAD binding dehydrogenase family protein	NONE	0.58	0.18	0.49	0.24	-1.42	-0.48	2.98	1.59	
spbc1685.08+	spbc1685.08+	serine protease (predicted)	YNL123W	-0.05	-0.39	0.01	-0.58	-2.61	-0.94	1.18	1.04	
spbc1685.13+	fhm1+	Fhm1 plasma membrane organization protein	YPR149W, YGR131W	0.44	-0.26	0.87	1.17	-0.32	0.59	1.46	1.04	
spbc16d10.08c+	spbc16d10.08c+	heat shock protein Hsp104 (predicted)	YLL026W	0.89	0.97	0.13	0.71	-5.93	-0.53	-0.07	-0.57	
spbc16e9.16c+	lsd90+	Lsd90 protein	NONE	0.61	-1.18	0.27	-2.34	-0.83	-2.83	3.92	2.00	
spbc20f10.03+	spbc20f10.03+	conserved eukaryotic protein	NONE	0.28	0.12	0.14	-0.13	-3.43	0.05	1.99	1.23	
spbc215.05+	gpd1+	glycerol-3-phosphate dehydrogenase Gpd1	YOL059W, YDL022W	1.10	0.67	0.44	0.12	-1.79	-0.40	3.14	1.81	
spbc21c3.19+	spbc21c3.19+	SBD5 family protein Rtc3 (predicted)	YHR087W	1.02	-0.56	0.52	-1.26	-5.18	-2.14	5.23	1.54	
spbc23g7.11+	mag2+	DNA-3-methyladenine glycosidase Mag2	YER142C	0.32	0.40	0.11	-1.32	-0.42	-0.42	1.77	1.83	
spbc27.08c+	suat1+	sulfate adenylyltransferase	YJR010W	1.38	1.38	0.65	2.12	-0.07	1.58	-1.26	0.14	
spbc2g2.17c+	spbc2g2.17c+	beta-glucosidase Psu2 (predicted)	YNL066W, YIL123W, YKR042W, YIL116C	-0.11	0.29	-0.38	0.69	-0.51	0.57	3.02	3.29	
spbc33f12.03c+	gpx1+	glutathione peroxidase Gpx1	YKL025C, YR037W, YBR244W, YOR348C, YCL025C, YGR191W, YPL274W, YLL061W, YBR069C, YKR039W, YDR508C, YFL055W, YNL270C	0.93	-0.18	0.77	-0.52	-1.81	-0.87	3.43	1.46	
spbc359.03c+	aaf1+	amino acid transporter Aat1 (predicted)	NONE	-1.08	0.49	-0.18	1.73	2.31	1.71	-0.97	0.87	
spbc3e7.02c+	hsp16+	heat shock protein Hsp16	YBR072W, YDR171W	0.45	2.32	1.43	1.24	-8.79	0.97	1.15	0.82	
spbc3h7.08c+	spbc3h7.08c+	conserved fungal protein	NONE	-0.06	-0.03	0.21	0.05	-0.15	-0.38	1.85	1.06	
spbc428.10+	spbc428.10+	sequence orphan	NONE	0.54	0.12	0.23	-0.82	-2.14	-0.38	2.79	1.65	
spbc428.11+	spbc428.11+	homocysteine synthase Met17	YLR303W	-0.36	-0.52	1.36	0.98	-2.28	-0.41	-1.88	-2.05	
spbc4b4.08+	glt2+	hexose transporter Glt2	YDL194W, YDL138W, YOL103W, YBR241C, YDR342C, YMR011W, YFL011W, YHR096C, YHR094C, YJL214W	-0.06	0.35	-0.08	0.67	0.90	0.77	2.65	1.72	
spbc4c3.08+	mug136+	acetylglucosaminyltransferase (predicted)	YOR320C, YJL146W	0.25	0.21	0.59	0.94	-0.82	0.14	0.55	0.67	
spbc557.05+	spbc557.05+	arrestin	NONE	0.01	0.00	0.36	0.29	-0.02	-0.30	1.34	1.38	
spbc562.06+	mug147+	sequence orphan	NONE	1.20	0.75	0.51	-0.84	-3.36	-0.85	3.31	1.66	
spbc649.04+	uvr15+	UV-induced protein Uvr15	YBR016W, YDL012C, YDR210W, YDR001C, YBR001C	0.16	-0.10	0.46	-0.29	-1.68	-0.68	1.86	1.05	
spbc660.07+	rip1+	alpha,alpha-trehalase Rip1	NONE	1.00	0.20	0.16	-0.27	-2.48	-0.44	2.04	1.05	
spbc660.09+	mug168+	sequence orphan	NONE	0.09	0.11	-0.04	0.05	-1.84	0.02	1.77	1.37	
spbc713.11c+	pmg3+	plasma membrane proteolipid Pmg3	YDR276C	0.52	-0.03	0.31	-0.10	-1.03	0.09	1.75	1.22	
spbc725.03+	spbc725.03+	pyridoxamine 5'-phosphate oxidase (predicted)	YGR017W	0.92	0.13	0.21	-0.92	-4.63	-1.31	3.86	1.29	
spbc839.06+	cta3+	P-type ATPase, calcium transporting Cta3	YGL167C, YDR036C, YGR039C, YDR040C	-0.03	0.11	-0.03	0.64	-2.20	0.36	5.40	3.13	
spbc8e4.01c+	spbc8e4.01c+	inorganic phosphate transporter (predicted)	YML123C	1.54	2.75	0.38	2.84	2.12	2.59	-0.90	1.61	
spbc8e4.02c+	spbc8e4.02c+	sequence orphan	NONE	-1.07	0.54	0.68	1.50	-1.57	1.22	1.98	2.26	

Table 3.3. Continued

ORF	NAME ^a	Description ^b	<i>S. cerevisiae</i> Ortholog(s) ^c	[PI+PI] ^d	[Pho7 ⁺ Pho7 ^Δ]PI	[Fe+Fe]	[Pho7 ⁺ Pho7 ^Δ]Fe	[Cu+Cu]	[Pho7 ⁺ Pho7 ^Δ]Cu	[1.2M:0.1M]	[Pho7 ⁺ Pho7 ^Δ]1.2M
<i>spbbp4g3.02+</i>	<i>pho1+</i>	acid phosphatase Pho1	YBR095C, YDL024C, YBR095C, YHR215W, YAR071W	2.56	2.72	0.34	1.44	0.93	2.52	-0.99	1.74
<i>spbbp4h10.10+</i>	<i>spbbp4h10.10+</i>	rhomoid family protease	NONE	0.41	0.35	0.15	-0.13	-2.46	-0.40	2.07	1.45
<i>spbbp10d8.01+</i>	<i>spbbp10d8.01+</i>	cysteine transporter (predicted)	YLL055W, YOL163W, YAL067C, YLL186C, YJR162W, YOL162W, YGR260W, YLR004C, YCR028C	-0.90	0.21	2.03	2.79	-2.90	0.86	-2.20	-1.62
<i>spbbp10d8.02c+</i>	<i>spbbp10d8.02c+</i>	arylsulfatase (predicted)	NONE	-0.63	-0.72	1.84	1.80	-3.23	0.13	-0.46	-0.90
<i>spbbp2b2.05+</i>	<i>spbbp2b2.05+</i>	peptidase family C26 protein	NONE	1.00	1.72	0.19	1.03	0.05	0.96	-0.96	-0.20
<i>spbbp2b2.08+</i>	<i>spbbp2b2.08+</i>	conserved fungal protein	NONE	-0.18	-0.05	2.80	2.69	-6.11	-0.16	-0.33	-0.50
<i>spcc1020.10+</i>	<i>oca2+</i>	serine/threonine protein kinase Oca2	YNL183C, YDL214C, YDL025C, YOR267C, YDL214C, YDL025C, YOR267C	0.32	1.15	0.16	1.34	-0.32	1.08	1.94	1.45
<i>spcc1235.01+</i>	<i>spcc1235.01+</i>	sequence orphan	NONE	0.28	-0.25	0.96	0.07	-1.81	-0.28	2.21	1.63
<i>spcc1235.03+</i>	<i>spcc1235.03+</i>	SMR domain protein, possibly involved in DNA repair	YKL090W	0.01	0.00	0.16	0.08	-1.42	0.13	1.35	1.23
<i>spcc1281.04+</i>	<i>spcc1281.04+</i>	pyridoxal reductase (predicted)	YPR127W	0.27	0.20	0.09	0.45	-0.23	-0.23	2.15	2.30
<i>spcc1393.12+</i>	<i>spcc1393.12+</i>	sequence orphan	NONE	0.10	-0.13	-0.27	-0.75	-0.53	-0.78	3.20	1.12
<i>spcc16a11.15c+</i>	<i>spcc16a11.15c+</i>	sequence orphan	NONE	0.29	-0.06	-0.02	-0.99	-2.89	-0.74	3.90	1.22
<i>spcc1739.06c+</i>	<i>spcc1739.06c+</i>	uroporphyrin methyltransferase (predicted)	YKR069W	-1.21	-0.44	2.38	1.30	-4.20	-0.57	-1.04	-1.90
<i>spcc1795.06+</i>	<i>map2+</i>	P-factor pheromone Map2	NONE	-0.08	0.10	-0.23	0.09	0.14	0.14	1.80	1.19
<i>spcc330.06c+</i>	<i>spcc330.06c+</i>	thioredoxin peroxidase (predicted)	YLR109W	0.86	1.34	0.56	1.15	-1.05	0.86	0.58	0.93
<i>spcc338.18+</i>	<i>spcc338.18+</i>	sequence orphan	NONE	0.17	0.61	-0.72	-0.84	-0.58	-0.58	3.32	1.44
<i>spcc569.09+</i>	<i>spcc569.09+</i>	sequence orphan	NONE	-0.22	0.29	0.17	0.01	-0.36	-0.16	2.31	1.23
<i>spcc576.16c+</i>	<i>wf12+</i>	pseudogene wtf element WF22	NONE	0.05	0.15	0.17	-0.13	-0.23	-0.26	1.49	1.58
<i>spcc74.09+</i>	<i>mug2+</i>	RNA-binding protein, rrm type	NONE	0.28	0.11	0.28	0.44	-0.16	0.18	1.53	1.03
<i>spcc757.03c+</i>	<i>spcc757.03c+</i>	ThJ domain protein	YDR533C, YMR322C, YOR391C, YPL280W	0.40	-0.63	0.23	-0.62	-1.64	-1.46	3.56	1.56
<i>spcc757.07c+</i>	<i>ctf1+</i>	catalase	YDR256C, YGR088W	0.55	-0.15	-0.80	-0.32	-1.40	-0.17	3.26	1.49
<i>spcc825.04c+</i>	<i>naa40+</i>	histone N-acetyltransferase Naa40 (predicted)	YMR069W	0.09	-0.03	-0.08	0.28	0.61	0.09	1.44	1.41
<i>spcc830.07c+</i>	<i>psr1+</i>	DNAJ domain protein, involved in translation initiation Psr1	YNL007C	0.73	0.87	0.06	0.43	-3.30	0.39	0.72	0.43
<i>spcp31f0.06+</i>	<i>mug190+</i>	C2 domain protein	YOR086C, YNL087W, YML072C	0.85	-0.16	0.44	-0.37	-2.77	-0.62	2.55	1.10
<i>spcpb16a4.06c+</i>	<i>spcpb16a4.06c+</i>	sequence orphan	NONE	0.22	0.44	0.01	-0.36	-4.29	-0.92	3.46	1.28
<i>spcpb1c11.03+</i>	<i>spcpb1c11.03+</i>	cysteine transporter (predicted)	YLL055W, YOL163W, YAL067C, YLL186C, YJR162W, YOL162W, YGR260W, YLR004C, YCR028C	-0.11	-0.52	1.37	1.89	-0.98	0.22	-1.94	-1.28

^aGenes that are regulated in multiple stress conditions by Pho7 are indicated in **bold**

^bDescriptions for each gene product are provided by the *S. pombe* GeneDB: <http://old.genedb.org/genedb/pombe/index.jsp> (Hietrz-Fowler et al., 2004)

^cOrthologs were identified using the Fungal Orthogroups Repository: <http://www.broadinstitute.org/regex/orthogroups> (Wapinski et al., 2006)

^dAll comparisons between stressed and non-stressed conditions (i.e. [-PI/+PI]) were done in a *pho7⁺ backgroud*. Genes passing the 2σ + median threshold for each array condition are have their values indicated in **bold**. +Pi: 10 mM H₂KPO₄, -Pi: 0 mM H₂KPO₄, +Fe: 100 μM Fe(III)Cl₃, -Fe: 250 μM DIP, +Cu: 100 μM Cu(II)SO₄, -Cu: 100 μM Cu(II)SO₄, -1.2M: 1.2 M NaCl, 0.1M: 0.1 M NaCl

Pho7 is a Bi-functional Regulator of Coding, Non-Coding, and Anti-Sense Transcription

Our initial ChIP-Seq analysis reveals a widespread enrichment of Pho7 throughout the *S. pombe* genome (Figure 3.3C). From the transcriptional microarrays we know that a subset of these peaks are located in the promoters of Pi- and *pho7*⁺-responsive genes, providing a correlation between Pho7 enrichment and *pho7*⁺-dependent induction. Yet, Pho7 is obviously bound to many regions not implicated in the transcriptional microarray analysis. We identified one possible explanation for this discrepancy through our analysis of additional *pho7*⁺, non-Pi dependent environmental stress responses (Figure 3.6 and 3.7). Alternatively, it is possible that we missed *pho7*⁺-regulated transcripts because they were not present in our microarray probe-set. This would certainly be the case for non-coding or anti-sense transcripts. Given that non-coding and anti-sense transcription plays a significant role in *S. pombe* global gene regulation (Dutrow et al., 2008; Ni et al., 2010; Wilhelm et al., 2008) we thought the widespread enrichment of Pho7 might suggest an involvement in non-coding transcriptional regulation.

To determine which transcripts might be subjected to *pho7*⁺ regulation, we first took the mid-point of each peak and identified transcriptional features within 800 bp on either the plus or minus strand. There are 968 transcriptional features within 800 bp downstream of one or more of the 1476 ChIP-Seq peaks in no-Pi conditions. ~13% (123) of the total Pho7-bound features are annotated as non-coding RNAs. Additionally, 596 annotated ORFs have at least one peak either 800 bp downstream or within gene coding regions that might drive transcription from the anti-sense strand. To ascertain the

actual prevalence of Pi- and *pho7*⁺-dependent transcription from these sites we performed direction specific global sequencing on the *S. pombe* transcriptome. Briefly, we collected total RNA from *pho7*⁺ cells grown in high-Pi or no-Pi media or from *pho7*Δ cells grown in no-Pi media. Total RNA was purified, converted into cDNA, and subjected to high-throughput sequencing (RNA-Seq, see materials and methods). Transcriptional output measured by RNA-Seq was then compared to Pho7 enrichment to define regions bound and regulated by Pho7.

We confirm that the core PHO regulon is represented in our dataset and behaves in a manner similar to that previously characterized. *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺ show significant up-regulation during phosphate starvation, and induction is dependent on *pho7*⁺ (Figure 3.9A, B). Next, we compared transcript abundance for genes with one or more Pho7 upstream binding sites – from *pho7*⁺ cells grown in no-Pi media – with the differential abundance in those same genes between *pho7*⁺ and *pho7*Δ backgrounds – also in no-Pi media (Figure 3.8A). As expected, the core PHO genes form a high transcript abundance cluster (upper right quadrant) that is significantly influenced by the loss of *pho7*⁺. We also find a set of 95 additional unique Pho7-bound genes that display a decrease in transcript abundance in the *pho7*Δ background (see materials and methods for threshold determination). Of the 98 Pho7-bound, *pho7*⁺-dependent genes, 47 are upregulated in response to phosphate starvation. Four of the genes that have a Pi and *pho7*⁺ component from the microarray analysis are present in this set (*pho1*⁺, *spbc8e4.01c*⁺, *spbc1271.09*⁺, and *spbpb2b2.05*⁺). The other four microarray-defined genes do not have Pho7 bound within 800 bp of their promoters and are not included in this analysis. Interestingly, we found an additional *PHO84* ortholog

(*spac23d3.12*⁺) with a Pho7 binding site whose regulation is *pho7*⁺-dependent in response to phosphate starvation (Figure 3.9C). *spac23d3.12*⁺ is right at the cutoff for the microarray analysis, which explains why it was previously overlooked.

There are 40 ncRNAs and 281 anti-sense regions with a Pho7-binding site whose transcription is dependent on *pho7*⁺ (Figure 3.8B and C). Nine of the ncRNAs are also regulated in a phosphate-dependent manner. Two representatives of Pho7-regulated ncRNA transcripts are shown in Figures 3.9D and E. The first, *spncRNA.72*⁺, is an anti-sense transcript for *eta2*⁺ (Figure 3.9D). However, the function of this ncRNA does not appear to be *eta2*⁺ inhibition at the level of transcription, as the *eta2*⁺ transcript abundance is the same in both the *pho7*⁺ and *pho7* Δ background. The second, *spncRNA.779*⁺, is located within a cluster of ncRNAs (Figure 3.9E). This region is flanked by two Pho7-binding sites and *pho7*⁺ is required to keep bi-directional transcription active in both high-Pi and no-Pi conditions. The function of *pho7*⁺-regulated ncRNA transcription within the cell is unknown. Moreover, our results with alternative non-Pi dependent stress response networks suggest that the non-coding and anti-sense transcripts unaffected by Pho7 deletion during phosphate starvation might be instead regulated in a *pho7*⁺-dependent manner in response to other stressors.

The presence of a Pho7-binding site is insufficient for defining Pi- and *pho7*⁺-dependent transcriptional boundaries. This led us to wonder: does the change in Pho7 enrichment during phosphate starvation correlate with a corresponding change in *pho7*⁺-regulated transcription? For each of the significantly enriched peaks we compared the change in peak area between *pho7*⁺ samples grown in high-Pi or no-Pi media with the change in transcript abundance from RNA-Seq (Figure 3.8D). We do not

see a correlation between these two variables. Further, Pho7-TAP peak size alone is not correlated with the magnitude of *pho7*⁺-dependent transcript abundance (Figure 3.8E). This is not surprising, as we know that *pho7*⁺-dependent transcription is combinatorially controlled by at least one additional factor (*csk1*⁺) and by at least four distinct stressors (Pi, Fe, osmolarity, and carbon utilization). To build a complete model linking Pho7 enrichment to transcriptional output we will need to understand better how the *cis*- and *trans*-regulatory elements control Pho7 activity. Future experiments combining Pho7 enrichment with additional regulatory networks should allow us to define the elements that lead to coding and non-coding output.

