Title

Systemic approaches to identifying substrates of the budding yeast cyclin [sic] dependendant kinanse Pho85

Permalink

https://escholarship.org/uc/item/6kj5q4vw

Author

Howson, Russell W.

Publication Date

2003

Peer reviewed|Thesis/dissertation

Systematic Approaches to Identifying Substrates of the Budding Yeast
Cyclin Dependendent Kinase Pho85

by

Russell W. Howson

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of
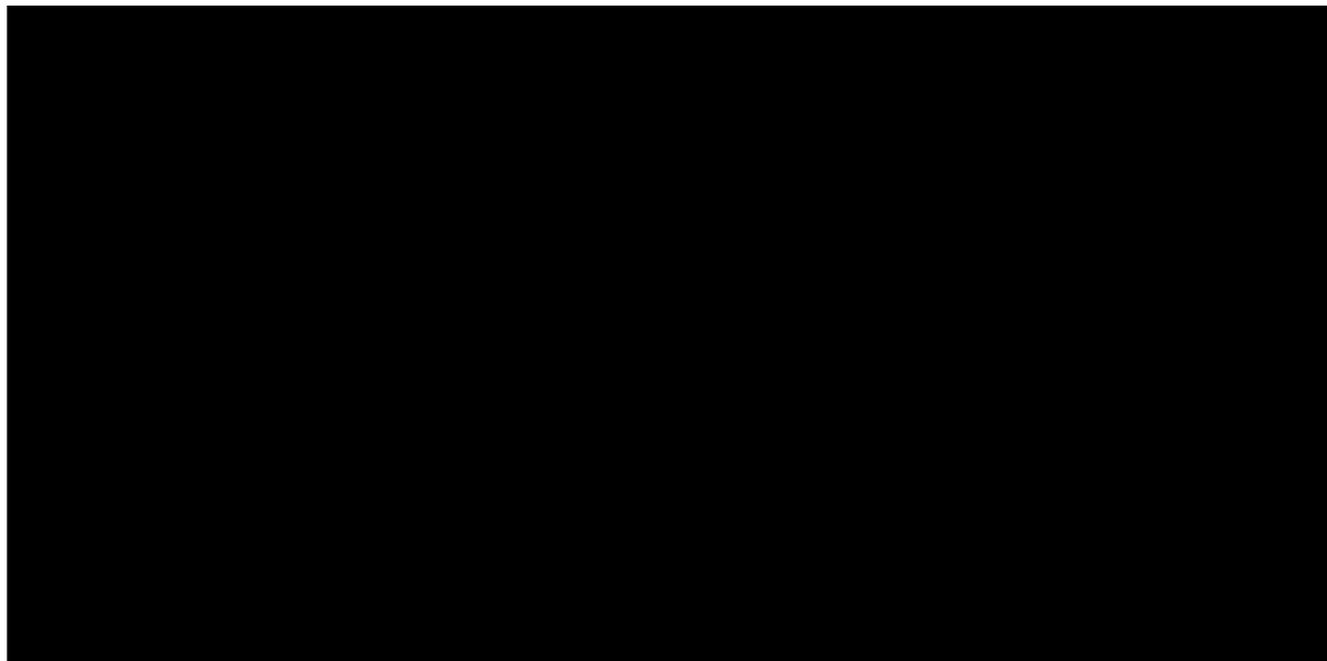
DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

For all the people who supported me during my scientific and personal development

My thesis committee members: Erin O'Shea, Jonathan Weissman, Peter Walter and Reg Kelly, for being patient with me even when I gave them plenty of reasons not to be

Susan Shih, my true love

Dr. Barbara Sharp

Abstract

PHO85 encodes a cyclin dependent kinase (CDK) in the budding yeast *S. cerevisiae*. Though best characterized in its role in regulating phosphate responsive genes, Pho85 has also been implicated in many other cellular processes, and associates with a family of ten cyclins. To understand the molecular functions of Pho85, I have attempted to identify novel substrates of Pho85 kinases, and have taken systematic approaches toward doing so. Here I describe these approaches, as well as the construction of two new collections of yeast strains suitable for global analysis of the proteome. Further, I discuss the identification of putative substrates of Pcl1/Pho85 by a novel substrate screening method.