Finally, our RNA-Seq analysis using Pho7-bound transcriptional features highlights a set of transcripts repressed by *pho7*⁺ activity (Figure 3.8A-C, found under the -2-F.C. threshold). We had noticed in our microarrays that certain ORFs were down-regulated in response to phosphate starvation and up-regulated in response to *pho7*⁺ deletion, but we thought these changes were caused by an indirect general stress response. Combining our Pho7 ChIP-Seq data with our RNA-Seq data suggests that this might not be the case. Of the 968 identified Pho7-bound features, 302 are up-regulated in the *pho7* Δ strain compared to the *pho7*⁺ strain. Fifty-four are also repressed in response to phosphate starvation. A representative of this class, *ptr2*⁺, is shown in Figure 3.9F. Whether this *pho7*⁺ bi-functionality is a consequence of active repression or the passive loss of transcriptional activation in the opposite direction is unknown. This bi-functional behavior is also observed for 7 ncRNAs, although what role this repression plays in phosphate starvation is unknown. Both *snf5*⁺ and *snf22*⁺ function as transcriptional activators and repressors *in vivo* (Monahan et al., 2008); therefore, *pho7*⁺

bi-functionality might simply be a consequence of its interaction with the SWI/SNF machinery. Future experiments defining Pho7 protein-protein interactions will allow us to address this question in further mechanistic detail.

Figure 3.8. *pho7*⁺ Regulates Non-Coding Transcription in Both a Pi-dependent and Pi-independent Manner.

(A) Shown is a scatter plot comparing the RPKM score (\log_{10}) for each annotated gene from a *pho7*⁺ strain incubated in no-Pi conditions versus the differential expression (\log_2 scale, using RPKM scores for each gene) between the *pho7*⁺ and *pho7* Δ strains in no-Pi media. Transcript abundance was measured using the RNA-Seq procedure detailed in the text. Only genes within 800 bp of a Pho7-TAP ChIP-Seq peak are plotted. Genes up-regulated greater than 14-fold are highlighted. Thresholds at (+)2 and (-)2-F.C. are indicated. (B) As in A, plotting ncRNAs instead. (C) As in A, plotting anti-sense regions instead. (D) Scatter plot comparing the change in Pho7-TAP enrichment between no- and high-Pi conditions (\log_2 scale) versus the differential expression (\log_2 scale, using RPKM scores for each transcript) for a *pho7*⁺ strain grown in no- or high-Pi media. RPKM scores for genes, ncRNAs, and anti-sense transcripts within 800 bp of a Pho7-TAP ChIP-Seq peak were determined using RNA-Seq. Thresholds at (+)2 and (-)2-F.C. (solid lines) as well as (+)4 and (-)4-F.C. (dashed lines) are indicated. (E) Scatter plot comparing the height of each Pho7-TAP ChIP-Seq peak (\log_{10}) versus the differential expression (\log_2 scale, RPKM) between the *pho7*⁺ and *pho7* Δ strains in no-Pi media. RPKM scores for genes, ncRNAs, and anti-sense transcripts within 800 bp of a Pho7-TAP ChIP-Seq peak were determined using RNA-Seq. Thresholds for gene expression at (+)2 and (-)2-F.C. (solid lines) as well as peak height $\geq 1\sigma$ or $2\sigma + \text{median peak height}$ (dashed lines) are indicated.

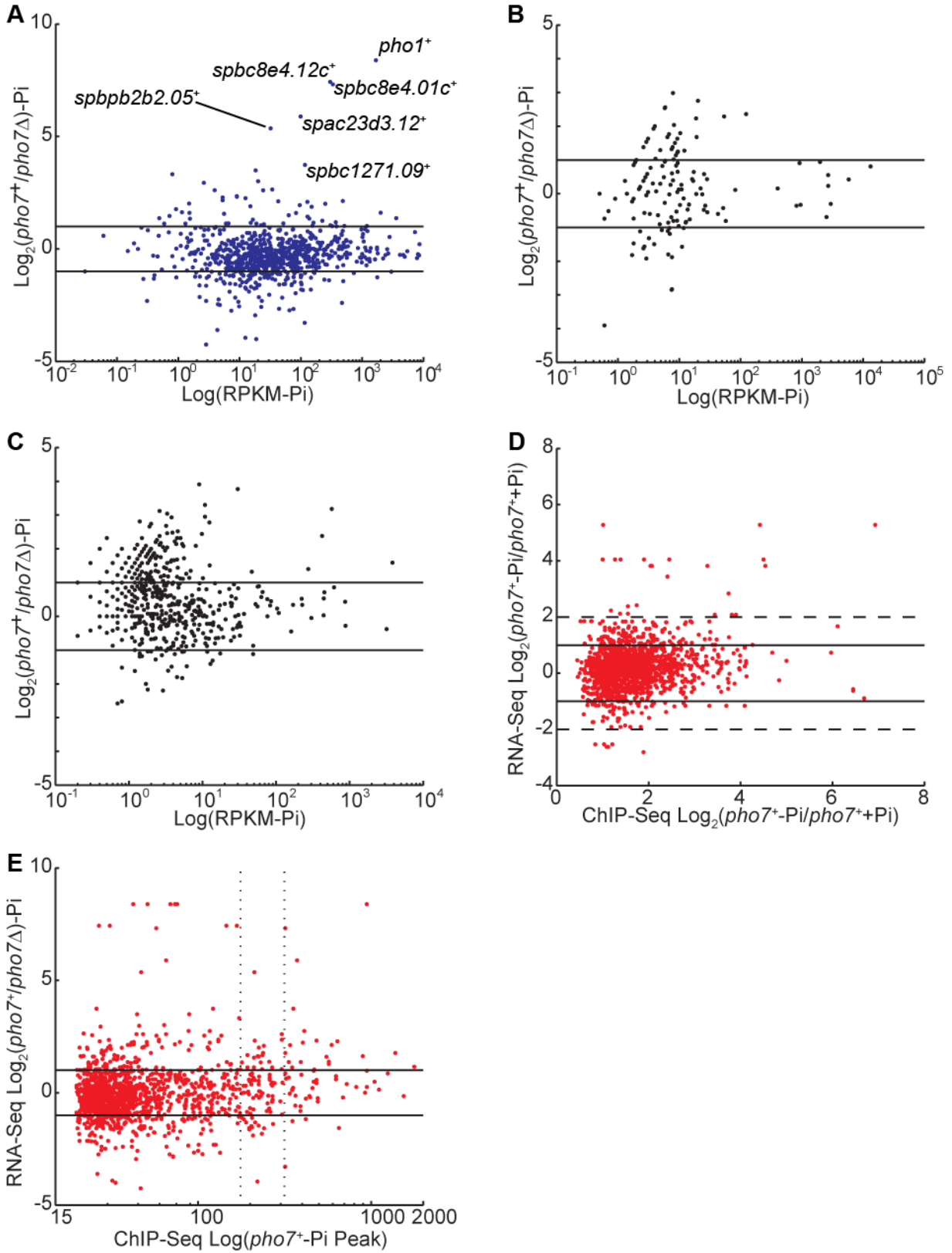


Figure 3.8. Continued

Figure 3.9. RNA-Seq Profiles for Representative *pho7*⁺ Regulated Transcripts.

Shown are RNA-Seq (top) and ChIP-Seq (bottom) profiles for representative classes of *pho7*⁺ regulated transcripts. *pho7*⁺ cells were grown in either high-Pi (blue) or no-Pi (red) conditions and transcript abundance was compared to *pho7* Δ cells grown in no-Pi conditions (black). Transcript abundance was measured via RNA-Seq as described in the text. RNA-Seq signals are normalized to nucleotide reads per million sequenced reads (RPM). Each RNA-Seq track for a given transcript is normalized to the same min-max scale. RPM scores are cropped to allow visualization of low-abundance signal. Signal from the plus strand is plotted in the positive direction, while the minus strand is plotted in the negative. Wild-type cells containing Pho7-TAP were grown in either high-Pi (blue) or no-Pi (red) conditions and ChIP-Seq libraries were prepared from purified DNA. For comparison, the ChIP-Seq signal from mock (black) cells grown in no-Pi is included. ChIP-Seq reads were normalized as previously described. The gene product of interest is plotted based on transcript direction with the plus (+) strand above and the minus (-) strand below. (A) *pho1*⁺ and *spbc8e4.01*⁺ (*PHO84* ortholog). (B) *spbc1271.09c*⁺ (*GIT1* ortholog). (C) *spac23d3.12*⁺ (*PHO84* ortholog). (D) The anti-sense transcript, *ncRNA.72*⁺. (E) A set of clustered intergenic ncRNAs. (F) A *pho7*⁺ repressed transcript, *ptr2*⁺

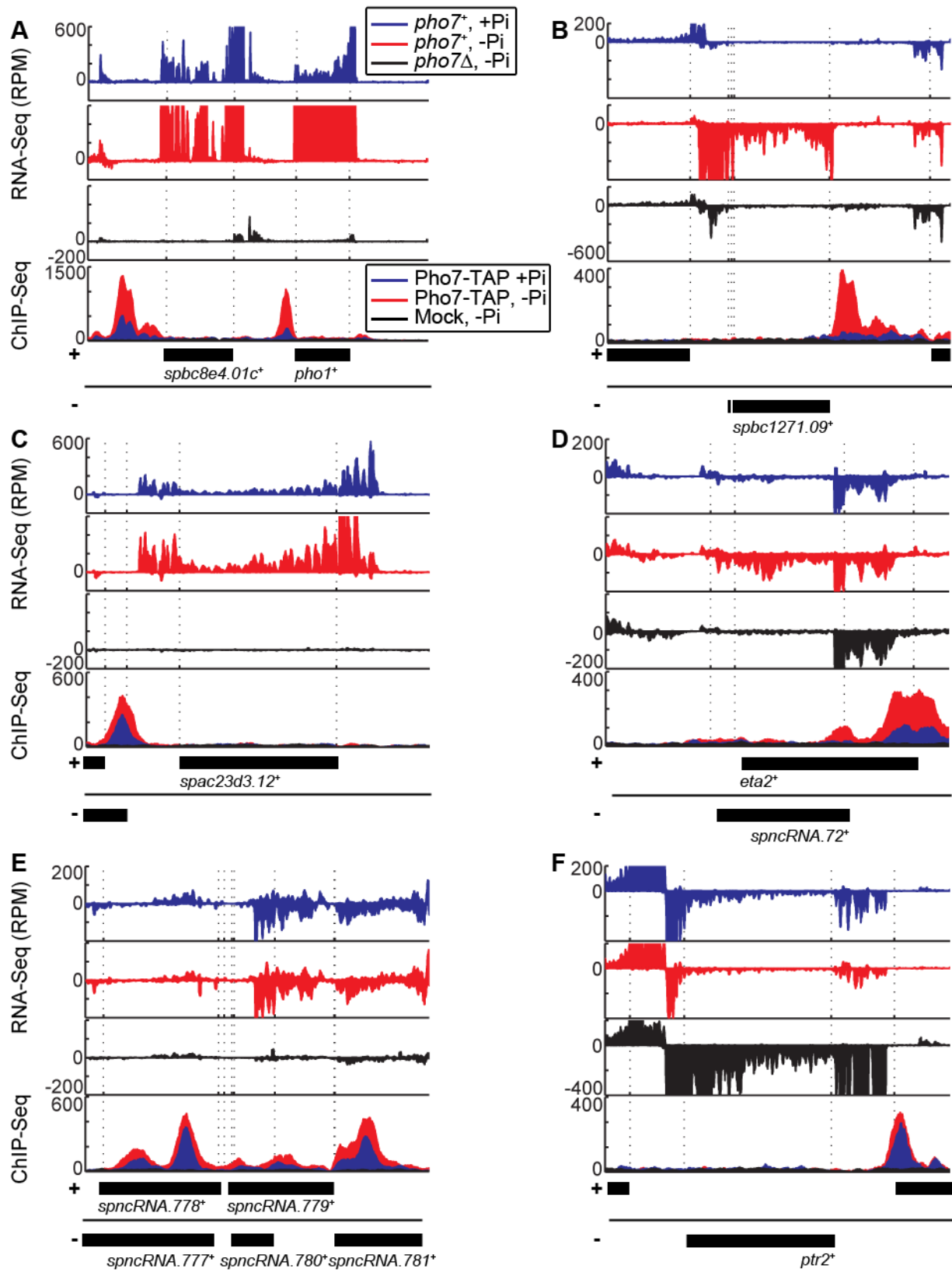


Figure 3.9. Continued

CONCLUSIONS

Using a number of whole-genome techniques we defined and characterized the gene regulatory network in *S. pombe* responsible for coordinating the response to inorganic phosphate starvation. There are two distinct temporal dimensions in the PHO pathway in *S. pombe*: a fast response concerned with immediately harvesting inorganic phosphate from the environment and transporting it into the cell, and a slower one dedicated to harvesting alternate forms of phosphate and regulating internal polyphosphate stores. Within the fast response we define a core PHO regulon comprised of the *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺ genes whose induction in response to phosphate starvation, and regulatory behavior, has been conserved between the *S. pombe* and *S. cerevisiae*.

Unlike the PHO response in *S. cerevisiae*, however, the positive regulator in *S. pombe* is bound to the *pho1*⁺ and *spbc8e4.01c*⁺ promoters irrespective of external phosphate availability. This interaction by Pho7 with the UAS leads to a basal expression of the secreted acid phosphatase in high-Pi conditions. Csk1 prevents full activation of Pho7 during phosphate depletion via an interaction at a site between -1kb and -2kb in the *pho1*⁺ promoter (the URS). During phosphate starvation this inhibition is relieved (through an unknown mechanism) and additional Pho7 is recruited to a number of sites throughout the genome, causing further induction of *pho7*⁺-dependent genes. In previous work, Csk1 was shown to regulate transcription by activating the positive transcription elongation factor b (P-TEFb) ortholog, Cdk9 (Pei et al., 2006). Cdk9 coordinates transcript elongation and processing, and its full activation by Csk1 leads to an increase in CTD kinase activity (Garriga and Grana, 2004). How this generally

positive regulatory network is switched to an inhibitory role in the PHO system remains an open question.

We were also surprised to find that Pho7 was bound throughout the genome in both high-Pi and no-Pi conditions. We had thought based on previous evidence that Pho7, like Pho4, would be specific to the PHO response. Instead we demonstrate that Pho7 binds within the promoters of additional stress responsive genes and plays a role in iron starvation, osmotic, and alternative carbon utilization stress. Its role in copper starvation stress is limited, suggesting that *pho7*⁺ activation in response to nutrient limitation is discriminatory. There must exist some mechanism to either direct Pho7 to the proper location for inducing the correct genes or activate Pho7 at only the appropriate locations (or some mixture of both). Additionally, widespread Pho7 binding occurs near transcription start sites for non-coding and anti-sense transcripts. We identified a cohort of non-coding RNAs that are regulated in a *pho7*⁺-dependent manner, though the function of these ncRNAs is unknown.

In *S. cerevisiae*, the osmotic, oxidative, and glucose limitation stress responses are mediated by the transcription factor Msn2 (Martinez-Pastor et al., 1996). In normal conditions, Msn2 is phosphorylated and its entry into the nucleus is limited (Gorner et al., 1998). Each individual stress causes a differential pattern of nuclear entry leading to a specified transcriptional output (Hao and O'Shea, 2011). Given that Pho7 is bound to the genome constitutively, we do not expect that nuclear exclusion will play as large a role as it does with Msn2 regulation, but it remains possible that differential post-translational modifications are responsible for this combinatorial activation by Pho7. Moreover, our evidence suggests that at least two members of the SWI/SNF chromatin

remodeling machinery (Snf5, Snf22) either play a direct role in activating the PHO response or are recruited as part of a regulated complex. Pho7 may be playing a more passive role in regulation, with additional factors determining Pho7 genomic localization. Perhaps, the make-up of different chromatin remodeling complexes containing Pho7 determines Pho7 activity.

Nonetheless, we have demonstrated that within the evolutionary parallel signal transduction networks that comprise the PHO pathway there exists a core PHO regulon. The specific mechanisms involved in regulating the PHO response in *S. cerevisiae* and *S. pombe* show remarkable plasticity. An interesting area for future research centers on the environmental factors that contributed to the development of these two parallel networks. Why is the PHO response in *S. pombe* under the control of a general stress transcription factor, Pho7, while *S. cerevisiae* has developed the phosphate starvation specific pathway for Pho4 activation? What are the environmental pressures that favor a “leaky” response in *S. pombe* and not one in *S. cerevisiae*? Broadly speaking, our study provides a framework for determining the fundamental requirements for regulating phosphate homeostasis in ascomycota and the specific points in the signal transduction pathway that can be altered as conditions merit.

MATERIALS AND METHODS

Growth Conditions and Strains

S. pombe cells were maintained in previously described YES or EMM media (Forsburg and Rhind, 2006). The yeast strains used were: DP1 (972 *h*⁻), DP81 (*pho7ΔKANMX6* 972 *h*⁻), DP94 (Pho7-TAPKANMX6 972 *h*⁻), DP106 (*csk1ΔNATMX6* 972 *h*⁻), DP113 (*pho7ΔKANMX6 csk1ΔNATMX6* 972 *h*⁻), DP115 (Pho7-TAPKANMX6

csk1ΔNATMX6 ura4⁻ 972 h⁺), and FWP289 (Monahan et al., 2008). The functionality of the Pho7-TAP allele was confirmed by both liquid phosphatase and RT-qPCR analysis and it behaves as *pho7⁺*. To tag Pho7 and delete *csk1⁺* we utilized a PCR fragment containing the marker of interest flanked by homologous regions for the specific gene target. Cells were transformed with lithium acetate and polyethylene glycol 8000 (Forsburg and Rhind, 2006). Primers used for deletion or tagging are found in Table S3.1. To provide consistency with previously published results for inorganic phosphate starvation, all starvation experiments were conducted with cells incubated in a 90%SD-10%EMM media, which has been previously described (Henry et al., 2011).

Microarray Analysis and Data Processing

Strains were grown in 90%SD-10%EMM medium containing 10 mM KH_2PO_4 (high-Pi) at 30°C until they reached early-log phase ($\text{OD}_{600}=0.15-0.25$). Cells were collected via filtration, washed twice, transferred to fresh media lacking Pi (no-Pi), and grown at 30°C for up to 4 hours. Immediately prior to starvation ($t=0$), 20 mL of cells were added to 30 mL of methanol kept at -65°C to prevent further transcription or RNA degradation. At 30, 60, 120, and 240 minutes post-starvation this process was repeated. Cells were left in methanol for 10 minutes, pelleted, washed quickly in autoclaved water, re-suspended in 750 uL of RNAlater (Ambion), and snap-frozen in liquid nitrogen. RNA was extracted using the RNeasy Mini kit (Qiagen). cDNA was generated in a reverse transcriptase reaction using 10 µg total RNA with a 1:1 mixture of oligo-dT and random hexamer primers (Operon) and a 2:3 ratio of amino-allyl-dUTP:dTTP (Sigma). Superscript II RT (Invitrogen) was added and the reaction mixture was incubated at 42°C for 2.5 hours. cDNA was purified using a PCR purification kit

(Qiagen) after completing hydrolysis of remaining RNA. An equal amount of cDNA from each time point was pooled to provide the reference sample. Purified cDNA samples were labeled using N-hydroxyl succinamide esters of either Cy3 or Cy5 dyes (GE Biosciences). 300 ng of the Cy3 (each individual time point) and 300 ng of the Cy5 (pooled reference) labeled sample was competitively hybridized to custom Agilent 8x15K *S. pombe* two-color expression microarrays (probe design by the Regev group) in 2xGEx Hybridization Buffer (Hi-RPM) (Agilent) for 17 hours at 60°C. Microarrays were washed and immediately scanned using an Axon 400B scanner (Vijayan et al., 2009). The mean intensity of each spot in the Cy3/Cy5 channels was extracted using the GenePix 5.1 software, followed by lowess and quantile normalization performed with the MATLAB bioinformatics toolbox. Expression ratios for each time point, x , were normalized to $t=0$ ($\text{Log}_2[\text{Cy3}_{t=x}/\text{Cy5}_{\text{pool}}] - \text{Log}_2[\text{Cy3}_{t=0}/\text{Cy5}_{\text{pool}}]$) and thresholds for induced genes were set at $\geq 2\sigma + \text{median-F.C.}$ for each time point (1.02-F.C. at 120 minutes, 1.24-F.C. at 240 minutes).

To determine the extent of *pho7*⁺ and *csk1*⁺ regulation within the PHO response we grew the relevant strains (DP1, DP81, DP106, DP113) as described above, with the exception that cells were split into either high-Pi or no-Pi media and grown for 2 hours prior to RNA collection. Two independent biological replicates were performed for the following comparisons: DP1 in high-Pi vs. DP1 in no-Pi, DP81 vs. DP1 in no-Pi, DP1 vs. DP81 in high-Pi, and DP106 vs. DP1 in high-Pi. For each of these arrays the two probes used to detect each ORF were treated as separate data points (for 4 total) and p-values were determined using a student's t-test with a one-tailed distribution against the null hypothesis in the MATLAB software. Thresholds were set at $\geq 1.8\text{-F.C.}$ to facilitate

comparison with the previously characterized *S. cerevisiae* data set (Zhou and O'Shea, 2011). Genes passing the induction threshold also had to pass a p-value threshold of ≤ 0.01 . For the epistasis analysis (Figure S3.1), thresholds for individual arrays were set at $\geq 2\sigma + \text{median-F.C.}$ Clustering analysis was completed using k-means clustering in the Cluster 3.0 program (Eisen et al., 1998) after empirically determining the optimal number of clusters using the MATLAB bioinformatics toolbox.

Chromatin Immunoprecipitation

ChIP experiments were carried out using DP1 (mock), DP94 (Pho7-TAP), DP115 (Pho7-TAP *csk1Δ*), and FWP289 (Snf22-TAP) strains as previously described (Lam et al., 2008; Monahan et al., 2008), with the following modifications. Cells were grown to early log-phase ($OD_{600} \sim 0.18$) in high-Pi media at 30°C and split into either 200 mL of high-Pi or no-Pi media and grown for 2 hours. Formaldehyde (Sigma) was added to a final concentration of 1% (v/v) to cross-link chromatin, and the reaction was allowed to proceed for 15 minutes. Glycine (Sigma) was then added to a final concentration of 125 mM and incubated for 5 minutes to quench cross-linking. Cells were lysed by bead beating (6 x 2 min on, 2 min off) and chromatin was sheared to 300-600 bp fragments using a Misonix Sonicator 3000. Immunoprecipitation was performed with 100 μ L of Protein G Dynabeads (Invitrogen) coupled to 4 μ L of anti-Protein A antibody (Sigma). Protein concentrations were measured using a Bradford Assay (Bio-Rad) and 650 μ g of each sample was subjected to IP. 50 μ g was set aside to measure total DNA (input). ChIP DNA was quantified using quantitative PCR on a Stratagene MX3000. qPCR was completed using a standard curve generated by a serial dilution of genomic DNA. Abundance of Pho7-TAP or Snf22-TAP bound to DNA was measured in respect to input

and was normalized against the signal from the *spo3*⁺ promoter. Enrichment for a given primer set, Y, is defined as: $[(Y_{\text{ChIP}}/Y_{\text{Input}})/(spo3^+_{\text{ChIP}}/spo3^+_{\text{Input}})]$. Values calculated represent the mean of three biological replicates \pm standard error. Primer sequences used are listed in Table S3.1.

***pho1*⁺ Promoter Deletion Analysis**

Segments of the *pho1*⁺ promoter were amplified using PCR and subjected to restriction digest. Fragments were ligated into a *yfp*⁺ plasmid containing the selectable *ura*⁺ marker, creating a *pho1*⁺pr-*yfp*⁺ fusion. Plasmids were transformed into *pho7*⁺, *pho7* Δ , or *csk1* Δ *ura*⁻ backgrounds using lithium acetate and polyethylene glycol 8000. Cells were selected based on their ability to grow in EMM-ura media. Cells containing the various plasmids were grown overnight in high-Pi media at 30°C. The next morning cells were diluted and grown to mid-log phase before being pelleted, washed, and then transferred into either high-Pi or no-Pi media for 4 hours. Total RNA was extracted from harvested cells using an acid phenol protocol (Kerwin and Wykoff, 2009). cDNA was generated from 1 μ g total RNA using the iScript cDNA synthesis kit (Bio-Rad) and was purified using a PCR clean up kit (Qiagen). Real-time PCR was performed using a Stratagene MX3000. Amplification was measured relative to a standard curve generated by a serial dilution of genomic DNA. *pho1*⁺ and *yfp*⁺ transcript levels were normalized to *act1*⁺. Shown is the average from three independent replicates \pm standard error. Primers used for amplification are listed in Table S3.1.

For fluorescent microscopy, cells containing the various plasmids were grown overnight in high-Pi media at 30°C. They were transferred to fresh high-Pi media, grown to mid-log phase, and split into either high-Pi or no-Pi media and cultured for 4 hours.

Cells were then harvested and subjected to fluorescent and confocal microscopy.

Representative images are shown in Figure S3.6.

Chromatin Immunoprecipitation of Pho7-TAP with High-Throughput Sequencing (ChIP-Seq)

ChIP-Seq was performed on the DP1, DP94, and DP115 strains as previously described (Vijayan et al., 2011; Zhou and O'Shea, 2011). Briefly, following the generation of ChIP lysate as detailed above, three aliquots of 650 μ g soluble protein were subject to immunoprecipitation and pooled just prior to elution from the beads. Samples were processed following the Illumina HTS guidelines with libraries of 200-300 bp selected via 2% agarose DNA gels. Libraries were amplified by PCR and purity was determined using an Agilent High-Sensitivity DNA kit on an Agilent Bioanalyzer. Libraries were sequenced on an Illumina HighSeq 2000 and 50bp reads were aligned to the *S. pombe* 972 *h*⁻ genome using ELAND. We obtained between 16 million and 54 million reads on average from our samples. Uniquely aligned reads were extended 80 bp from the read start site to cover the average length of insert as determined by the Agilent Bioanalyzer.

To determine which of our enriched regions were actually attributable to a Pho7-TAP binding event we used a modified method from (Vijayan et al., 2011). For each condition analyzed we set a lower threshold for peak discovery equal to the genomic average of reads per base. We set the upper threshold equal to the highest observed read count within the given sample. Using 380 equal increments between these two thresholds we defined peaks that were larger than 100 nucleotides and separated by at least 20 nucleotides. Peaks were compiled at the highest threshold at which they met

those standards and full peaks were required to be at least 150 nucleotides distant from the nearest neighbor. Statistical analysis comparing sample enrichment to mock enrichment was performed in MATLAB using previously described methods (Vijayan et al., 2011). Peaks used for subsequent analysis had a ≥ 2 -fold enrichment over the genome average and a p-value ≤ 0.005 compared to the mock sample.

In Silico Motif Discovery

The set of significantly enriched peaks we identified in the ChIP-Seq analysis were smoothed and fit to a binomial distribution using the MATLAB bioinformatics toolbox. Maximums from each distribution were identified and FASTA sequences for 50-200 bp surrounding the max were submitted to MEME for motif discovery (Bailey et al., 2009). We attempted many iterations varying the requirements for the number of motifs within each sequence, the length of sequence, the proximity to known transcription factor binding sites, gap penalty, *pho7*⁺-regulation from microarray analysis, etc. We were unable to produce a motif through this analysis that was significantly enriched in more than 15% of the top 200 peaks.

***pho7*⁺ Regulation in Additional Stress Response Pathways**

For each individual stress response, initial cultures of DP1 and DP81 were grown in 90%SD-10%EMM media containing 10 mM KH₂PO₄ to early-log phase at 30°C. Cells were collected, washed twice with autoclaved water, and split into the following conditions (all modifications of the high-Pi media): +Pi – 10 mM KH₂PO₄, -Pi – 0 mM KH₂PO₄, +Fe – 100 uM Fe(III)Cl₃, -Fe – 250 uM 2-2'-dipyridine (DIP) (Sigma), +Cu – 100 uM Cu(II)SO₄, -Cu – 100 uM bathocuproine disulphonate (BCS) (Sigma), Osmotic Shift – 1.2 M NaCl instead of 0.1M NaCl, and Carbon Switch – 2% glycerol/1% ethanol

instead of 2% glucose. Cells were grown for 2 hours and harvested as described above. Recovered RNA was converted into cDNA using the iScript cDNA synthesis kit (Bio-Rad) and subjected to RT-qPCR. Amplification of the *gpd1*⁺, *fiol*⁺, *ctr4*⁺, *hxx2*⁺, and *pho1*⁺ transcripts were measured for the three independent replicates and transcript abundance was normalized to *act1*⁺. Shown is the average \pm standard error. Primers used in the RT-qPCR analysis can be found in Table S3.1.

Extracted RNA was also subjected to microarray analysis as detailed above. Expression from *pho7*⁺ cells in replete conditions was compared to that in stress conditions for each individual stressor to determine the base set of genes that respond in each given stress. The dependence upon *pho7*⁺ was determined by comparing the levels of induction in *pho7*⁺ cells in stress to induction in *pho7* Δ cells in stress. Following extraction of Cy3-Cy5 fluorescence intensity using the GenePix 5.1 software and normalization through the MATLAB bioinformatics toolbox, expression ranges for each individual condition were normalized to a min-max of -1 to 1 to assist in cluster analysis. Genes that had a *pho7*⁺-regulated stress response were identified as having induction values $\geq [2\sigma + \text{median}]$ -F.C. between both the *pho7*⁺ replete vs. stress and *pho7*⁺ stress vs. *pho7* Δ stress conditions. This set of 124 unique genes was then grouped using k-means clustering. The optimum number of clusters for subsequent analysis was determined empirically using the MATLAB bioinformatics toolbox and we performed k-means clustering on the combined data set using Cluster 3.0 (Eisen et al., 1998).

High-Throughput Sequencing of the *S. pombe* Transcriptome (RNA-Seq)

RNA-Seq was performed on samples from *pho7*⁺ cells grown in high-Pi and no-Pi conditions and *pho7* Δ cells grown in no-Pi conditions using previously described

methods (Ingolia et al., 2009). Briefly, DP1 or DP81 strains were grown to early-log phase in high-Pi media at 30°C. Cells were filtered, washed twice with no-Pi media, and split into either high-Pi or no-Pi media and grown for another 2 hours. Samples were frozen in 2:3 ratios of cells: cold methanol (-65°C) to prevent further transcription or degradation. RNA was extracted from harvested cells using warm acid phenol and ethanol precipitation. Total RNA was treated with oligo-dT Dynabeads (Invitrogen) to enrich for poly-A tailed RNA and to decrease the abundance of rRNA. Enriched RNA was fragmented using an alkaline fragmentation solution and transcripts between 50-100 nucleotides were purified from a 15% TBE-Urea polyacrylamide gel. RNA libraries were then subjected to dephosphorylation using T4 polynucleotide kinase (NEB) and poly-A tailing using poly-A polymerase (NEB). RNA was mixed with template DNA containing linker sequences for PCR amplification of the first-strand cDNA product, an oligo-dT(20) primer, and two degenerate nucleotides. Following reverse transcription with Superscript II (Invitrogen), the cDNA-containing template was circularized and subjected to PCR. Library purity was analyzed using the Agilent Bioanalyzer. Libraries were sequenced on an Illumina High-Seq 2000 and 50 bp single-end reads were used for all subsequent analysis.

Library output was first trimmed for any poly-A stretch and sequences were shortened to between 18-23 nucleotides. Our samples generated an average of between 130 and 170 million trimmed reads. Reads aligning to any region of rRNA were then removed from the dataset and the remaining sequences were aligned to the *S. pombe* 972 h⁻ genome using Bowtie (Langmead et al., 2009). On average this left us with between 30 and 50 million uniquely mapped, non-rRNA reads. Utilizing the most

current annotation of the *S. pombe* genome (Wood et al., 2012), we calculated reads per kilobase of exon model per million mapped reads (RPKM) scores for each annotated feature in both the sense and anti-sense direction (Mortazavi et al., 2008). Transcript enrichment between high-Pi and no-Pi conditions in the *pho7⁺* background and between *pho7⁺* and *pho7 Δ* backgrounds in no-Pi conditions are expressed as the \log_2 ratio of RPKM values. Statistical significance for expression differences between either high-Pi and no-Pi conditions in a *pho7⁺* background, or the expression difference between a *pho7⁺* and *pho7 Δ* background grown in no-Pi media were calculated with slight modifications to the previously described method (Bitton et al., 2011; Bloom et al., 2009). Briefly, for each annotated feature a 20kb sliding window was used to generate direction specific read totals for the surrounding region. We then applied a two-tailed Fischer's exact test with a 2x2 contingency table. The rows were separated by conditions assayed and the columns were separated by sum read totals within the feature and read totals in the total 20kb sliding frame. The statistical test was performed using MATLAB with code from Michael Boedigheimer (Boedigheimer, 2008). Total transcript abundance was normalized for each nucleotide as nucleotide reads per million sequenced reads (RPM).

ACKNOWLEDGEMENTS

The authors wish to thank C. Daly and J. Zhang for help with Illumina sequencing, V. Vijayan and X. Zhou for assistance with MATLAB and statistical analysis, B. Zid for assistance with RNA-Seq, and all members of the O'Shea lab for thoughtful discussion and commentary. This work was supported by the Howard Hughes Medical Institute.

REFERENCES

- Abel, S. (2011). Phosphate sensing in root development. *Curr Opin Plant Biol* 14, 303-309.
- Aiba, H., Yamada, H., Ohmiya, R., and Mizuno, T. (1995). The osmo-inducible *gpd1+* gene is a target of the signaling pathway involving Wis1 MAP-kinase kinase in fission yeast. *FEBS Lett* 376, 199-201.
- Askwith, C., and Kaplan, J. (1997). An oxidase-permease-based iron transport system in *Schizosaccharomyces pombe* and its expression in *Saccharomyces cerevisiae*. *J Biol Chem* 272, 401-405.
- Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* 27, 1653-1659.
- Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37, W202-208.
- Bailis, J.M., and Forsburg, S.L. (2002). RNAi hushes heterochromatin. *Genome Biol* 3, REVIEWS1035.
- Bellemare, D.R., Sanschagrín, M., Beaudoin, J., and Labbe, S. (2001). A novel copper-regulated promoter system for expression of heterologous proteins in *Schizosaccharomyces pombe*. *Gene* 273, 191-198.
- Bergwitz, C., and Juppner, H. (2011). Phosphate sensing. *Adv Chronic Kidney Dis* 18, 132-144.
- Bitton, D.A., Grallert, A., Scutt, P.J., Yates, T., Li, Y., Bradford, J.R., Hey, Y., Pepper, S.D., Hagan, I.M., and Miller, C.J. (2011). Programmed fluctuations in sense/antisense transcript ratios drive sexual differentiation in *S. pombe*. *Mol Syst Biol* 7, 559.
- Bloom, J.S., Khan, Z., Kruglyak, L., Singh, M., and Caudy, A.A. (2009). Measuring differential gene expression by short read sequencing: quantitative comparison to 2-channel gene expression microarrays. *BMC Genomics* 10, 221.
- Boedigheimer, M. (2008). Fisher's Exact Test (<http://www.mathworks.com/matlabcentral/fileexchange/22550-fishers-exact-test>), MATLAB Central File Exchange. Retrieved February 21, 2012.
- Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., and Lewis, S. (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* 25, 288-289.
- Chen, D., Toone, W.M., Mata, J., Lyne, R., Burns, G., Kivinen, K., Brazma, A., Jones, N., and Bahler, J. (2003). Global transcriptional responses of fission yeast to environmental stress. *Mol Biol Cell* 14, 214-229.

Cohen, A., Perzov, N., Nelson, H., and Nelson, N. (1999). A novel family of yeast chaperons involved in the distribution of V-ATPase and other membrane proteins. *J Biol Chem* 274, 26885-26893.

Doerner, P. (2008). Phosphate starvation signaling: a threesome controls systemic P(i) homeostasis. *Curr Opin Plant Biol* 11, 536-540.

Dutrow, N., Nix, D.A., Holt, D., Milash, B., Dalley, B., Westbroek, E., Parnell, T.J., and Cairns, B.R. (2008). Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA-DNA hybrid mapping. *Nat Genet* 40, 977-986.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95, 14863-14868.

Ekwall, K. (2004). The roles of histone modifications and small RNA in centromere function. *Chromosome Res* 12, 535-542.

Fisher, E., Almaguer, C., Holic, R., Griac, P., and Patton-Vogt, J. (2005). Glycerophosphocholine-dependent growth requires Gde1p (YPL110c) and Git1p in *Saccharomyces cerevisiae*. *J Biol Chem* 280, 36110-36117.

Forsburg, S.L., and Rhind, N. (2006). Basic methods for fission yeast. *Yeast* 23, 173-183.

Garriga, J., and Grana, X. (2004). Cellular control of gene expression by T-type cyclin/CDK9 complexes. *Gene* 337, 15-23.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11, 4241-4257.

Gorner, W., Durchschlag, E., Martinez-Pastor, M.T., Estruch, F., Ammerer, G., Hamilton, B., Ruis, H., and Schuller, C. (1998). Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes Dev* 12, 586-597.

Gould, K.L., Ren, L., Feoktistova, A.S., Jennings, J.L., and Link, A.J. (2004). Tandem affinity purification and identification of protein complex components. *Methods* 33, 239-244.

Hao, N., and O'Shea, E.K. (2011). Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nat Struct Mol Biol* 19, 31-39.

Henry, T.C., Power, J.E., Kerwin, C.L., Mohammed, A., Weissman, J.S., Cameron, D.M., and Wykoff, D.D. (2011). Systematic screen of *Schizosaccharomyces pombe* deletion collection uncovers parallel evolution of the phosphate signal transduction pathway in yeasts. *Eukaryot Cell* 10, 198-206.

Hermend, D., Pihlak, A., Westerling, T., Damagnez, V., Vandenhoute, J., Cottarel, G., and Makela, T.P. (1998). Fission yeast Csk1 is a CAK-activating kinase (CAKAK). *EMBO J* 17, 7230-7238.

Hirota, K., Mizuno, K., Shibata, T., and Ohta, K. (2008). Distinct chromatin modulators regulate the formation of accessible and repressive chromatin at the fission yeast recombination hotspot *ade6-M26*. *Mol Biol Cell* 19, 1162-1173.