Table of Contents

List of Figures and Tables

Chapter 1:

Introduction

The phosphoproteome

Phosphorylation is a reversible covalent modification which is ubiquitous in the signaling pathways of all biological organisms. It can regulate proteins in a variety of ways: by altering the catalytic activity of a protein, changing its subcellular localization, promoting or interfering with protein-protein interactions, protein-DNA interactions, as well as influencing protein stability. The versatility of this modification makes phosphorylation a useful mode of protein regulation under many cellular contexts. As such, phosphorylation serves as a major source of regulation in virtually all cellular pathways.

The importance of this modification is underscored by the fact that some one third of all eukaryotic gene products are covalently modified by phosphorylation(1, 2). Further, kinases, the enzymes which catalyze this modification, are implicated in innumerable processes and, along with transcription factors, comprise one of the largest families of proteins in eukaryotes(3, 4).

Despite the importance of this modification, discovery of the phosphoproteins, and the specific kinases responsible for their phosphorylation, has proved difficult. In yeast, there are 122 predicted protein kinases(3), and roughly 2000 phosphoproteins [based on (1, 2)]. Kinases can be easily identified by protein sequence, as all of these proteins contain significant homology in the ATP binding pocket. The substrates of these kinases, however, have not been so easy to identify, due to the poorly understood chemistry of kinase-substrate

interactions. as the chemical requirements for phosphorylation are not stringent. As such, neither the identity of the phosphoproteins or the specific kinase-substrate interactions can be predicted from primary protein sequence. This is highlighted by the fact that few of these phosphoproteins have been identified, and in even fewer cases is the relevant kinase known.

As a result of this difficulty, the regulatory kinases for many biological processes have been identified, but the substrates of these enzymes are not known. One example of this is in the cell division cycle, for which the sequence of events of this process is exquisitely controlled by cyclin dependent kinases (CDKs). Although the identities of most of the players involved is known due to the extraordinary amount of work has gone into understanding this process in the past decades, and much is known, but the substrates of these kinases, and how they carry out the nuts and bolts of the cell cycle, remain a mystery.

Traditional biochemistry has not proved successful for identification of kinase substrate interactions. This is partly due to two factors. First, identification of protein kinase substrates is the reverse of typical biochemical purifications, in which an activity is generally purified for a known substrate. Second, the large number of kinases present in an extract makes development of a specific assay difficult. The ubiquity of this enzymatic activity also complicates any *in vitro* approaches.

The approaches towards identifying substrates of protein kinases which have been successful have mostly been in cases where a candidate substrate is

already suggested by other lines of genetic or cell biological evidence [for example, see (5)]. In these cases, fine tuned biochemical assays, which at time are laborious, have been employed to confirm the validity of the candidate substrate. As a general approach, the obvious failing of these methods is that a candidate substrate is not always suggested by other lines of evidence, and the validation methods utilized are not readily amenable to larger scale or unbiased analyses of potential substrates.

Some inroads into understanding the phosphoproteome have also been made through the use of mass spectrometry. Both general characterization of the yeast phosphoproteome(1, 6) and identification of phosphopeptides associated with specific growth conditions(7) have been performed with some success. These approaches have the advantage of directly examining physiologically relevant conditions, and could potentially be used to identify the phosphorylated effectors of a biological process without knowledge of the associated kinase or signaling pathway. These methods, however, have significant limitations in their ability to deal with complex mixtures of proteins, and thus far have only been able to detect the most abundant phosphoproteins.

A general approach towards identifying the substrates of protein kinases, then, requires both the sensitivity and accuracy of a fine tuned biochemical assay and the global, unbiased nature of a mass spectrometry based approach. Thus far, this combination has been difficult to achieve, resulting in the paucity of knowledge of kinase-substrate interactions.

The budding yeast CDK Pho85

The kinase I have chosen to study is with the budding yeast cyclin

dependent kinase Pho85*(8)*. *PHO85* encodes one of seven CDKs in yeast, and

is the only one which is both non-essential and binds to multiple cyclin partners.