Huang, D., Moffat, J., and Andrews, B. (2002). Dissection of a complex phenotype by functional genomics reveals roles for the yeast cyclin-dependent protein kinase Pho85 in stress adaptation and cell integrity. *Mol Cell Biol* 22, 5076-5088.

Hulett, F.M. (1996). The signal-transduction network for Pho regulation in *Bacillus subtilis*. *Mol Microbiol* 19, 933-939.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218-223.

Johnston, M. (1987). A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. *Microbiol Rev* 51, 458-476.

Kaffman, A., Herskowitz, I., Tjian, R., and O'Shea, E.K. (1994). Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science* 263, 1153-1156.

Kaffman, A., Rank, N.M., O'Neill, E.M., Huang, L.S., and O'Shea, E.K. (1998a). The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature* 396, 482-486.

Kaffman, A., Rank, N.M., and O'Shea, E.K. (1998b). Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes Dev* 12, 2673-2683.

Kerwin, C.L., and Wykoff, D.D. (2009). *Candida glabrata* PHO4 is necessary and sufficient for Pho2-independent transcription of phosphate starvation genes. *Genetics* 182, 471-479.

Kim, D.U., Hayles, J., Kim, D., Wood, V., Park, H.O., Won, M., Yoo, H.S., Duhig, T., Nam, M., Palmer, G., *et al.* (2010). Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* 28, 617-623.

Komeili, A., and O'Shea, E.K. (1999). Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science* 284, 977-980.

Labbe, S., Pena, M.M., Fernandes, A.R., and Thiele, D.J. (1999). A copper-sensing transcription factor regulates iron uptake genes in *Schizosaccharomyces pombe*. *J Biol Chem* 274, 36252-36260.

- Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. *Nature* **453**, 246-250.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25.
- Lee, Y.S., Huang, K., Quijcho, F.A., and O'Shea, E.K. (2008). Molecular basis of cyclin-CDK-CKI regulation by reversible binding of an inositol pyrophosphate. *Nat Chem Biol* **4**, 25-32.
- Lee, Y.S., Mulugu, S., York, J.D., and O'Shea, E.K. (2007). Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science* **316**, 109-112.
- Lenburg, M.E., and O'Shea, E.K. (1996). Signaling phosphate starvation. *Trends Biochem Sci* **21**, 383-387.
- Liang, S.D., Marmorstein, R., Harrison, S.C., and Ptashne, M. (1996). DNA sequence preferences of GAL4 and PPR1: how a subset of Zn₂ Cys₆ binuclear cluster proteins recognizes DNA. *Mol Cell Biol* **16**, 3773-3780.
- MacPherson, S., Larochele, M., and Turcotte, B. (2006). A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiol Mol Biol Rev* **70**, 583-604.
- Maerkl, S.J., and Quake, S.R. (2007). A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233-237.
- Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356**, 408-414.
- Martinez-Pastor, M.T., Marchler, G., Schuller, C., Marchler-Bauer, A., Ruis, H., and Estruch, F. (1996). The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *EMBO J* **15**, 2227-2235.
- Maudrell, K., Nurse, P., Schonholzer, F., and Schweingruber, M.E. (1985). Cloning and characterization of two genes restoring acid phosphatase activity in *pho1-* mutants of *Schizosaccharomyces pombe*. *Gene* **39**, 223-230.
- Monahan, B.J., Villen, J., Marguerat, S., Bahler, J., Gygi, S.P., and Winston, F. (2008). Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nat Struct Mol Biol* **15**, 873-880.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628.
- Mouillon, J.M., and Persson, B.L. (2006). New aspects on phosphate sensing and signalling in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **6**, 171-176.

- Ni, T., Tu, K., Wang, Z., Song, S., Wu, H., Xie, B., Scott, K.C., Grewal, S.I., Gao, Y., and Zhu, J. (2010). The prevalence and regulation of antisense transcripts in *Schizosaccharomyces pombe*. *PLoS One* 5, e15271.
- O'Neill, E.M., Kaffman, A., Jolly, E.R., and O'Shea, E.K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* 271, 209-212.
- Ogawa, N., DeRisi, J., and Brown, P.O. (2000). New components of a system for phosphate accumulation and polyphosphate metabolism in *Saccharomyces cerevisiae* revealed by genomic expression analysis. *Mol Biol Cell* 11, 4309-4321.
- Ohmiya, R., Yamada, H., Nakashima, K., Aiba, H., and Mizuno, T. (1995). Osmoregulation of fission yeast: cloning of two distinct genes encoding glycerol-3-phosphate dehydrogenase, one of which is responsible for osmotolerance for growth. *Mol Microbiol* 18, 963-973.
- Okada, H., and Toh-e, A. (1992). A novel mutation occurring in the PHO80 gene suppresses the PHO4c mutations of *Saccharomyces cerevisiae*. *Curr Genet* 21, 95-99.
- Oshima, Y. (1997). The phosphatase system in *Saccharomyces cerevisiae*. *Genes Genet Syst* 72, 323-334.
- Pei, Y., Du, H., Singer, J., Stamour, C., Granitto, S., Shuman, S., and Fisher, R.P. (2006). Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIH-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 in vivo. *Mol Cell Biol* 26, 777-788.
- Pelletier, B., Beaudoin, J., Mukai, Y., and Labbe, S. (2002). Fep1, an iron sensor regulating iron transporter gene expression in *Schizosaccharomyces pombe*. *J Biol Chem* 277, 22950-22958.
- Petit, T., Blazquez, M.A., and Gancedo, C. (1996). *Schizosaccharomyces pombe* possesses an unusual and a conventional hexokinase: biochemical and molecular characterization of both hexokinases. *FEBS Lett* 378, 185-189.
- Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., *et al.* (2012). The Pfam protein families database. *Nucleic Acids Res* 40, D290-D301.
- Reddy, V.S., Singh, A.K., and Rajasekharan, R. (2008). The *Saccharomyces cerevisiae* PHM8 gene encodes a soluble magnesium-dependent lysophosphatidic acid phosphatase. *J Biol Chem* 283, 8846-8854.
- Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* 147, 1408-1419.

- Schmidt, G., Bartsch, G., Laumont, M.C., Herman, T., and Liss, M. (1963). Acid phosphatase of bakers' yeast: an enzyme of the external cell surface. *Biochemistry* 2, 126-131.
- Schweingruber, M.E., Edenharter, E., Zurlinden, A., and Stockmaier, K.M. (1992). Regulation of *pho1*-encoded acid phosphatase of *Schizosaccharomyces pombe* by adenine and phosphate. *Curr Genet* 22, 289-292.
- Sonnhammer, E.L., Eddy, S.R., and Durbin, R. (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 28, 405-420.
- Springer, M., Wykoff, D.D., Miller, N., and O'Shea, E.K. (2003). Partially phosphorylated Pho4 activates transcription of a subset of phosphate-responsive genes. *PLoS Biol* 1, E28.
- Tanaka, K., and Okayama, H. (2000). A *pcl*-like cyclin activates the Res2p-Cdc10p cell cycle "start" transcriptional factor complex in fission yeast. *Mol Biol Cell* 11, 2845-2862.
- Thomas, M.R., and O'Shea, E.K. (2005). An intracellular phosphate buffer filters transient fluctuations in extracellular phosphate levels. *Proc Natl Acad Sci U S A* 102, 9565-9570.
- Tsakraklides, V., and Solomon, M.J. (2002). Comparison of Cak1p-like cyclin-dependent kinase-activating kinases. *J Biol Chem* 277, 33482-33489.
- Turunen, O., Seelke, R., and Macosko, J. (2009). In silico evidence for functional specialization after genome duplication in yeast. *FEMS Yeast Res* 9, 16-31.
- Verdel, A., and Moazed, D. (2005). RNAi-directed assembly of heterochromatin in fission yeast. *FEBS Lett* 579, 5872-5878.
- Vijayan, V., Jain, I.H., and O'Shea, E.K. (2011). A high resolution map of a cyanobacterial transcriptome. *Genome Biol* 12, R47.
- Vijayan, V., Zuzow, R., and O'Shea, E.K. (2009). Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc Natl Acad Sci U S A* 106, 22564-22568.
- Vogel, K., Horz, W., and Hinnen, A. (1989). The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Mol Cell Biol* 9, 2050-2057.
- Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453, 1239-1243.
- Wilson-Grady, J.T., Villen, J., and Gygi, S.P. (2008). Phosphoproteome analysis of fission yeast. *J Proteome Res* 7, 1088-1097.

Wood, V., Harris, M.A., McDowall, M.D., Rutherford, K., Vaughan, B.W., Staines, D.M., Aslett, M., Lock, A., Bahler, J., Kersey, P.J., *et al.* (2012). PomBase: a comprehensive online resource for fission yeast. *Nucleic Acids Res* 40, D695-699.

Wykoff, D.D., and O'Shea, E.K. (2001). Phosphate transport and sensing in *Saccharomyces cerevisiae*. *Genetics* 159, 1491-1499.

Yoshida, K., Ogawa, N., and Oshima, Y. (1989). Function of the PHO regulatory genes for repressible acid phosphatase synthesis in *Saccharomyces cerevisiae*. *Mol Gen Genet* 217, 40-46.

Zhou, X., and O'Shea, E.K. (2011). Integrated approaches reveal determinants of genome-wide binding and function of the transcription factor Pho4. *Mol Cell* 42, 826-836.

CHAPTER 4

Conclusions Regarding the *S.cerevisiae* Pho85-Pho80 Structure and the PHO

Pathway in *S.pombe*

Structure and Function of the Pho85-Pho80 Complex

After solving the crystal structure of Pho85-Pho80 we successfully identified a number of unique features that provide insights into the function of the Pho85-Pho80 complex. We elucidated the structural reason for T-loop dispensability, demonstrating that residues R132 on Pho85 and D136 on Pho80 form a salt bridge activating the CDK-cyclin complex. We found that the phenylalanine at position 138 in Pho80 was responsible for optimal phosphorylation of Pho4, and plays a partial role in recognizing the phosphorylation sequence motif SPXI/L. Utilizing rational mutagenesis we identified a distal site on Pho80 that contributes to the binding and processive phosphorylation of Pho4 by the Pho85-Pho80 complex. Finally, based on previous mutagenesis studies we identified the surface on Pho80 that interacts with the CKI Pho81 and isolated a putative binding pocket for IP₇.

There are two primary directions that would be interesting for continuing research on the structure and function of the Pho85-Pcl complexes. First, we would like to solve the structure of the novel inhibitory complex Pho85-Pho80-Pho81MD-IP₇-Mg²⁺. Besides informing our understanding of the switch-like behavior of Pho81 – which would be the first structural elucidation of a non-canonical CKI mechanism – such a structure could provide clues regarding protein-polyphosphate inositol interactions. Recently, inositol polyphosphates have been implicated in a number of cellular processes – chromatin remodeling, chemotaxis, signal transduction, among others (Irvine et al., 1999; Luo et al., 2003; Schell et al., 1999; Steger et al., 2003) – and it would be interesting to know if there were key structural features that are conserved between inositol polyphosphate interacting proteins.

We have made some progress towards this aim using multi-dimensional NMR. We have successfully created vectors for recombinantly expressing the full Pho81MD – as well as the peptides S1, S3, and S3S1 – with cleavable GST and His₆ tags in *E.coli*. We were able to produce each construct in minimal media supplemented with ¹⁵NH₄Cl as the sole nitrogen source, as well as ¹⁵N-labeled Pho85-Pho80. Using recombinant sources of Pho85-Pho80 and IP₇ we formed the Pho85-Pho80-¹⁵N-S3S1-IP₇-Mg²⁺ complex *in vitro* and subjected the labeled peptides to HSQC analysis (data not shown). We also collected TROSY spectra for the ¹⁵N-labeled Pho85-Pho80 either alone or in complex with S3S1-IP₇-Mg²⁺ (data not shown).

Unfortunately, we encountered serious obstacles when we attempted reconstitution of the inhibitory complex at concentrations above 100 μM. Based on previous attempts at increasing the solubility of inositol polyphosphate species (Grynspan and Cheryan, 1983) and truculent proteins (Golovanov et al., 2004) we used a number of different buffers containing a range of NaCl (30-150 mM), pH (6-8), MgCl₂ (0.3-10 mM), and free arginine and glutamic acid (up to 50 mM). In each case precipitation followed complex formation, with almost complete precipitation of peptide following 12 hours (data not shown). By lowering the initial concentration of ¹⁵N-S3S1 to 50 μM we were able to perform a crude titration with unlabelled Pho85-Pho80 (0.2-1.8 molar equivalents), but the addition of IP₇ once again caused precipitation of the complex. Attempts to verify the published K_D for the S1 or S3S1 interaction with Pho85-Pho80 at concentrations above 100 μM via isothermal titration calorimetry were also met with precipitation issues (data not shown). Moreover, while we could increase the bulk preparation of IP₇ using recombinant Vip1, there were complications scaling the

preparation beyond 30 mL and a final yield of ~10 uM. These significant challenges have led us to the belief that further structural analysis should be pursued using x-ray crystallography.

One possible strategy involves crystal formation of two distinct Pho85-Pho80-Pho81-IP₇ complexes: (1) Pho85-Pho80-S3 and (2) Pho85-Pho80-S1-IP₇-Mg²⁺. The first should provide insight regarding the constitutive interaction between Pho81 and Pho85-Pho80 independent from external Pi concentrations; the second would provide the inhibited complex. Early attempts at forming crystals of Pho85-Pho80 and soaking in Pho81MD failed because the requirement for IP₇ was not known. Further, it is possible that the Pho85-Pho80-Pho81MD structure undergoes conformational changes upon binding IP₇, which would cause problems within the crystal lattice for the full length Pho81MD during switching. By using S1 and S3 separately, this problem should be avoided. Once obtained, the crystal structure could be solved using molecular replacement with our Pho85-Pho80 structure.

The second direction for the functional analysis of Pho85-Pcl complexes involves the biochemical nature of Pcl substrate specificity. As mentioned in chapter 2, each of the different Pcl family members targets Pho85 to specific phosphorylation motifs (Dephoure et al., 2005; Huang et al., 2007). Previous reports have demonstrated that peptide libraries can be used to build energy landscapes for CDK-cyclin specificity (Songyang et al., 1994). Using peptides with the Pho4 SP6 sequence modified with all 20 amino acids at the +3 position (i.e. SPXN) we could build landscapes for each of the Pho85-Pcl complexes. With this information we could then determine the contribution from the F138 residue of Pho80 on substrate selection. We also know from alignments

of the Pcl family members the residues that are equivalent to F138 in each Pcl (see Figure S2.1). Such residues might be important for each individual Pcl's substrate specificity and mutating each would allow us to build a relative contribution index for each of the Pcl residues. We could then determine whether a single swap at this position is sufficient for altering substrate specificity or if we will need to build larger Pcl chimeras to re-wire Pho85-Pcl targeting. This experimental set-up would elucidate the structural features responsible for diverse substrate selection by the Pcl family.

The *S.pombe* PHO Pathway and Beyond: The *pho7*⁺ Regulatory Mechanism

Using a systems biology approach we characterized the PHO transcriptional response in the fission yeast *S.pombe*. The PHO response has a fast element – genes responsible for inorganic phosphate harvesting and transport – and a slow element – genes responsible for intracellular phosphate management and alternate organic phosphate harvesting. The core PHO regulon, comprising the *PHO5*, *PHO84*, and *GIT1* orthologs *pho1*⁺, *spbc8e4.01c*⁺, and *spbc1271.09*⁺, is activated by the transcription factor *pho7*⁺ in response to phosphate starvation. *csk1*⁺ represses the activation of these three genes by *pho7*⁺ in high-Pi conditions. Fusing varying lengths of the *pho1*⁺ promoter to a *yfp*⁺ expression construct we identified a segment between -2 kb and -1 kb that is necessary for *csk1*⁺ mediated inhibition; which we have termed the URS. We also used this system to verify the *pho7*⁺ UAS between nucleotides -245 and -225 in the *pho1*⁺ promoter.

Based on our CHIP-Seq data we found evidence of widespread Pho7 binding in regions that are not involved in the phosphate starvation response. GO Term

enrichment suggested that these Pho7-bound genes might play a role in responding to other forms of environmental stress. Therefore, we grew *pho7⁺* and *pho7Δ* cells in iron starvation, copper starvation, high osmotic pressure, and glucose-free conditions and assayed gene expression using microarray hybridization. We found that differential sets of the Pho7-bound genes responded to the varying stressors and, for ~15-18% of those genes, maximal expression depended on *pho7⁺*. Our RNA-Seq analysis implicated *pho7⁺*-dependent regulation in activating non-coding and anti-sense transcription. Moreover, we discovered evidence of a bi-functional role for *pho7⁺*, activating certain regions of the genome for transcription and repressing others. In sum, our work proves that *pho7⁺* plays a role in global transcript regulation in response to multiple environmental stressors.

There are a number of experimental directions that would be fruitful for continuing work on phosphate, *pho7⁺*, and *csk1⁺* dependent transcriptional regulation. First, while we have defined a 20 bp region in the *pho1⁺* promoter necessary for *pho7⁺* activity (the UAS), we still have not demonstrated a direct interaction between Pho7 and DNA. Using recombinant Pho7 we could verify this interaction *in vitro* using an electrophoretic mobility shift assay (EMSA) with radiolabeled DNA probe. This method has been validated with the related Gal4 protein (also containing a ZC domain), and was used to distinguish motif preferences for multiple ZC containing transcription factors (Liang et al., 1996). Thus, if Pho7 binds directly to DNA, this experiment would allow us to define the DNA binding motif for Pho7, which has been recalcitrant to *in silico* discovery.

Second, we would like to understand the biochemical mechanism behind *pho7*⁺ activation and repression. There are two possible scenarios that could be controlling Pho7 activity. In the first, a post-translational modification of Pho7 could be required to fully activate or repress Pho7 functionality. We know that the URS in the *pho1*⁺ promoter is necessary for *csk1*⁺-mediated repression. Our initial attempts at identifying a Csk1 phosphorylation target on Pho7 resulted in putative Csk1 target site mutants that were not defective for *pho1*⁺ transcriptional regulation; yet, this does not preclude Pho7 phosphorylation by Csk1 at the promoter as a requirement for repression. Recombinant Csk1 has been successfully purified from both *E.coli* and insect cell sources and is fully capable of phosphorylating Mcs6 and Cdk9 (Hermand et al., 1998; Pei et al., 2006). Using recombinant Pho7 we could determine whether Csk1 phosphorylates Pho7 *in vitro*, and compare the enzyme kinetics to known Csk1 targets. If the URS requirement extends to the *in vitro* system we could try recombinant Csk1 kinase assays with lysate from *csk1*Δ cells. Starting with synthetic stretches of the *pho1*⁺ promoter (i.e. plasmid vectors), we would add whole cell extract to reconstitute the chromatin architecture, and assay epitope-tagged Pho7 phosphorylation in the presence or absence of Csk1.

It is possible that either Csk1 is not the enzyme responsible for phosphorylating Pho7 or that the post-translational modification controlling Pho7 activity is not a phosphorylation event. Purification of epitope-tagged Pho7 from *csk1*⁺ and *csk1*Δ cells grown in high-Pi or no-Pi conditions followed by post-translational mass spectrometry (PT-MS) would allow us to determine: (1) whether Pho7 is post-translationally modified in response to phosphate starvation, (2) what modifications are occurring, and (3) the

extent of Csk1 involvement in Pi-dependent Pho7 modification. This would also provide validation for the *in vitro* kinase assays described above.

In the second - not mutually exclusive - scenario, Pho7 activity is determined by the proteins in complex with Pho7 in either high-Pi or no-Pi conditions. In this scenario Pho7 is directed to the correct locations for transcriptional activity based on the DNA binding affinities of the other, non-Pho7 complex members. We know from the ChIP-Seq results that Pho7 is recruited to Pi-responsive promoters during phosphate starvation. It is possible that the difference in DNA binding affinity is attributable to a non-Pho7 component of the activating complex that is itself regulated by phosphate starvation. Besides the differential recruitment of Pho7 to DNA, it is known that the opening and closing of transcribed regions in *S.pombe* is a complex process relying on the coordinated efforts of many different DNA and histone modification enzymes (Grewal and Jia, 2007; Suganuma and Workman, 2008). Perhaps the ability of Pho7 to switch from an inactive form to an active one at PHO promoters is a function of the successful recruitment of one such factor to the Pho7-containing complex. To address this mechanism we could once again utilize an epitope tagged version of Pho7. Growing the tagged-Pho7 strain in either high-Pi or no-Pi conditions we could prepare whole cell extracts that would be subjected to affinity-purification followed by mass spectrometry (complex-MS). This would allow us to identify proteins bound to Pho7 and assay the composition of protein complexes during phosphate starvation.

We have made some progress towards addressing these experimental concerns, but we are stymied by the difficulties in producing recombinant Pho7 in *E.coli* or pure, full-length Pho7-TAP in *S.pombe*. We have vectors for expressing MBP and GST

fusions of Pho7 in *E.coli*, but the recombinant protein is not expressed correctly and is predominately found precipitated in the pelleted debris. Other groups working with ZC domains have observed that working with shorter constructs (i.e. only the first 60-100 residues of the Gal4 protein) or expressing/purifying the protein with excess zinc greatly increases solubility and yield (Lenouvel et al., 1997; Liang et al., 1996; Marmorstein et al., 1992). We could modify our method and create deletion constructs of Pho7 containing variable sequences surrounding the ZC domain. These constructs could then be swapped into *S.pombe* expression vectors and Pho7 complementation could be confirmed in the *pho7* Δ background. Hopefully this would provide a ready source of recombinant Pho7 for *in vitro* studies.

As we demonstrated in Figure S3.4, full-length Pho7-TAP purified from *S.pombe* extract is not pure and has a low yield. To overcome these limitations we could try tagging Pho7 with a suite of different tags (FLAG, Myc, HA, etc.) at both the termini, and determine the most efficient tag system for purification. If our work with the ZC deletion constructs produces a truncated form of Pho7 that is necessary and sufficient for PHO pathway regulation, we could use this deleted form instead of the full-length Pho7. Finally, there are plasmid based systems for producing protein in *S.pombe* from promoters of varying strengths that we could use to drive Pho7 expression (Siam et al., 2004). While this would cause disruption of the normal stoichiometries for Pho7-containing protein complexes, we might be able to use over-expressed Pho7 as bait for complex-MS or PT-MS.

Third, we want to detail the mechanism for combinatorial control of Pho7 in response to diverse environmental stressors. Using the methods outlined above, we

could determine whether or not Pho7 undergoes different post-translational modifications and/or interacts with different protein complexes based on the nature of the perceived stress. We already know the transcription factors responsible for activation or repression of the key stress response genes in a few of the signal transduction pathways identified in our microarray analysis (Labbe et al., 1999; Mercier et al., 2008; Ohmiya et al., 1997; Pelletier et al., 2002). It is possible that these previously identified factors are responsible for regulating Pho7 activity. For example, during iron depletion Fep1, a transcriptional repressor, binds upstream of both *fip1*⁺ and *fio1*⁺ preventing induction of the iron uptake system (Pelletier et al., 2002). In our ChIP-Seq data we discovered a binding site for Pho7 directly over the Fep1 site (Figure S3.5A). It would be useful to know whether Fep1 represses Pho7 and if so, whether that interaction is direct or indirect. Further, we would like to know if Pho7 repression is mediated through Csk1 in all of these additional stress response pathways or if there are negative regulators specific to each stress. Using the genome wide *S.pombe* deletion collection and phenotypic markers for the different stressors (akin to secreted acid phosphatase plate assays) we could attempt an identification of the factors responsible for gene regulation and their epigenetic relationships to Pho7.

Finally, we would like to explore the functional role of the non-coding and anti-sense transcription regulated by Pho7. Before attempting an in-depth exploration of ncRNA function we would first confirm that Pho7 binding is necessary for ncRNA transcription. Swapping the selected ncRNA promoters into our *pho1*⁺*pr-yfp*⁺ system will allow us to delete the appropriate Pho7 binding site and measure its impact on transcriptional output. This will help us avoid ncRNAs that are up- or down-regulated in

response to indirect effects of Pho7 deletion. Previous genome wide studies have remarked on the widespread prevalence of non-coding transcription in *S.pombe*, though the functionality of many transcripts remains unknown (Ni et al., 2010; Rhind et al., 2011). Bitton et al., recently outlined the role for ncRNA and anti-sense RNAs in regulating gene silencing during mating type switching (Bitton et al., 2011). They developed a method for determining the effects of anti-sense transcription in *trans* using ectopic locus transplantation, a method that could be applied to our cohort of ncRNAs. Further, they demonstrated that regulation by their ncRNAs required the RNAi machinery, which is something we could assay with our ncRNAs.

Regardless of the directions pursued in the wake of this research, we believe that our characterization of the PHO pathway in *S.pombe* has revealed a multifaceted and novel stress response network. At the center of this network lies the transcription factor *pho7⁺*, which plays a major role in the phosphate starvation response, and a minor role in multiple other stress response pathways. The biological and biochemical details surrounding *pho7⁺* regulation should provide ample material for future study, and should hopefully provide insights into how diverse environmental conditions have influenced the evolution of phosphate responsive signal transduction pathways in diverse fungal species.

REFERENCES

- Bitton, D.A., Grallert, A., Scutt, P.J., Yates, T., Li, Y., Bradford, J.R., Hey, Y., Pepper, S.D., Hagan, I.M., and Miller, C.J. (2011). Programmed fluctuations in sense/antisense transcript ratios drive sexual differentiation in *S. pombe*. *Mol Syst Biol* 7, 559.
- Dephoure, N., Howson, R.W., Blethrow, J.D., Shokat, K.M., and O'Shea, E.K. (2005). Combining chemical genetics and proteomics to identify protein kinase substrates. *Proc Natl Acad Sci U S A* 102, 17940-17945.
- Golovanov, A.P., Hautbergue, G.M., Wilson, S.A., and Lian, L.Y. (2004). A simple method for improving protein solubility and long-term stability. *J Am Chem Soc* 126, 8933-8939.
- Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. *Nat Rev Genet* 8, 35-46.
- Grynspan, F., and Cheryan, M. (1983). Calcium Phytate - Effect of Ph and Molar Ratio on In Vitro Solubility. *J Am Oil Chem Soc* 60, 1761-1764.
- Hermand, D., Pihlak, A., Westerling, T., Damagnez, V., Vandenhoute, J., Cottarel, G., and Makela, T.P. (1998). Fission yeast Csk1 is a CAK-activating kinase (CAKAK). *EMBO J* 17, 7230-7238.
- Huang, K., Ferrin-O'Connell, I., Zhang, W., Leonard, G.A., O'Shea, E.K., and Quijcho, F.A. (2007). Structure of the Pho85-Pho80 CDK-cyclin complex of the phosphate-responsive signal transduction pathway. *Mol Cell* 28, 614-623.
- Irvine, R.F., McNulty, T.J., and Schell, M.J. (1999). Inositol 1,3,4,5-tetrakisphosphate as a second messenger--a special role in neurones? *Chem Phys Lipids* 98, 49-57.
- Labbe, S., Pena, M.M., Fernandes, A.R., and Thiele, D.J. (1999). A copper-sensing transcription factor regulates iron uptake genes in *Schizosaccharomyces pombe*. *J Biol Chem* 274, 36252-36260.
- Lenouvel, F., Nikolaev, I., and Felenbok, B. (1997). In vitro recognition of specific DNA targets by AlcR, a zinc binuclear cluster activator different from the other proteins of this class. *J Biol Chem* 272, 15521-15526.
- Liang, S.D., Marmorstein, R., Harrison, S.C., and Ptashne, M. (1996). DNA sequence preferences of GAL4 and PPR1: how a subset of Zn²⁺ Cys₆ binuclear cluster proteins recognizes DNA. *Mol Cell Biol* 16, 3773-3780.
- Luo, H.R., Huang, Y.E., Chen, J.C., Saiardi, A., Iijima, M., Ye, K., Huang, Y., Nagata, E., Devreotes, P., and Snyder, S.H. (2003). Inositol pyrophosphates mediate chemotaxis in *Dictyostelium* via pleckstrin homology domain-PtdIns(3,4,5)P₃ interactions. *Cell* 114, 559-572.

- Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* 356, 408-414.
- Mercier, A., Watt, S., Bahler, J., and Labbe, S. (2008). Key function for the CCAAT-binding factor Php4 to regulate gene expression in response to iron deficiency in fission yeast. *Eukaryot Cell* 7, 493-508.
- Ni, T., Tu, K., Wang, Z., Song, S., Wu, H., Xie, B., Scott, K.C., Grewal, S.I., Gao, Y., and Zhu, J. (2010). The prevalence and regulation of antisense transcripts in *Schizosaccharomyces pombe*. *PLoS One* 5, e15271.
- Ohmiya, R., Aiba, H., Yamada, H., and Mizuno, T. (1997). Clarification of the promoter structure of the osmoregulated *gpd1+* gene encoding an isozyme of NADH-dependent glycerol-3-phosphate dehydrogenase in fission yeast. *Biosci Biotechnol Biochem* 61, 553-555.
- Pei, Y., Du, H., Singer, J., Stamour, C., Granitto, S., Shuman, S., and Fisher, R.P. (2006). Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIH-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 in vivo. *Mol Cell Biol* 26, 777-788.
- Pelletier, B., Beaudoin, J., Mukai, Y., and Labbe, S. (2002). Fep1, an iron sensor regulating iron transporter gene expression in *Schizosaccharomyces pombe*. *J Biol Chem* 277, 22950-22958.
- Rhind, N., Chen, Z., Yassour, M., Thompson, D.A., Haas, B.J., Habib, N., Wapinski, I., Roy, S., Lin, M.F., Heiman, D.I., *et al.* (2011). Comparative functional genomics of the fission yeasts. *Science* 332, 930-936.
- Schell, M.J., Letcher, A.J., Brearley, C.A., Biber, J., Murer, H., and Irvine, R.F. (1999). PiUS (Pi uptake stimulator) is an inositol hexakisphosphate kinase. *FEBS Lett* 461, 169-172.
- Siam, R., Dolan, W.P., and Forsburg, S.L. (2004). Choosing and using *Schizosaccharomyces pombe* plasmids. *Methods* 33, 189-198.
- Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnica-Worms, H., and Cantley, L.C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol* 4, 973-982.
- Steger, D.J., Haswell, E.S., Miller, A.L., Wentz, S.R., and O'Shea, E.K. (2003). Regulation of chromatin remodeling by inositol polyphosphates. *Science* 299, 114-116.
- Suganuma, T., and Workman, J.L. (2008). Crosstalk among Histone Modifications. *Cell* 135, 604-607.

APPENDIX A

Chapter 2 Supplemental Data

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Bacterial Expression and Purification of Pho85-Pho80

Co-expression of Pho85-His₆ and Pho80 in *E. coli* essentially follows the published procedure (Jeffery et al., 2001). Briefly, the full length Pho85-His₆ and Pho80 were co-expressed in BL21 DE3 (Novagen) by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to final concentration of 200 μ M at 24°C after the cells are grown to A_{595} of 0.4 – 0.6. Following cell lysis by microfluidizer processor (Microfluidics), the His₆-Pho85-Pho80 was purified through Ni Sepharose High Performance and HiTrap Heparin High Performance columns (Amersham Biosciences). Similar procedure was used to express and purify the seleno-methionine variant of Pho85-Pho80 using *E. coli* B834 methionine auxotroph strain (Novagen). The purified native and Se-Met His₆-Pho85-Pho80 are essentially monodisperse as judged by dynamic light scattering measurement.

Crystallization and Diffraction Data Collection

Long needle-shaped crystals of the complex were grown in hanging drops at 10°C by mixing 3 volumes of the concentrated Pho85-Pho80 and 2 volumes of the reservoir solutions of 9 to 12% PEG 10K, 10% glycerol, 0.2 – 0.25 M strontium chloride (SrCl₂), 10 mM Tris (2-carboxyethyl) phosphine (TCEP) and 0.1 M 2-morpholinoethanesulfonic acid (MES), pH6.5. Crystals of the complex of Pho85-Pho80 in the presence of ATP- γ -S (Sigma) were obtained by microseeding. The crystals were cryo-protected by coating with paratone-N oil before flash-freezing in liquid nitrogen.

Both ligand-free and ATP- γ -S-bound crystals belong to the same space group and have very similar unit cell dimensions (Table 2.1). Based on the Matthews

coefficient analysis (Kantardjieff and Rupp, 2003; Matthews, 1968), the number of Pho85-Pho80 heterodimer molecules in the asymmetric unit could not be determined with certainty; assuming 2, 3 and 4 heterodimers in the asymmetric unit, the calculated solvent content are approximately 75, 62 and 50%, respectively.

Determination of the Crystal Structure of Pho85-Pho80

Although Pho85 shares 55% identity with CDK2, the use of molecular replacement technique to obtain initial phases, with the main chain of the crystal structure of CDK2 (PDB code: 1QMZ) as the search model, was not successful. Therefore, a single-wavelength anomalous dispersion (SAD) data set (Table 2.1) was collected at the microfocus beamline ID29, ESRF, Grenoble, France and processed and analyzed using MOSFLM (Leslie, 1992). The Pho85-Pho80 heterodimer has a total of twelve methionines (nine in Pho85 and three in Pho80), including the initiator M1 residues in both proteins. Excluding the two N-termini methionines, most (~80%) of the expected selenomethionine (or Se) sites resides in Pho85. Twelve Se sites were identified in SOLVE (Terwilliger and Berendzen, 1999) using the SAD data at the peak wavelength. These sites were refined in SHARP (La Fortelle and Bricogne, 1997) using the data between 30-3.5 Å. However, the overall experimental figure of merit was quite poor (~0.3), and the SAD-phased electron density map was not good enough to discern useful structural features and intermolecular boundaries.

We then used pattern matching in attempts to assess the correctness and relationship of the initial twelve sites and the existence of non-crystallographic symmetry (NCS). After measuring the distances between every two Se sites, we could identify two sets of three Se sites with nearly identical triangular geometries. Noting the locations of

the residues in the CDK2 crystal structure (PDB code: 1QMZ) that align with the eight internal methionines of the Pho85 sequence (two being identical to CDK2 sequence), we could also find a set of three methionine positions in the CDK2 structure that match with either of the two sets of three Se sites identified in the pattern matching. Taken together, the results indicate that the two sets of three Se sites are related by a NCS and that there are two heterodimers in the asymmetric unit. Indeed, the NCS matrix was found by FINDNCS (Lu, 1999). A SAD phased (based on the 12 seleno sites), two-fold NCS averaged and solvent flattened (75% solvent) electron density map at 3.5 Å showed considerable improvement and was very encouraging. Molecular boundaries could be made out. The C α trace of the crystal structure of CDK2, which has a significant sequence identity and similarity to Pho85, could be fitted reasonably well as an entire rigid body to the Pho85 density in the two heterodimers. More importantly, we could confirm not only the positions of the initial twelve Se sites but also found seven more sites from SHELXD (Schneider and Sheldrick, 2002) and SnB (Weeks and Miller, 1999). The nineteen total sites (ten in one heterodimer complex and nine in the other complex) were refined in SHARP and used to calculate another 2-fold averaged and solvent flattened map (figure of merit of 0.85). The densities of a significant number of side chains were revealed. Model building was done with the programs COOT (Emsley and Cowtan, 2004) and O (Jones et al., 1991). The model was refined against the native dataset, collected at the APS ID19 beamline, with CNS (Brunger et al., 1998) and CCP4 (CCP4, 1994) package.

Site-Directed Mutagenesis and Kinase Assays

Fusion PCR was used to generate Pho80 mutants, F138A with the following primers: 5'-AAAGGCTTATGTGATTCGGCCTCAACAAACGCCCATAT-3' and 5'-ATAATGGGCGTTTGTGAGGCCGAATCACATAAGCCTTT-3' and F138E with the following primers: 5'-AAAGGCTTATGTGATTCGCAATCAACAAACGCCCATAT-3' and 5'-ATAATGGGCGTTTGTGATTGCGAATCACATAAGCCTTT-3'. In both cases, the underlined sequence represents the mutated codon that was inserted into the PCR product. Subsequently the following primers were utilized for both mutants: 5'-TCACAACATATGGAAAGCACATCAGGA-3' and 5'-CATCAAGGATCCTTAATCTGGCTTTGATCG-3', which are complimentary to nucleotides +1 to +18 and +865 to +882 of the PHO80 gene, respectively. All PCR reactions were carried out with VentR polymerase according to the manufacturer's instructions (New England Biolabs). After cutting the PCR product into the pSBET-a vector (Schenk et al., 1995), the presence of the appropriate mutation was verified by sequencing.

Quick change mutagenesis was performed on EB1285 using the Quick Change II kit (Stratagene) to generate the Pho80 distal mutants with the following primer pairs: 5'-CAAAGTAATTCTGCCCGCTGATTTAATAAATAATTCTAGAACTGACCTAGT-3' and 5'-ACTAGGTCAGTTCTAGAAATATTTATTAATAATCAGCGGGCA-GAATTAATTTT-3' for C30Y, 5'-CTAGAACTGACCTAGTGGTGITCATA-TCACGAATGTTAGTA-3' and 5'-TACTAACATTCGTGATATGAACCACTAGG-TCAGTTCTAG-3' for L38F, 5'-CCTAGTGGTGCTCATATCACAATGTTAGTATCGCTGATAG-3' and 5'-CTATCAGCGATACTAACATTIGTGATATGAGCACCACTAGG-3' for R41Q, and 5'-

GAATAGTTCAGCTGGTGGATTCCTTTAACGCTTCACC-3' and 5'-

GGTGAAGCGTTAAAGGAAICCACCAGCTGAACTATTC-3' for G229D. In each case, the underlined sequence represents the nucleotide(s) that was modified. These EB1285 based vectors were then cut with BamHI and NdeI (New England Biolabs) and the Pho80 fragment was ligated into a pSBET-a vector (Schenk et al., 1995). Following ligation, all mutations were verified by sequencing. Coexpression of Pho85-His6 and the Pho80 mutants in *E. coli*, and their purifications, essentially follows the published procedure (Jeffery et al., 2001). The Pho80 distal site mutants were further purified using an isocratic elution from a HiLoad 16/60 75 µg gel filtration column (GE Healthcare), and the 1:1 stoichiometry of Pho85-His6:Pho80 was confirmed via Safe Blue staining (Invitrogen).