As such, it represents a unique opportunity to take advantage of the genetic and

biochemical approaches available in yeast to understand the molecular functions

of CDKs. The best characterized function of Pho85 is in the regulation of

phosphate responsive genes*(9)*, but the pleiotropic phenotype of a *pho85*Δ strain

implicates Pho85 in the regulation of many different cellular processes, including

phosphate and glycogen metabolism, polarized growth and morphogenesis, G1

progression, carbon source utilization, response to DNA damage, and others

[reviewed in *(8)*]. A *pho85*Δ strain is also sensitive to several chemical agents,

including hydroxyurea, hygromycin B, G418, salt, Calcofluor white, calcium

chloride, cycloheximide, rapamycin, and tunicamycin*(10)*. These sensitivities

implicate Pho85 in responding to the various stimuli.

A chemical genetics approach coupled with transcriptional profiling

identified a plethora of transcriptional effects of instantaneous loss of Pho85

function, including a role in the environmental stress response (ESR)*(11)*. By

comparison with a *pho4*Δ mutant, Carroll was able to distinguish these responses

from those involved in phosphate responsiveness. Interestingly, many of the

transcriptional effects were transient, and were not observed in a strain with

genomic disruption of PHO85. This suggests that the *pho85Δ* strain may have adapted in some ways, and that the processes in which Pho85 has been previously implicated by this strain may be an underestimate.

Ten cyclin binding partners for Pho85, or PCLs (<u>P</u>ho85 <u>C</u>yc<u>L</u>ins) have been identified (*PCLs 1,2,5-10, PHO80,* and *CLG1*), primarily through sequence homology and two-hybrid screens*(12)*. They can be grouped by sequence homology into two subfamilies, the Pcl1 subfamily (*PCLs 1,2,5,9,* and *CLG1*) and the Pho80 subfamily (*PHO80, PCLs 6,7,8,10*). In some cases, a single Pcl, or group of Pcls, is implicated in a Pho85 mediated process, but in most cases, even this type of general link cannot be made. In even fewer cases is the actual Pho85 substrate known. Only four substrates have been shown to be phosphorylated in a Pho85 dependent manner *in vivo*: Pho4, Gsy2, Gcn4, and Rvs167*(5, 13-15)*.

The transcription factor Pho4, which is phosphorylated by Pho80/Pho85 in regulating expression of phosphate responsive genes, provides a rare well characterized example of the different ways in which phosphorylation can affect a protein. Phosphorylation of Pho4 promotes binding to the export receptor Msn5*(16)*, inhibits binding to the import receptor Pse1*(17)*, prevents binding to the transcriptional coactivator Pho2*(18)*, and selectively modulates its activity at different promoters by an unknown mechanism (M. Springer and D. Wykoff, unpublished data).

6

Unlike its phosphorylation of Pho4, a great deal has yet to be discovered about how Pho85 is functioning at a molecular level in the processes outlined above. Even the identity of many of the downstream effectors of Pho85 function, the first level of which are the substrates of the various Pcl/Pho85 kinases, remains a mystery.

Pcl1/Pho85 is one Pho85 kinase for which little is known about its downstream effectors. It has been known that *pho85Δ* is synthetic lethal with disruption of the G1 cyclins *CLN1* and *CLN2(19, 20)*. The synthetic lethality of *pho85Δcln1Δcln2Δ* is recapitulated by disruption of *PCL1* and *PCL2*; the *pcl1Δpcl2Δcln1Δcln2Δ* strain is inviable, implicating these two Pcls in G1 progression, which is supported by the fact that expression of these genes peaks in G1. *pcl1Δpcl2Δ* is also synthetic lethal with *bck1Δ, mpk1Δ, cdc42-1,* and *cla4Δ,* and *pcl1Δ* is synthetic lethal with *bem2Δ(21),* providing strong evidence that these Pcls have functions involved in polarized growth and morphogenesis. This is also consistent with the random budding pattern and abnormal morphology observed in the *pho85Δ* strain*(12, 22)*.

Further synthetic lethal analysis with the *pho85Δ* strain identified 53 gene disruptions synthetic lethal with *pho85Δ(10)*. About half of these are disruptions in genes known to be involved polarized growth and in cell wall maintenance and regulation. In particular, many members of the PKC pathway (including *BCK1, MPK1* and *BEM2*) were identified. Though the authors of this study did not evaluate the synthetic lethality of these disruptions with *pcl1Δ* or *pcl1Δpcl2Δ,* we

can hypothesize that a number of them would be synthetic lethal with these strains given the observations above.