The kinase assays were carried out following published procedure (Jeffery et al., 2001). The Pho4 substrate was purified according to a published procedure (Jeffery et al., 2001). The peptides used in kinase assays (identified as SPVI and SPVA) were made by Biosynthesis, Inc. and Bethyl Laboratories, Inc. respectively. SPVI, which corresponds to the SP6 phosphorylation site of Pho4 (O'Neill et al., 1996), has the sequence 213-SAEGVVASE**SPVI**APHGSTHARSY-237, where the consensus sequence (**SPVI**) is in bold. SPVA, the non-specific substrate, is identical to the SPVI peptide substrate but with the I residue at the +3 position of the consensus sequence changed to A. For analysis of SPVI phosphorylation by the Pho80 distal site mutants, the dried p81 discs were exposed to a storage phosphor screen and imaged using the Typhoon Trio system (GE Healthcare).

Identification of a Shorter Segment of the 80 Residue Fragment of Pho81 that Binds Pho85-Pho80

We sought to determine whether there is a shorter segment of the 80-residue (644 to 723) segment/domain of Pho81, initially shown to be necessary for the inhibition of the kinase activity (Huang et al., 2001), that will bind Pho85-Pho80. Eight different truncation constructs (shown in Figure S2.4) of the 80-residue domain were cloned into the pTYB11 expression vector of the IMPACT T7 System (New England Biolabs) as an intein fused to a chitin binding domain (CBD). A pull down assay was used to determine the interaction between each truncation mutant and Pho85-Pho80. About 50 μ l of the chitin beads was mixed with 1 ml cell lysate of each truncation mutant-INTEIN-CBD. The beads were then centrifuged for 30 seconds at 1,000g, and the pellet suspended in 1 ml of 500 mM NaCl, 5% glycerol and 20 mM Tris, pH 8.0; this procedure was repeated five times to remove protein contaminants bound non-specifically to the chitin beads. The beads were then mixed with excess amount of purified Pho85-Pho80 and incubated for one hour. After subjecting the beads to extensive washing as described above, a sample was chromatographed on a SDS gel electrophoresis and the gel stained with Coomassie Blue. As shown in Figure S2.4, a shorter segment of the primary sequence of Pho81, which comprises of residues 665 to 701, can bind Pho85-Pho80. The fact that Pho85-Pho80 remained bound to the segment despite extensive washing of the beads indicates a tight complex. The 37-residue segment includes residues 690 to 695 and 699 to 701 which when mutated in Pho81 caused significant defects in inducing Pho5 expression (Huang et al., 2001). Interestingly, RXL motif is not present in the shorter segment that binds Pho85-Pho80. Although there is one putative RXL motif

at residue 647 to 649 (RPL) of the 80-residue domain, the three truncation mutants (644-652, 644-658 and 644-664) which harbor the RXL motif do not bind Pho85-Pho80 (Figure S2.4).

Cyclin Subunit	Substrate			Sequence of D-loops in Pcls ^b
	Name	Sequence ^a	Reference	
Pho80	Pho4	SPXI/L (5 sites)	(O'Neill et al., 1996)	DSFSF
"	Rim15 kinase	T ¹⁰⁷⁵ PPL	(Reinders et al., 1998)	DSFSF
Pcl1	Sic1	T ⁵ PPR	(Nishizawa et al., 1998)	DS <u>S</u> SPL
Pcl2	Rvs167	S ²⁰⁹ PVS, S ³⁷⁹ PPL	(Friesen et al., 2003)	DS <u>S</u> SPL
Pcl5	Gcn4	T ¹⁰⁵ PMF, T ¹⁶⁵ PVL	(Meimoun et al., 2000; Shemer et al., 2002)	DN <u>I</u> YS
Pcl6/7	Gcl8	T ¹¹⁸ PYQ	(Tan et al., 2003)	DF <u>F</u> YS
Pcl10	Gsy2	S ⁶⁵⁴ PRD	(Wilson et al., 1999)	DFV <u>H</u> S

^aSequence of the peptide segment containing the phosphorylatable S or T residue.


^bBased on amino sequence alignment of the Pho80/Pcls (Supplementary Figure S2.1). The first residue (D), which is involved in a functional salt link with Pho85 R132 (see Text), is invariant in all ten cyclins. The third residue (F138) of the Pho80 D-loop is potentially involved in recognizing I/L at position +3 of the phosphorylatable sites on Pho4 (see Text).

Figure S2.1. Amino Acid Sequence Alignment of Pho80 and the Other Nine Pho85 Cyclins (Pcls)

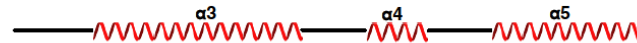
The alignment was done using Clustal W program and colored by the Boxshade Program. The α helices, loops and the two short strands of one β sheet as they occur in the Pho80 structure are placed above the alignment. The residues that are completely invariant, identical and similar are highlighted in green, yellow and cyan backgrounds, respectively. The numbers on the right hand side correspond to the amino acid positions in the primary sequences. The star identifies the position of an invariant aspartic acid residue (D136 in Pho80). Dashed lines indicate disordered segments in the structure, consisting of an eight-residue segment in the middle of the α NT- α 1 loop and a five-residue segment in the middle of the α CT1- α CT2 loop (see also Figure 2.1A in the main paper). The Pcl6 sequence (residues 230 to 390) could be aligned only after excluding the 70-residue segment from 268 to 327. The excluded segment, which is predicted to be between α 2 and α 3, is considerably longer than the corresponding segments in Pho80 and other Pcls.

----------

Pho80	--MESTSGERSENIHEDQGIPKVLIPADFNKCSRDTLVLVLSRMLVSLIAINENSATKK	57
pcl1	-----	
pcl2	-----VSKEMVQYLASTTAS	29
pcl5	-----	
pcl6	-----	
pcl7	EAKTHSLDEETNEQTDVKILNIADFPPTDELILMISALLNRIITANDETTDVSQQVSDETE	140
pcl8	-----ESIPKIPELSDDEALSKFRENIELILQLSKKINDNANTLAISSSE	348
pcl9	-----	
pcl10	-----PVDVEVPHISVDEALANFKETIELELLKLSG--NRKCTGFNTRVE	293
Clg1	-----	


----------

Pho80	SDDQITLTRYHSKIPENISIFNYFIRLTKFSSLEHCWLMTSLYIDLLQT-----	107
pcl1	-----SLMTFITRLVRYNVVYPTLLTAACYLNKLR-----	85
pcl2	IIKKKTNMIDIALFAPLTKFINRLIKHSNVQPTLMATSVYLAKLRS-----	79
pcl5	-----ILKFLNEVLKRSKCSKENAVLATYFQKIHQSR-----GV	114
pcl6	-----QIGLDQYFQRICKYCPPTNDVFLSLLVYFDRISKR-----	264
pcl7	DELLTPILAFYGNVPEIAVQYLERIQKCYPTNDIFLSLLVYFDRISKNYGHSRERNG	200
pcl8	DP--QKFVNFVMKNPFLSFRDFIDRIQNKCMFGAVWYLGATVYLLQLVFLTR---DEM DG	403
pcl9	-----PPDLSIFIKNVVIQSNVQPTLMATSVYLNKLS-----	79
pcl10	K--KEYSNFYMKSKFTLSSADFLKRIQDKCEYQPTWYLVAFVLDLFLTR---DGNN-	346
Clg1	-----YLNKLSNGIHSIGGHSIN-ITVQM-----	214

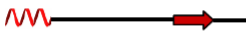
----------

Pho80	-VYPDFTLNSLTAHRFLLTATTVATKGLCSFSTNAHYAKVGG--VRCHELNILNDFLK	164
pcl1	ILPRDATGLPSTIHRIFLACLILSAKPHNDSPLNKHWAAYTDGLFTLEDINLMERQLLQ	145
pcl2	IIPSNVYGIETTRHRIFLGCCLILAAKTLNDSPLNKHWAAYTDGLFLRELVNTISRELLG	139
pcl5	RDESSLPEFESHCSRRIPLCCLILSHKFLNNTYSMKNWQIISG--LHAKDLSLMEWCLG	172
pcl6	CNSQMFVMDSHNIHRLIAGITVSTKFLSDFFYSNSRYAKVGG--ISLQELNHLQLQFLV	382
pcl7	CAKQLFVMDSGNIHRLITGVITICTKFLSDFFYSNSRYAKVGG--ISLQELNHLQLQFLI	258
pcl8	PIKLLKAKLQEDQAHRIYISTIRIATKLEDFVHSONYICKVFG--ISKRLLTKLISFMA	461
pcl9	VIPKNVYGINTRHRIFLGCCLILAAKTLNDSPLNKHWAAYTEGLRIRREVNTISRELLG	139
pcl10	ILQLKLNQLQEKVHRMIIAAVRLSTKLEDFVHSHYFYSKVCG--ISKRLLTKLISVLLI	404
Clg1	ILIAKRLNQLQEKVHRMIIAAVRLSTKLEDFVHSHYFYSKVCG--ILINVTINDFERQWLR	256

★

----------

Pho80	RVNYRIIPRDHNITLCSIEQKQKQKVIDKNALGSLDLSYSYVNRPKSGYNVLDKYYRRI	224
pcl1	LLNW-----	149
pcl2	YFDWDTISTDDLITC-----	155
pcl5	KLNYELAIIPYDE-----	184
pcl6	LCDFELLI-----	390
pcl7	LCDFKLLVSVVEEMQK-----	273
pcl8	SNVF-----	465
pcl9	YLNNDV-----	145
pcl10	CVCV-----	408
Clg1	IFNH-----	260

----------

Pho80	VQLVGSFNASPKSRKVDYVLPFNIDIVSESGSQTTQLKGSPPNSHSSQKRYSEAKDAH	284
pcl1	-----	
pcl2	-----	
pcl5	-----	
pcl6	-----	
pcl7	-----	
pcl8	-----	
pcl9	-----	
pcl10	-----	
Clg1	-----	

Pho80	IYNKRSEKPDY
pcl1	-----
pcl2	-----M
pcl5	-----
pcl6	-----
pcl7	-----F
pcl8	-----F
pcl9	-----
pcl10	-----F
Clg1	-----

Figure S2.1. Continued

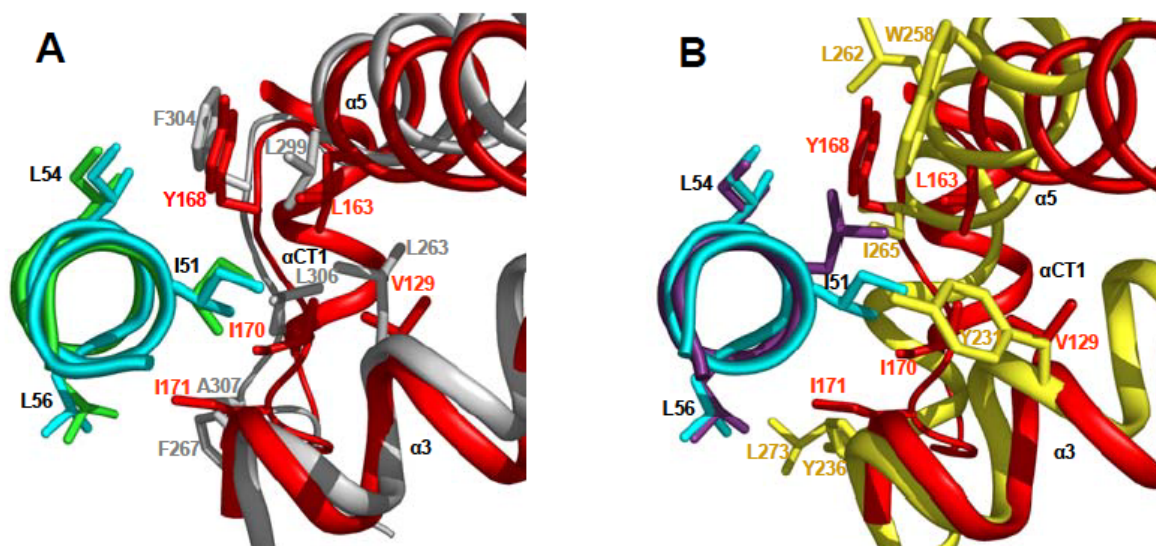


Figure S2.2. Hydrophobic Interactions of the Aliphatic Residues on the PSTAIRE/PSSALRE Helices of Pho85, CDK2, and CDK5 with Residues of Their Respective Cyclins

The PSTAIRE in Pho85 and CDK2 are shown in cyan and green, respectively, and the PSSALRE in CDK5 is shown in purple. The cyclins are depicted in red for Pho80, grey for cyclin A and yellow for p25. Residue identifications (black color for the PSTAIRE helix and red for the cyclin) are for only the Pho85-Pho80 complex. The residues of cyclin A and p25 are shown in grey and orange, respectively. The overlaps in panels A and B were obtained following superposition of the kinase subunit structures. The three PSTAIRE/PSSALRE helices adopt virtually identical position. The hydrophobic interactions between the aliphatic residues on the PSTAIRE/PSSALRE and cyclin residues are summarized below. The locations of the residues on the cyclins are in parenthesis.

Pho85	Pho80	CDK2	Cyclin A	CDK5	P25
I51	V129 ($\alpha 3$) L163 ($\alpha 5$) I170 (loop after $\alpha 5$)	I49	L263 ($\alpha 3$) L299 ($\alpha 5$) L306 (loop after $\alpha 5$)	L49	Y231 ($\alpha 3$) W258 ($\alpha 5$) I265 ($\alpha 5$)
I54	Y168 (loop after $\alpha 5$)	I52	F304 (loop after $\alpha 5$)	I52	L262 ($\alpha 5$)
L56	I171 (loop after $\alpha 5$)	L54	A307 (loop after $\alpha 5$) F267 ($\alpha 3$)	L54	Y236 ($\alpha 3$) L273 ($\alpha CT1$)

(A) Overlap of the Pho85 and CDK2 PSTAIRE helices and parts of Pho80 and cyclin A. Both $\alpha 3$ and $\alpha 5$ in Pho80 and cyclin A adopt roughly a similar position. $\alpha CT1$ in Pho80 has no counterpart in cyclin A. Note that the locations of the three aliphatic residues on the PSTAIREs involved in the interactions with the cyclins are almost identical (summarized in the table above). Moreover, the residues on the cyclins involved in the interactions originate from identical segments (see table). (B) Overlap of Pho85-Pho80 and CDK5-p25. In contrast to that shown in panel A, the $\alpha 5$ helices of Pho80 and p25 barely overlap (see also Figure S2.3). Moreover, in p25, residues from almost entirely the C-termini of helices $\alpha 3$ and $\alpha 5$ and the N-terminus of $\alpha CT1$ make interactions with the large aliphatic side chains on PSSALRE. This difference is attributed mainly to the differences in the geometries of $\alpha 5$ and its neighboring helices between p25 and Pho80 or cyclin A (further discussed in Figure S2.3 legend). Among the three different aliphatic side chains on PSTAIRE or PSSALRE helix, only L49 of CDK5 fails to overlap with the corresponding I51 of Pho85 or I49 of CDK2. This difference may be attributed to the interactions in p25 of L49 with two bulkier aromatic residues (Y231 and W258), thereby possibly nudging L49 farther to one side, as compared with the interactions of I51 of Pho85 or I49 of CDK2 with only aliphatic side chains.

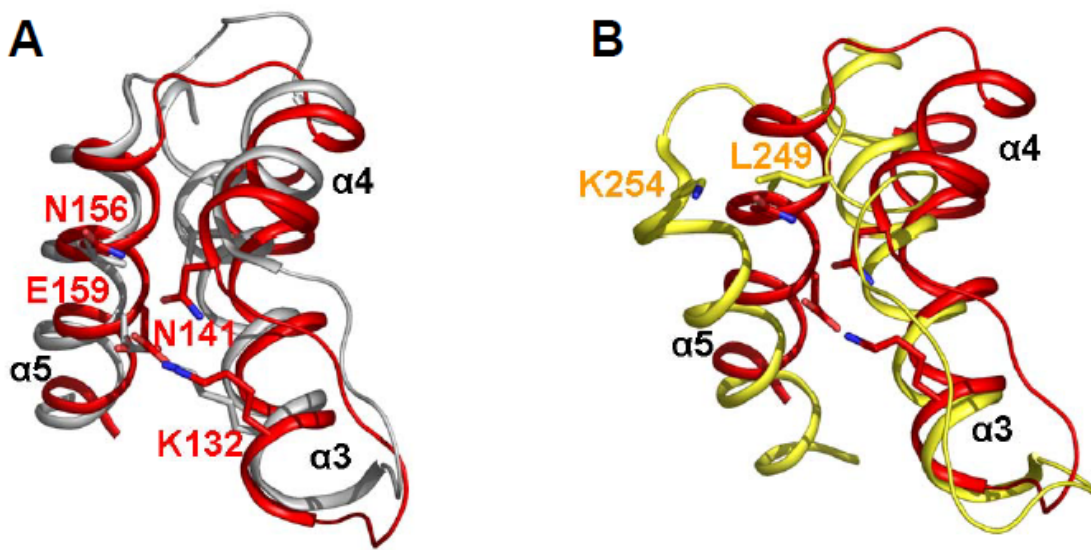


Figure S2.3. The Relationship of the $\alpha 5$ Helices with $\alpha 3$ and $\alpha 4$ of Pho80 or Cyclin A and with $\alpha 3$ of p25

The color schemes for the helices and residue numbers are identical to those described in Figure S2.2. The overlaps shown in panels A and B were obtained following superposition of the kinase subunit structures. (A) The geometries of the helices $\alpha 3$, $\alpha 4$ and $\alpha 5$ in Pho80 (red) and cyclin A (grey). Note the similarities of the geometries in the two complexes, notably all three helices in both cyclins adopting quite similar positions. Moreover, as shown in Figure 2.1A in main text, the $\alpha 5$ helices in Pho80 and cyclin A lie nearly parallel to the PSTAIRE helices. A salt link and a hydrogen bond or hydrophobic interaction contribute to maintaining the close packing of $\alpha 5$ to the cyclin box in both Pho80 and cyclin A. The salt link (lower pair) in Pho80, between E159 on $\alpha 5$ and K132 on $\alpha 3$, is equivalent to that in cyclin A, between E295 on $\alpha 5$ and K266 on $\alpha 3$ (data not shown). A hydrogen bond is also formed in Pho80 between N156 on $\alpha 5$ and N141 on $\alpha 4$ (the upper pair). In place of the hydrogen bond, a hydrophobic interaction is formed in cyclin A, between L292 on $\alpha 5$ and V275 on $\alpha 4$ (data not shown). (B) The relationship between $\alpha 5$ and $\alpha 3$ in Pho80 (red) and p25 (yellow). Despite the claim of the presence of $\alpha 4$ helix in p25 (Tarricone et al., 2001), PDB analysis indicates clearly that it is missing (see also Figure 2.1B in the main text). In contrast to the similar geometry shown in panel A for Pho80 and cyclin A, in p25, the position of $\alpha 5$ differs from that in Pho80 or cyclin A. The p25 $\alpha 5$ helix is farther away from the rest of the cyclin box helix bundle. The angle between the axes of $\alpha 5$ helices in p25 and Pho80 is 27° . Moreover, whereas the $\alpha 5$ helices in Pho80 and cyclin A lie parallel to the PSTAIRE, the $\alpha 5$ helix in p25 is rotated at an angle of about $\sim 25^\circ$ with respect to PSSALRE (Figure 2.1B in the main text). The reasons for the geometrical difference are not clear, but one possibility is the absence in p25 of the two stabilizing interactions between helices $\alpha 3$ and $\alpha 5$ in Pho80 and cyclin A (panel A). Additionally, the juxtaposition and noncomplementarity between L249, which resides on the long loop following $\alpha 3$ and roughly assumes a similar position to that of N156 on $\alpha 5$ of Pho80, and K254 on $\alpha 5$ could be another determinant in attaining the final geometry in p25. Since $\alpha 3$, $\alpha 4$ and $\alpha 5$ of the cyclin box in Pho85 and p25 are engaged extensively in the interaction with the cognate kinase subunits, the differences in the geometries of these helices may account for the plasticity of the CDK-cyclin interfaces. The difference in the position of $\alpha 5$ in p25 as compared to that in cyclin A (or now Pho80) has been attributed to be the consequence of the tight packing of αNT against $\alpha 1$, $\alpha 2$, and the last portion of $\alpha 7$, the equivalent of $\alpha CT2$ (Tarricone et al., 2001). A problem of this explanation is that the packing geometry of the four helices in p25 is not too different from that in Pho80 (see Figure 2.1B in the main text).