As described above, a compelling amount of genetic information implicates Pcl1 and/or Pcl2 in G1 progression, polarized growth and morphogenesis, and suggests a connection with the PKC pathway. How Pho85 acts in these processes at a molecular level, however, is not at all clear. Discovery of the substrates of these kinases will be the first step in understanding how Pho85 carries out its various cellular functions. Identification of substrates of the Pcl1/Pho85 kinase, therefore, is the goal of the course of study outlined in this dissertation.

The methods I employed in my search for Pcl1/Pho85 substrates took advantage of knowledge of the complete sequence of the *S. cerevisiae* genome*(23)*. By knowing all of the ORFs in budding yeast, we were able to approach this problem in a systematic way, applying biochemical assays, either standard or novel, in a parallel and efficient manner. In chapter 2, I describe studies using the GST collection *(24)*, and the difficulties which we encountered in interpreting the data. In Chapter 3, I discuss the construction of two new collections of yeast, in which essentially every ORF is tagged with the TAP and GFP tags respectively. I adapted the TAP collection to a novel high throughput kinase assay described in Chapter 4.

An underlying theme of the dissertation is the potential power and pitfalls in undertaking approaches of a global or systematic nature, and the fact that the

resulting data is of a new genre, and requires a different conceptual framework to

interpret and understand.

Literature cited

1. Ficarro, S.B. et al. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat Biotechnol* **20**, 301-305 (2002).

2. Ahn, N.G. & Resing, K.A. Toward the phosphoproteome. *Nat Biotechnol* **19**, 317-318 (2001).

3. Hunter, T. & Plowman, G.D. The protein kinases of budding yeast: six score and more. *Trends Biochem Sci* **22**, 18-22 (1997).

4. Lander, E.S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).

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

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

7. Zhou, H., Watts, J.D. & Aebersold, R. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* **19**, 375-378 (2001).

8. Carroll, A.S. & O'Shea, E.K. Pho85 and signaling environmental conditions. *Trends Biochem Sci* **27**, 87-93 (2002).

9. Lenburg, M.E. & O'Shea, E.K. Signaling phosphate starvation. *Trends Biochem Sci* **21**, 383-387 (1996).

10. Huang, D., Moffat, J. & Andrews, B. 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 (2002).

11. Carroll, A.S., Bishop, A.C., DeRisi, J.L., Shokat, K.M. & O'Shea, E.K. Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proc Natl Acad Sci U S A* **98**, 12578-12583 (2001).

12. Measday, V. et al. A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. *Mol Cell Biol* **17**, 1212-1223 (1997).

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

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

15. Huang, D. et al. 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 (1998).

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

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

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

19. Measday, V., Moore, L., Ogas, J., Tyers, M. & Andrews, B. The PCL2 (ORFD)-PHO85 cyclin-dependent kinase complex: a cell cycle regulator in yeast. *Science* **266**, 1391-1395 (1994).

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

21. Lenburg, M.E. & O'Shea, E.K. Genetic evidence for a morphogenetic function of the Saccharomyces cerevisiae Pho85 cyclin-dependent kinase. *Genetics* **157**, 39-51 (2001).

22. Gilliquet, V. & Berben, G. Positive and negative regulators of the Saccharomyces cerevisiae 'PHO system' participate in several cell functions. *FEMS Microbiol Lett* **108**, 333-339 (1993).

23. The Yeast Genome Directory. *Nature* **387**, 3-105 (1997).

24. Martzen, M.R. et al. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153-1155 (1999).

Chapter 2:

Studies with the GST collection

Introduction

One approach which we took to attempt to identify novel substrates of Pho85 was to screen a library of purified proteins which theoretically represent the entire proteome. The GST collection constructed by Phizicky *et al* (*1*) is a library of N-terminal GST fusions to every ORF in the yeast genome. The fusions are maintained on high copy plasmids, under the control of a copper inducible promoter. The fusions are readily purified by induction of a culture with copper, lysis and subsequent pulldown with glutathione agarose. An important feature of the GST library is its organization: the strains are organized in a matrix and two different pooling schemes are used. In one, all strains in a given row are pooled; in the other all strains in a given column are pooled. In the resulting pools, each ORF in the genome is uniquely represented in one row and one column.