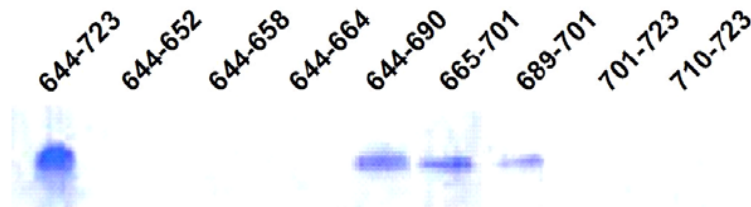


Figure S2.4. Identification of a Shorter Segment of the 80 Residue Fragment of Pho81 that Binds Pho85-Pho80

Binding was demonstrated by pull down assays with the segment-INTEIN-CBD bound to chitin beads serving as a bait to retrieve Pho85-Pho80. The stronger intensity of the Pho85-Pho80 band pulled down by the 80-residue fragment (644-723) is probably a reflection of a better expression of the 80-residue-INTEINCBD construct and a much larger fragment.

SUPPLEMENTAL REFERENCES

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., et al. (1998).

Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D54*, 905-921.

CCP4 (1994). Collaborative Computational Project No.4. The CCP4 programs for protein crystallography. *Acta Crystallogr D50*, 760-763.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D60*, 2126-2132.

Friesen, H., Murphy, K., Breitkreutz, A., Tyers, M., and Andrews, B. (2003). Regulation of the yeast amphiphysin homologue Rvs167p by phosphorylation. *Mol Biol Cell 14*, 3027-3040.

Huang, S., Jeffery, D. A., Anthony, M. D., and O'Shea, E. K. (2001). Functional analysis of the cyclin-dependent kinase inhibitor Pho81 identifies a novel inhibitory domain. *Mol Cell Biol 21*, 6695-6705.

Jeffery, D. A., Springer, M., King, D. S., and O'Shea, E. K. (2001). Multi-site phosphorylation of Pho4 by the cyclin-CDK Pho80-Pho85 is semi-processive with site preference. *J Mol Biol 306*, 997-1010.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A47*, 110-119.

Kantardjieff, K. A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci 12*, 1865-1871.

La Fortelle, E. D., and Bricogne, G. (1997). SHARP program for statistical heavyatom refinement. *Methods Enzymol* 276, 472-494.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 ESF-EACBM News Protein Crystallogr* 26.

Lu, G. (1999). FINDNCS: a program to detect non-crystallographic symmetries in protein crystals from heavy-atom sites. *J Appl Crystallogr* 32, 365-368.

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-497.

Meimoun, A., Holtzman, T., Weissman, Z., McBride, H. J., Stillman, D. J., Fink, G. R., and Kornitzer, D. (2000). Degradation of the transcription factor Gcn4 requires the kinase Pho85 and the SCF(CDC4) ubiquitin-ligase complex. *Mol Biol Cell* 11, 915-927.

Nishizawa, M., Kawasumi, M., Fujino, M., and Toh-e, A. (1998). Phosphorylation of Sic1, a cyclin-dependent kinase (Cdk) inhibitor, by Cdk including Pho85 kinase is required for its prompt degradation. *Mol Biol Cell* 9, 2393-2405.

O'Neill, E. M., Kaffman, A., Jolly, E. R., and O'Shea, E. K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* 271, 209-212.

Reinders, A., Burckert, N., Boller, T., Wiemken, A., and De Virgilio, C. (1998). *Saccharomyces cerevisiae* cAMP-dependent protein kinase controls entry into stationary phase through the Rim15p protein kinase. *Genes Dev* 12, 2943-2955.

Schenk, P. M., Baumann, S., Mattes, R., and Steinbiss, H. H. (1995). Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. *Biotechniques* 19, 196-198, 200.

Schneider, T. R., and Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr D* 58, 1772-1779.

Shemer, R., Meimoun, A., Holtzman, T., and Kornitzer, D. (2002). Regulation of the transcription factor Gcn4 by Pho85 cyclin PCL5. *Mol Cell Biol* 22, 5395-5404.

Tan, Y. S., Morcos, P. A., and Cannon, J. F. (2003). Pho85 phosphorylates Glc7 Protein phosphatase regulator Glc8 in vivo. *J Biol Chem* 278, 147-153.

Tarricone, C., Dhavan, R., Peng, J., Areces, L. B., Tsai, L. H., and Musacchio, A. (2001). Structure and regulation of the CDK5-p25(nck5a) complex. *Mol Cell* 8, 657-669.

Terwilliger, T. C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr D* 55, 849-861.

Weeks, C. M., and Miller, R. (1999). The design and implementation of SnB version 2.0. *J Appl Crystallogr* 32, 120-124.

Wilson, W. A., Mahrenholz, A. M., and Roach, P. J. (1999). Substrate targeting of the yeast cyclin-dependent kinase Pho85p by the cyclin Pcl10p. *Mol Cell Biol* 19, 7020-7030.

APPENDIX B

Chapter 3 Supplemental Data

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Pho7-TAP Expression and Purification

Purification of Pho7-TAP from DP94 cultures grown in either high-Pi or no-Pi media was attempted using modified procedures from (Gould et al., 2004). Pho7-TAP cells were grown to mid-log ($OD_{600}=0.4$) phase at 30°C in 1 L high-Pi media. Cells were collected, washed in no-Pi media, and transferred to either 10 L high-Pi or no-Pi media and grown for another 2 hours. Cells were harvested and frozen drop-by-drop in liquid nitrogen. Samples were lysed using a Retsch Ball Mill (6 x 3 min, 400 RPM) with continuous chilling in liquid nitrogen between pulses. Milled sample was re-solubilized in NP-40 buffer (6 mM Na_2HPO_4 , 4 mM NaH_2PO_4 , 1% NP-40, 150 mM NaCl, 2 mM EDTA, 50 mM NaF, 4 ug/mL leupeptin, 0.1 mM Na_3VO_4 , 1.3 mM benzamidine, 1 mM PMSF, and 1 EDTA-free Complete Protease Inhibitor Tablet (Roche) per 50 mL buffer) and cell debris was pelleted by centrifugation at 20,000 g for 10 minutes at 4°C. Supernatant was subjected to either polyethylenimine precipitation followed by ammonium sulfate (AS) precipitation (Burgess, 1991) or just AS precipitation from 20% and 60% ammonium sulfate (w/v) fractions. The 60% ammonium sulfate pellet was brought up in 25 mL NP-40 for further purification.

The re-suspended 60% AS fraction was incubated with 700 uL of 50% IgG Sepharose 6B resin, rotating, for 1 hour at 4°C. Lysate was packed into a Bio-Rad Poly-Prep Chromatography column and washed with 3 x 10 mL of IPP150 buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP-40) and 10 mL TEV CB (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% NP-40, 0.5 mM EDTA, and 1.0 mM DTT). Remaining IgG beads were then incubated with 400 U acTEV Protease (Invitrogen) in 1 mL TEV CB for

1.5 hours at 16°C. Elute was drained from the column and a second elute fraction was collected with 1 mL TEV CB.

Elute fractions were adjusted to 20% trichloroacetic acid (v/v) and precipitated overnight at -20°C. Samples were precipitated at 20,000 g for 30 minutes at 4°C and washed once in cold acetone-HCl and once in cold acetone. Precipitates were dried to completion using a speedvac and re-solubilized in 50 uL 2xSB. Appropriate samples from the purification process were analyzed using SDS-PAGE and stained with Invitrogen SafeStain, or subjected to western blot analysis using an anti-CBP antibody (Open Biosystems). Representative gels from this process are shown in Figure S3.4.

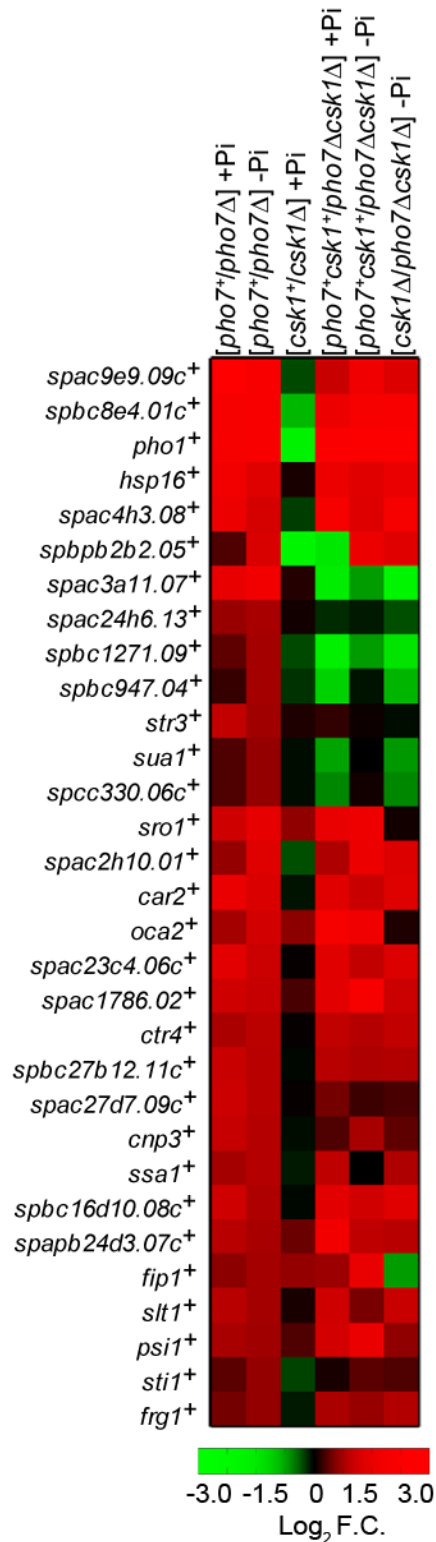


Figure S3.1. Microarray Epistasis Analysis for Pho7-Regulated Genes

Heat map displaying the fold change (log₂ scale) of Pho7-regulated genes (> 1.8-F.C. between *pho7*⁺ and *pho7*Δ strains in no-Pi conditions) in either a *pho7*Δ, *csk1*Δ, or *pho7*Δ*csk1*Δ strain incubated in the indicated conditions as measured by microarray analysis.

Figure S3.2. Secondary Structure Prediction for Pho7.

Displayed is the amino acid sequence of Pho7 overlaid with a secondary structure and solvent accessibility prediction performed using the Predict Protein software suite (Rost and Liu, 2003). A.A.: amino acid sequence (negatively charged (red) and positively charged (blue) residues are shown), R.I.S.: reliability index for secondary structure prediction (0:low, 9:high), P.S.S.: predicted secondary structure (L=random coil, E=extended sheet, H=helix) for residues that pass a $\geq 82\%$ expected average accuracy threshold, R.I.A.: reliability index for solvent accessibility prediction (0:low, 9:high), P.A.: predicted solvent accessibility (b=0-9%, i=9-36%, e=36-100%) for residues that pass a ≥ 0.69 expected average correlation threshold.


```

.....43.....44.....45.....46.....47.....48
A.A-- AVVHHDHFNSTSMDGVAVSNMDETGTSSAGSKPFNRKSRNRSFTNPVGMTEEHFLREYAQH
R.I.S-- 445444555555554445556566666666765555444555566565543344444455
P.S.S-- ..L...TTTTTTTT...LLLLLLLLLLLLLLLLLLLL...LLLLLLLLLL.....L
R.I.A-- 000001010100000001001010100000001010010000010010010100110001
P.A-- .....

.....49.....50.....51.....52.....53.....54
A.A-- SVANPSLLIHQIHGLPSEQVHGLLSHTELGNAMHNQPTYNESSIAAENVNNWMLETNDHE
R.I.S-- 555654333444556643323333433223333456655554434333344444555655
P.S.S-- LLLLL.....LLLT.....LLLLLLL.....LLLLLL
R.I.A-- 100000000000101010100000000100000110000110000011011000110010
P.A-- .....

.....55.....56.....57.....58.....59.....60
A.A-- NLSMQSHFEVPDLKMNHHDSSFFDRHIDQTAMPGQNQHGTVKNMETMHHFYPDVHNSEFP
R.I.S-- 443333334555444555444443333334565545554433321222344555544455
P.S.S-- .....LLL...LLL.....LLL...LLL.....LLL...L
R.I.A-- 10010000100011000001000010001000000010100010000000010101201
P.A-- .....

.....61.....62.....63.....64.....65.....66
A.A-- AAPNFVKSQVPYYYQSQAADDEEEDVPDHQPSWRGRIHSFSIATDSSQHVVERPCIHSLR
R.I.S-- 666665444555555443445446675556542333231122354322344543200123
P.S.S-- LLLLLL...LLLLL.....L...LLLLLLLL.....L.....L.....
R.I.A-- 010000101000000010110100010111001100000111111000100021230211
P.A-- .....

.....67.....68.....69.....70.....71.....72
A.A-- GIHGQDGGLEQHDGDHVNMLPDTHAEELAYTSMLLFHDIPTRDIRPDFNVHELVDHGTY
R.I.S-- 432333575321244422435342277877765433202653344457561110013566
P.S.S-- .....LLL.....L.....HHHHHHH.....LL.....LLL.....LLL
R.I.A-- 020300010000311000101101331454314538820411202110004016411421
P.A-- .....bbb..bb..bb..b.....b..bb..b..

.....73.....
A.A-- PNFHQNQADSFKNHPFRQ
R.I.S-- 431021333111387779
P.S.S-- .....LLLLL
R.I.A-- 000111203203443226
P.A-- .....ei...e

```

Figure S3.2. Continued

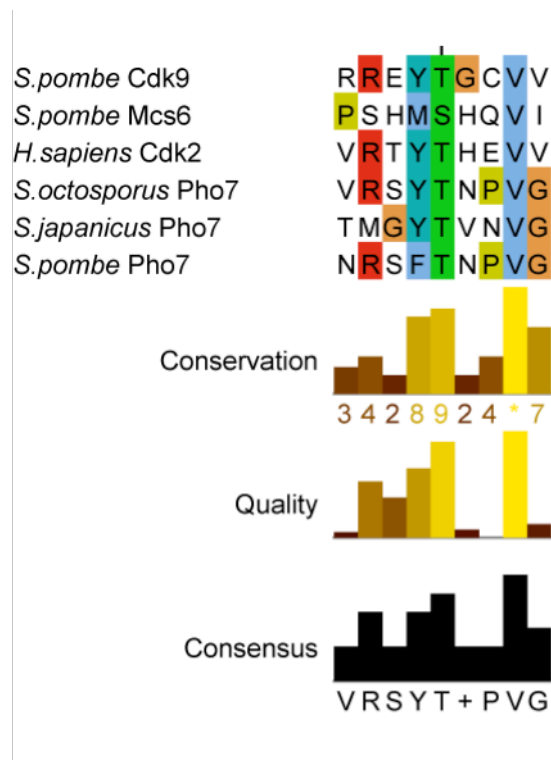


Figure S3.3. Conservation of Residue T463 in Pho7.

Shown is an alignment performed on the 9 residue sequence surrounding the T463 residue in *S. pombe* Pho7. T463 was identified as a phosphorylated residue in a global proteome study (Wilson-Grady et al., 2008). The Pho7 motif was aligned against both its orthogroup members (*S. octosporus* and *S. japonicus*) and previously identified targets for Csk1 kinase activity (see text) using the ClustalW2 algorithm (Goujon et al., 2010; Larkin et al., 2007). Residues are colored based on identity, and the consensus sequence is shown below. Jalview was used to produce the conservation plot (Waterhouse et al., 2009).

Figure S3.4. Purification of Pho7-TAP.

(A) Schematic of the Pho7-TAP purification workflow. For specific details pertaining to each step in the purification see the supplemental materials and methods. (B) Representative SDS-PAGE gel (top) and western blot (WB, bottom) from the Pho7-TAP workflow. See supplemental materials and methods for further details. Fractions analyzed were: WCE: whole cell extract, Post-Spin: supernatant from the hard spin, Post-AS: re-solubilized 60% ammonium sulfate pellet, IgG FT: flow-through from the IgG sepharose 6B resin, Pre-TEV Cut: pooled IgG bound material before addition of acTEV protease, TCA Elute: final TCA precipitated elute. Box in the top image represents the area where Pho7-CBP is thought to elute. The *-arrow in the bottom image represents the predicted Pho7-TAP molecular weight, the arrow shows the predicted size for Pho7-CBP post-TEV cleavage. The antibodies and imaging kit used are listed.