By purifying the GST fusions from pools of cells, one can perform the purification of many proteins in parallel. This enables the systematic purification of all yeast proteins with a relatively small number of purifications. In searching for a biochemical activity, then, one simply screens all the protein pools with a chosen assay. A given hit should be represented once in the rows and once in the columns, so the intersections of these hits in the matrix immediately define unique ORFs.

This library presented an opportunity to efficiently screen for kinase substrates, using purified kinase and the purified protein pools as potential substrates. We screened the pools with both Pho80/Pho85 and Pcl1/Pho85. We identified many substrates, though of questionable relevance, since we did not

14

identify Pho4 as a substrate of Pho80/Pho85. We also investigated the quality of the library and the reason for the lack of identification of Pho4.

Methods

Purification of GST fusion proteins

Purification of proteins in the GST collection was performed as described(*1*).

Purification of Pho85 kinases

Purification of Pho80/Pho85 and Pcl1/Pho85 kinases was performed as described for Pho80/Pho85(*2*).

Pho80/Pho85 and Pcl1/Pho85 kinase assays

Kinase assays used for the screens were essentially as described for kinase assays on Pho4(*2*), except using purified GST pools as potential substrates, and with minor modifications to account for the higher glycerol content of the protein pools. We used 100 nM of each kinase in the reactions.

Results

Purification of the GST library

The lab undertook the task of purifying the protein pools which, although made more efficient by the approach, still was quite laborious. In analyzing the efficiency of purification, as well as the final product, we were somewhat

15

disappointed: the GST purification was not very efficient, and the resulting purified proteins were not readily detectable by Bradford assay, and only occasionally visible on silver stain gels. GST protein itself, however, was frequently visible by silver stain in these pools, suggesting that much of what was purified was simply GST alone.

Screening purified GST pools with Pcl1/Pho85 and Pho80/Pho85

Despite the uncertainty over what these protein pools actually represented, we proceeded with the kinase screen. We purified Pcl1/Pho85 and Pho80/Pho85 from E. coli, and developed a relatively efficient kinase assay to be used with the pools. We screened all 160 pools with both kinases side by side, reasoning that true substrates would display some specificity for one kinase. As seen in Figure 1, quite a few proteins were phosphorylated in these screens, especially when Pcl1/Pho85 was used as the kinase, so initially this approach appeared promising.

At this point, however, we made several important observations about the screen and about the utility of the GST pools. Most importantly, we did not identify the known Pho80/Pho85 substrate Pho4 in this screen, which cast some doubt on the accuracy and/or the completeness of the screen. Second, we did not see a one to one correspondence of phosphorylated proteins between the two sets of pools. This not only makes it apparent that some differences exist between the quality of the pools of the rows and columns, but makes it difficult to identify substrates by the intersection of a row and column. Exacerbating this

16

problem was the number of positives identified in the screen, many of which were of similar apparent molecular weights.

We initially chose to focus on Pcl1/Pho85 hits within a defined size range which appeared to have possible corresponding bands in both sets of pools. We selected the individual strains from the GST collection representing these possible intersections, and repooled them into a smaller matrix. After purifying proteins from these new pools, we repeated the kinase assay on the purified proteins. The intersections of this screen definitively identified the unique GST fusions phosphorylated. We noticed that we obtained better yields of proteins, as well as more robust signal in the kinase assays, with the new, smaller, pools.

We also followed up one hit from the Pho80/Pho85 screen, though by a slightly different method, as this hit was only represented in the column pools and had no corresponding hit in the row pools. We identified the substrate by screening all the proteins represented in this column, again in a reorganized matrix fashion.

After this deconvolution, we ended up with six candidate substrates for Pcl1/Pho85 (ARP8, MED2, MVP1, YLR125W, YNL215W, and YGL242C) and one candidate for Pho80/Pho85 (YLR257W). No clear pattern emerges from the functions of these proteins (where known), and only tenuous links can be made to known Pho85 phenotypes.

We followed up these intersections by confirming phosphorylation in the individual strains, and then attempted to assess whether or not these were true physiological targets of Pho85. We first genomically integrated an epitope tag at

the C-terminus of each putative substrate in wild type, *pho85Δ*, and

PHO85(F82G) strains. Using these strains, we metabolically labeled cells with

^{32}P labeled orthophosphate, immunoprecipitated the substrate and compared the

degree of phosphorylation in conditions where Pho85 activity is present (wild

type or *PHO85(F82G)* without inhibitor present) to conditions in which it is absent

(*pho85Δ* or *PHO85(F82G)* strain with inhibitor).

With this approach, we had a great deal of difficulty confirming the

physiological relevance of these substrates. Some of this can be attributed to the

difficulty in the technique, but there was also a great deal of variation in the

results obtained. In particular, we had significant difficulty in consistently labeling

the cells, so accurate quantitation of phosphorylation was made impossible.

Investigation into the failure to identify Pho4 as a substrate of Pho80/Pho85

We investigated why Pho4 was not identified as a substrate of

Pho80/Pho85. We considered three possibilities: first, Pho4 may not have been

successfully tagged during the construction of the collection. Second, GST fusion

to Pho4 could either render the protein nonfunctional, or somehow inhibit its

proper phosphorylation. Lastly, the fusion protein may not have been

successfully purified, either because the GST tag is not accessible or because of

some failure in the pooled format.

As the strains in the this collection are not derived from individual colonies,

and were not verified for successful tag fusion, it is quite possible that Pho4 is not

tagged correctly in this collection. We addressed this possibility by rescuing

plasmids from the individual GST-Pho4 strain; we transformed into E. coli and picked 18 individual colonies which should represent individual plasmids in the starting strain. After preparing plasmid from each of these colonies, we assessed the presence of the PHO4 ORF by restriction digest. 17 of 18 isolates displayed the correct digest pattern (data not shown), so we conclude that failure during construction is not the cause of the failure of Pho4 to be identified.

As Pho4 seems to have been successfully GST tagged, we assessed whether this represents a functional fusion, capable of being phosphorylated by Pho80/Pho85. To do so, we purified GST-Pho4 individually. We were able to successfully purify GST-Pho4; in our preparation, GST-Pho4 is detectable by silver stain, as well as by both anti-Pho4 and anti-GST antibodies. We also performed a kinase assay with Pho80/Pho85 and Pho4 and compared phosphorylation with recombinant Pho4 purified from E. coli. We found that GST-Pho4 is successfully phosphorylated by Pho80/Pho85 (data not shown). Though it is difficult to quantify the amount of GST-Pho4, at appears to be phosphorylated somewhat less efficiently than recombinant Pho4.

To address the last possibility, we probed for Pho4 in the two pools (column 60, row 5) which should contain it by immunoblot with a Pho4 antibody. We were not able to detect GST-Pho4 in the row 5 pool, and were able to detect only a very faint band in the column 60 pool (data not shown). When compared to individually purified GST-Pho4 diluted 1:64 (the theoretical yield in the pool), the amount of protein appears significantly less, at best 1/10 of the theoretical yield.

We conclude that the failure of the screen to identify Pho4 as a substrate of Pho80/Pho85 is a combination of two factors. A small contributor is the decreased efficiency with which GST-Pho4 is phosphorylated. The major cause, however, is the near absence of GST-Pho4 in the purified protein pools. The GST purification was not particularly efficient in our hands to begin with, and the amount of GST-Pho4 represented in column 60 is even below the theoretical yield based on this poor purification. This could be due to some asymmetry in the growth rates of the strains in these pools, or perhaps due to competition between different GST fusions in the purification step.

Conclusions

The failure of the screen to identify Pho4 as a substrate, as well as the apparent cause of this, the seeming lack of GST-Pho4 in the purified protein pools, raises some significant concerns about this approach. Though we were able to identify some putative substrates of both Pho80/Pho85 and Pcl1/Pho85, it is difficult to understand the significance of these candidates. We were never able to repeatably demonstrate physiological relevance.