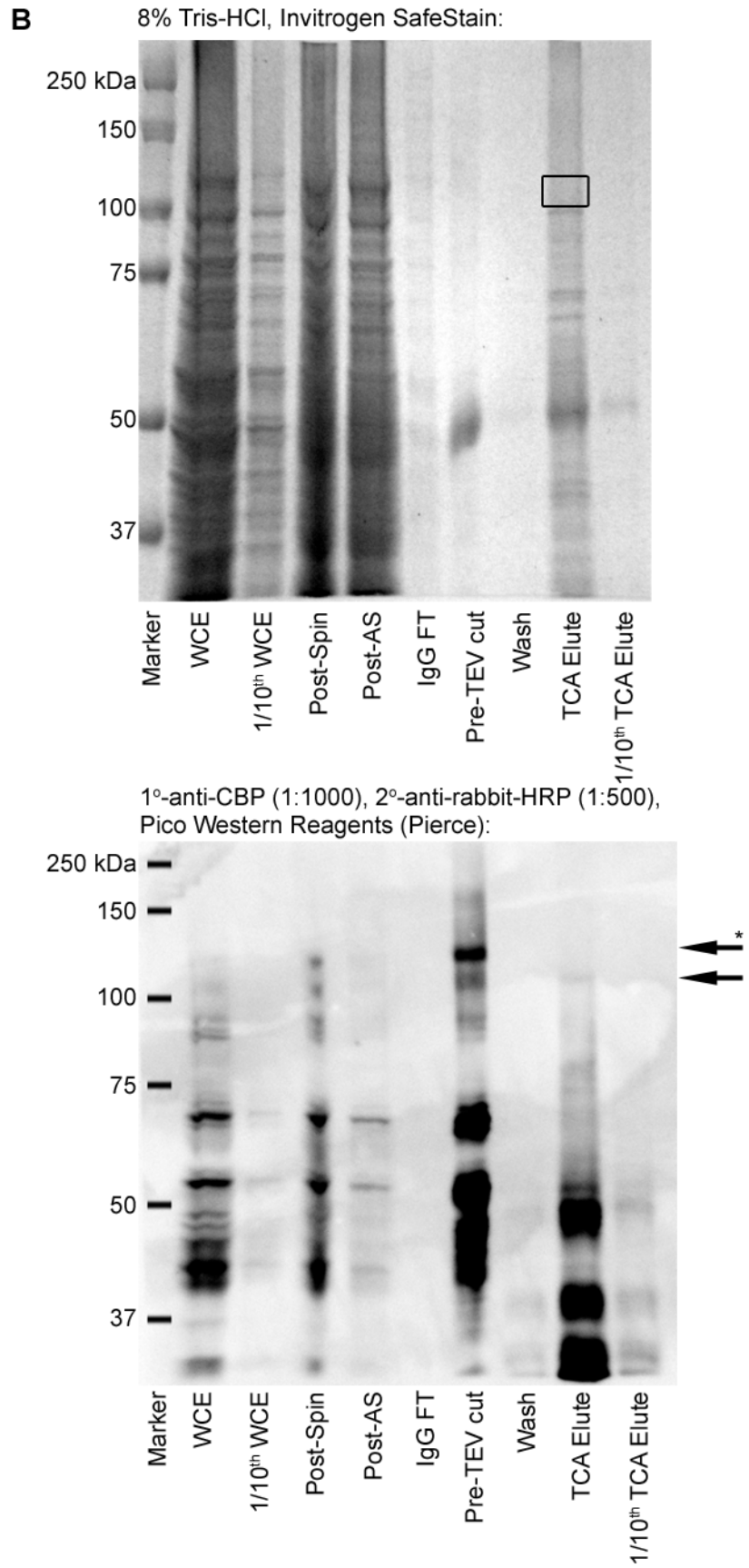
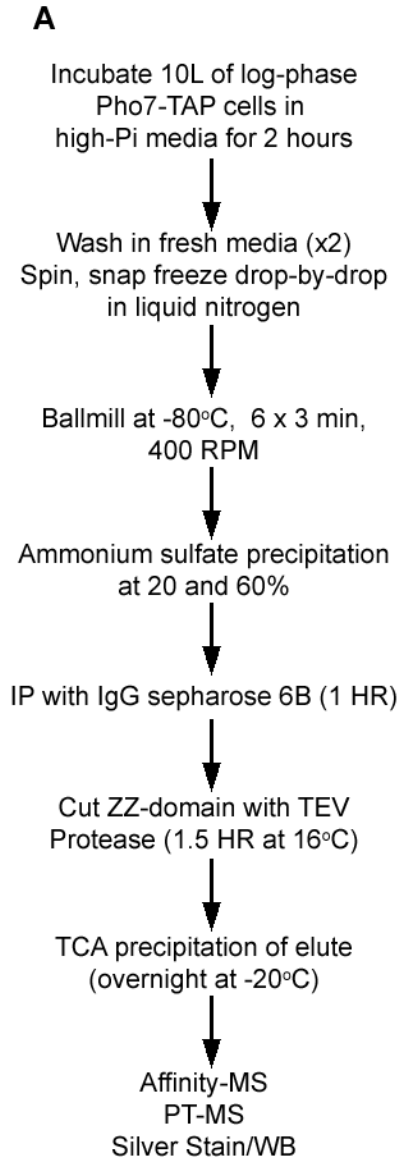


Figure S3.4. Continued

Figure S3.5. ChIP-Seq Binding Profiles for Pho7-TAP at Non-PHO Promoters. Shown are ChIP-Seq profiles for the genes identified in Figure 3.6. As previously described, wild-type cells containing Pho7-TAP were grown in either high-Pi (blue) or no-Pi (red) conditions and ChIP-Seq libraries were prepared from purified DNA. For comparison, the ChIP-Seq signal from mock (black) cells grown in no-Pi and *csk1Δ* (cyan) cells incubated in high-Pi is included. The gene product of interest is plotted based on transcript direction with the plus (+) strand above and the minus (-) strand below. Reads were normalized as described in the text and the location within the genome is plotted on the x-axis.

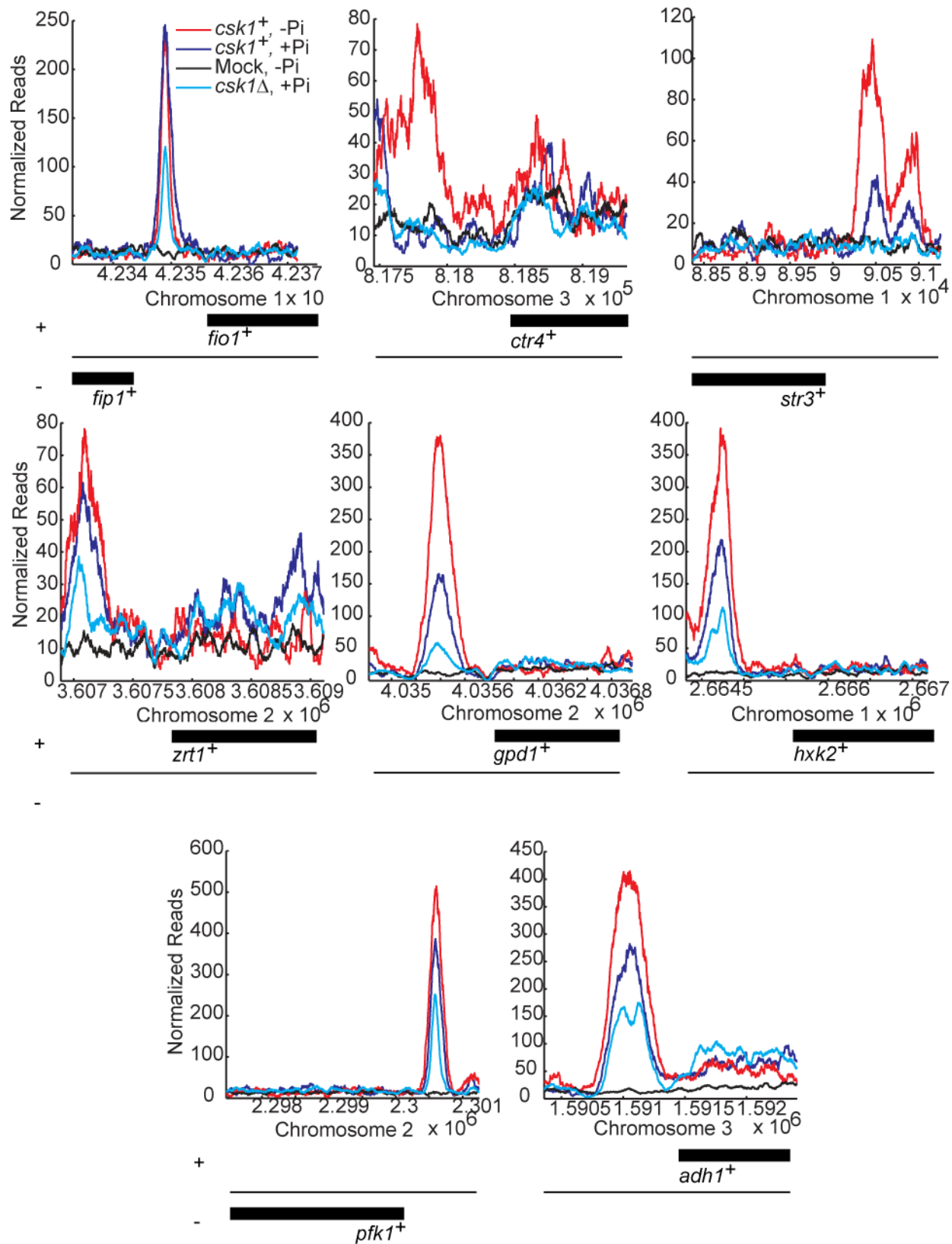


Figure S3.5. Continued

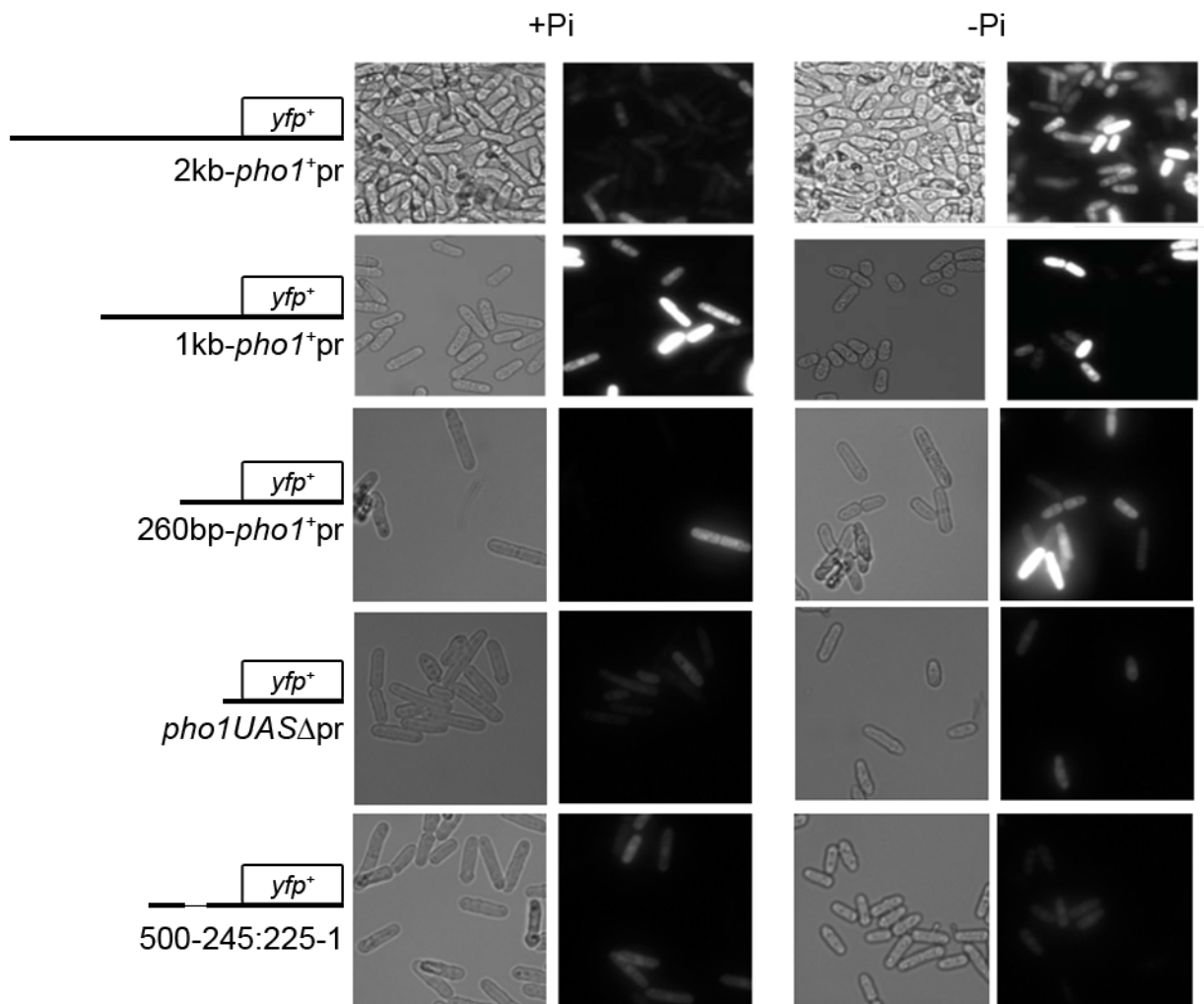


Figure S3.6. Fluorescent Microscopy of *pho1*⁺pr-*yfp*⁺ Constructs.

Shown are representative images for *pho7*⁺ strains transformed with the listed *pho1*⁺pr-*yfp*⁺ constructs. Cells were grown overnight in high-Pi media then split into high- or no-Pi media (+Pi, -Pi), incubated for 4 hours, harvested, and subjected to fluorescent microscopy. Confocal images of the magnified fields are shown to the left for clarity.

Table S3.1. Primers Used in Chapter 3 Study

Primer	Sequence	Purpose
ISP001	CGGAAACCCAGTGATTTTGT	ChIP-qPCR on <i>pho1</i> ⁺
ISP002	TTGTTGCTGACCACCATTTT	
ISP003	TTACGGGAAGTGCCGATATT	
ISP004	GGATGTTTCAAGCAGCGACT	
ISP005	TTTTGGTTCGGGACACATTT	
ISP006	AACTAGCACATTGCCGTTCC	
ISP007	TCGTTTGTGACAGGTTGGAA	
ISP008	GCCAAAAAGCCAAGGAATAA	
ISP009	TGGTTCGGGACACATTTACA	
ISP010	CACAAACGAAACCAACAACG	
ISP011	GTGCGCTTCAATGAGTTAACACTC	ChIP-qPCR on <i>spo3</i> ⁺
ISP012	AGTGTGGGGAAACTTTTATTAATC	
ISP013	GGTATTCATGAGGCTACTTACAAC	RT-qPCR on <i>act1</i> ⁺
ISP014	CCGCTCTCATCACTCTTGCTTG	
ISP015	GCTAATAGCCTTGCAAATTCCTC	RT-qPCR on <i>pho1</i> ⁺
ISP016	CTAGCATTGCAAAGAGTGTCAAAG	
ISP019	GACCCGACTTTAACGTTTCATGAGC	<i>pho7</i> ⁺ sequencing
ISP020	CGGCTATGCCTGGACAAAAATC	
ISP021	GTGGCGGTATCTAACATGGACG	
ISP022	ATGACGCTCGTCTTGTGACG	
ISP023	ACAACCTCTCCCTATCGTCTG	
ISP024	CATATCTCTTCTCCCTCTGC	
ISP083	GGATCAAGAGTCGCACACATACTA	RT-qPCR on <i>fio1</i> ⁺
ISP084	ATCGTCATTCTCGTCTAACTCCTC	
ISP085	TTGGAACCTGGTATACTATCGATGC	RT-qPCR on <i>ctr4</i> ⁺
ISP086	TGAACAGATGAAGAGGCCCTTAGTA	
ISP087	CTTACTTCTCAGTTGTCTCCGTTG	RT-qPCR on <i>gpd1</i> ⁺
ISP088	GTTATAAACGGCAGTGAACAAGG	
ISP089	AACTGATTTCTCTTGCCTTCACTC	RT-qPCR on <i>hvk2</i> ⁺
ISP090	ACACCAACATCACTCTTCTTACCA	
O0375	TAGGTTATGGTTAATGTG	RT-qPCR on <i>yfp</i> ⁺
O0376	TAAACATTGTGAGAGTTATA	
O0393	CATATTTCTTTTGTCTTTTTCGTTTGTGCGGATCCCCGGGTTAATTAA	<i>pho7</i> Δ
O0394	GAAAAACAAAGGTAATATGCGCAAAAAGAAATTCGAGCTCGTTTAAAC	
O0395	GTACCAAGTTGTTGATCATAAAGG	
O0396	TTAATTAACCCGGGGATCCGACAAACGAAAAAGACAAAAGAAATATG	
O0397	GTTTAAACGAGCTCGAATTCCTTTGCGCATATTACCTTTGTTTTTC	
O0398	CATCTGTACGATCACATTTATCC	
O0399	CGACTTCGCCCTAACACGAT	
O0568	GGGACAACACCAGTGAATAATTCTTACCTTTAGACATTGAACGCTACCTTAAAAAGTTG	Gap repair with <i>yfp</i> ⁺ plasmid for <i>pho1</i> ⁺ promoter
O0569	GCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCAGAGTGACATATAGTCTAC	
O0570	GCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCAGTGGTGGTGGATTTGCTTG	
O0571	GCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCTTGGTTTTGAGCGTTCATGC	
O0604	ATGTCTAAAGGTGAAGAATTATTC	Getting <i>yfp</i> ⁺ fusion product for pombe gap repair
O0605	GCAGTGATGTAACATTGTGAG	
O0869	TTTAACGTCAAGGAGAAGGAATTCAGCTGACCACCATGCAGAAGAAAGTCGTTTCAG	Pho7-TAP tagging
O0870	CATCTTCTCTTCGATCCCCGGGAATTGCCATGCCCTTATTGTCTAAAGGGATGAT	
O0935	ATGAAATCAGTCGGTCATTTCCGTTCCGTTGGTCACTGATACGGATCCCCGGGTTAATTAA	<i>csk1</i> Δ
O0936	CTTTAGGGAATGACTCAATCACGAAGGATGGTGAAGCCCGGAATTCGAGCTCGTTTAAAC	
O0969	CGAAACCGAAGTTTTGCTAATCCAGTTGGTATG	Mutate T463 to A in <i>pho7</i> ⁺
O0970	CATACCAACTGGATTAGCAAACCTCGGTTTCG	
O1103	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGAGAATCGGCATTACGTG	Gap repair with <i>yfp</i> ⁺ plasmid for <i>pho1</i> ⁺ promoter
O1104	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGACATGTTTGAACAACAGGG	
O1105	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGAAATAGTCGCTGCTTGAAC	
O1106	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCATCCGAAATAGTTTTATTC	
O1107	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGAAATTAACAAACAGTTTATG	
O1108	GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGTTGTTGGTTTCGTTTGTG	
O1109	GGACATGTTTGAACAACAGGCATCCGAAATAGTTTTATT	
O1110	AATAAAACTATTTCCGGATGCCTGTTGTTCAAACATGTCC	

SUPPLEMENTAL REFERENCES

Burgess, R.R. (1991). Use of polyethyleneimine in purification of DNA-binding proteins. *Methods Enzymol* 208, 3-10.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38, W695-699.

Gould, K.L., Ren, L., Feoktistova, A.S., Jennings, J.L., and Link, A.J. (2004). Tandem affinity purification and identification of protein complex components. *Methods* 33, 239-244.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.

Rost, B., and Liu, J. (2003). The PredictProtein server. *Nucleic Acids Res* 31, 3300-3304.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189-1191.

Wilson-Grady, J.T., Villen, J., and Gygi, S.P. (2008). Phosphoproteome analysis of fission yeast. *J Proteome Res* 7, 1088-1097.