We are also concerned about the concentrations of proteins used in these screens, especially the relative concentrations of the kinases and the potential substrates. We used higher than physiological concentrations of the kinases, and the protein content of the GST pools was very low, possibly far below the physiological concentrations of these proteins. This results in a situation in which spurious kinase activity is not only possible, but probable. We had hoped to skirt

20

around this problem by screening with both kinases and selecting substrates which exhibit some specificity for only one, but even in cases where some specificity exists, we cannot be sure that this represents physiological activity.

We are left at a loss as to how to interpret the candidate substrates identified in these screens. The primary source of this confusion is the lack of knowledge of what the GST collection, or the associated purified protein pools, actually represent with respect to the entire yeast proteome. The collection is not clonal, was never checked genetically, and proteins derived from it are not readily detected. It therefore makes it extremely difficult to assess the coverage of the proteome.

This also raises a related and significant point about proteomic approaches in general, that it is imperative to understand the quality and completeness of whatever library, collection, or approach is used, as well as the accuracy of the assay used. Clearly, the GST collection would have benefited from some quality control at the genetic level; if individual isolates were chosen and assayed for fusion of the ORF to GST we would have a better idea of the coverage of the proteome which this collection represents. However, ultimately a genetic test like this is somewhat meaningless; the only truly relevant metric for a collection such as this one is how well the proteins themselves are represented in whatever assay is used. Without some knowledge of this metric, interpretation of the resulting data, especially in relation to the proteome in general, is near impossible.

Literature cited

1. Martzen, M.R. et al. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153-1155 (1999).

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

Figure 1: Screening the GST pools with Pho80/Pho85 and Pcl1/Pho85. We screened all the GST row (R) and column (C) pools with Pho80/Pho85 and Pcl1/Pho85 in an *in vitro* kinase assay, and analyzed results by SDS-PAGE and phosphorimaging. Results are shown, with squares denoting Pcl1/Pho85 and circles denoting Pho80/Pho85.
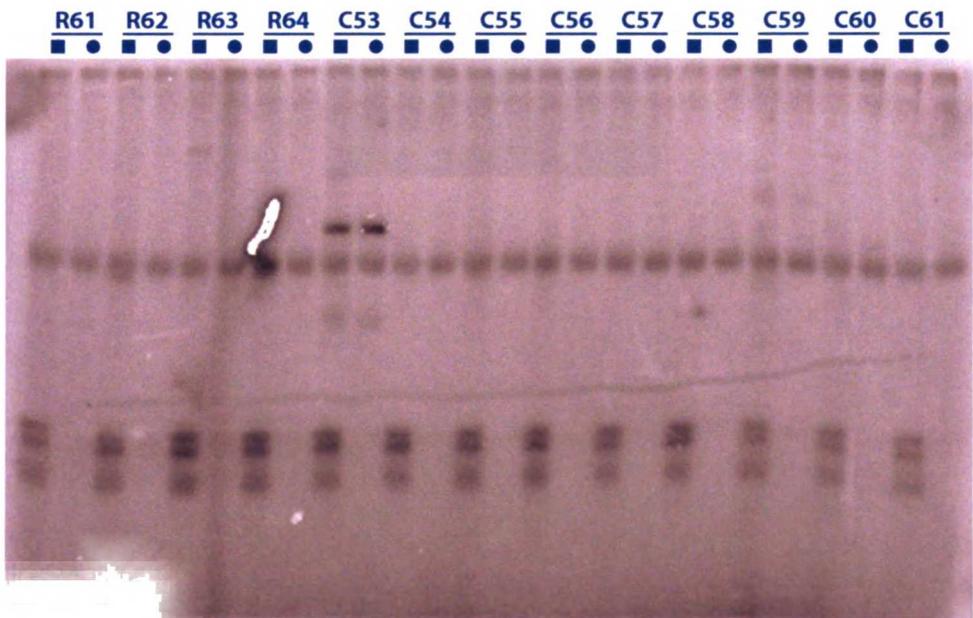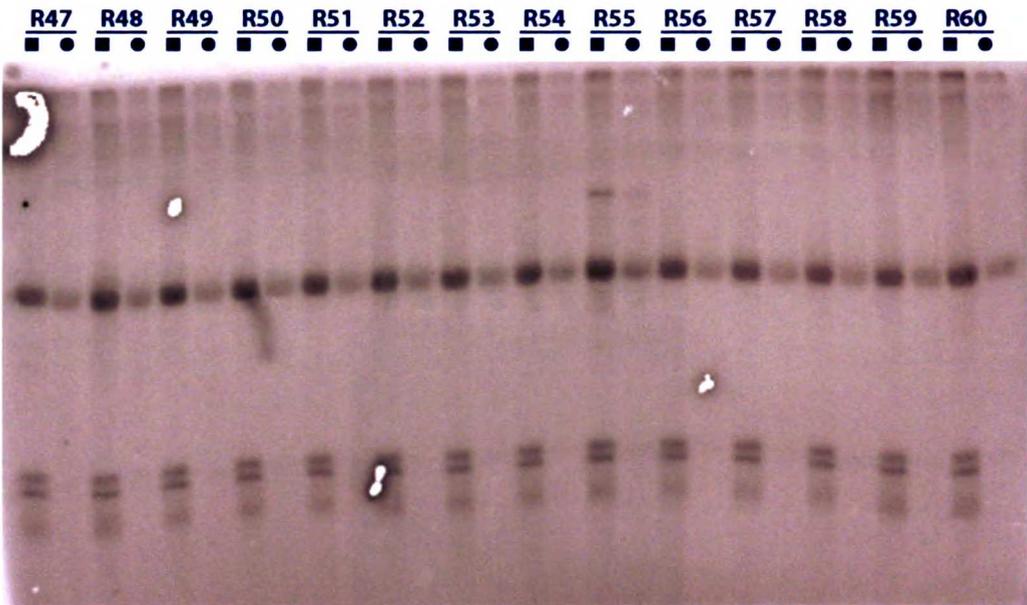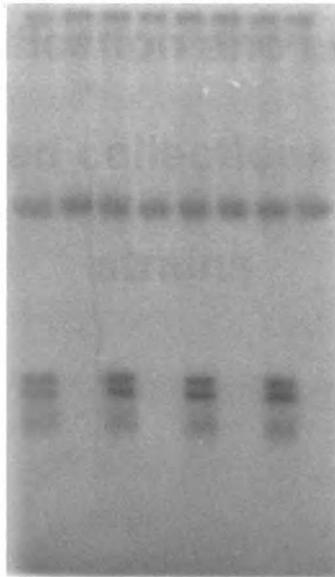
24

25

29
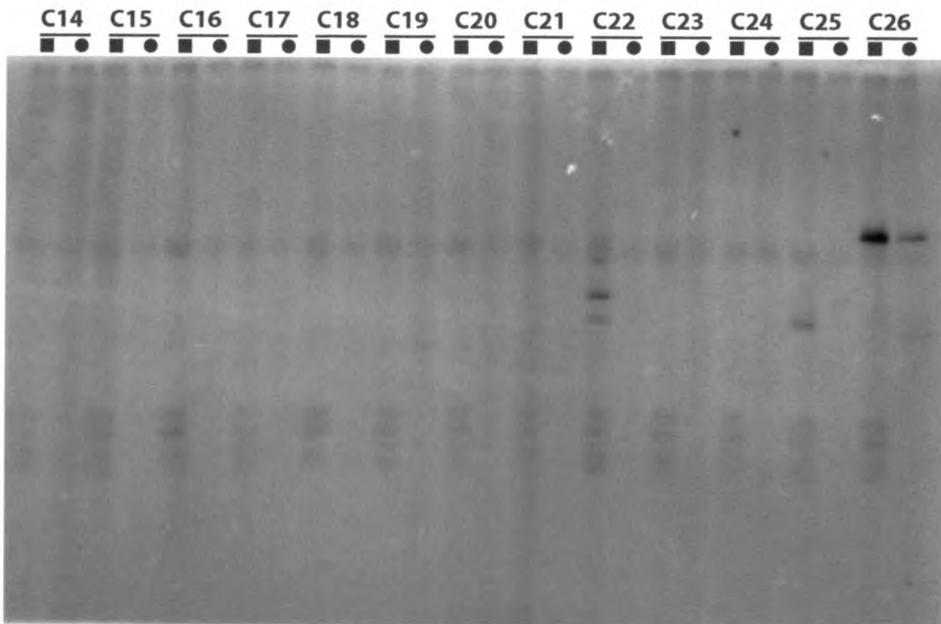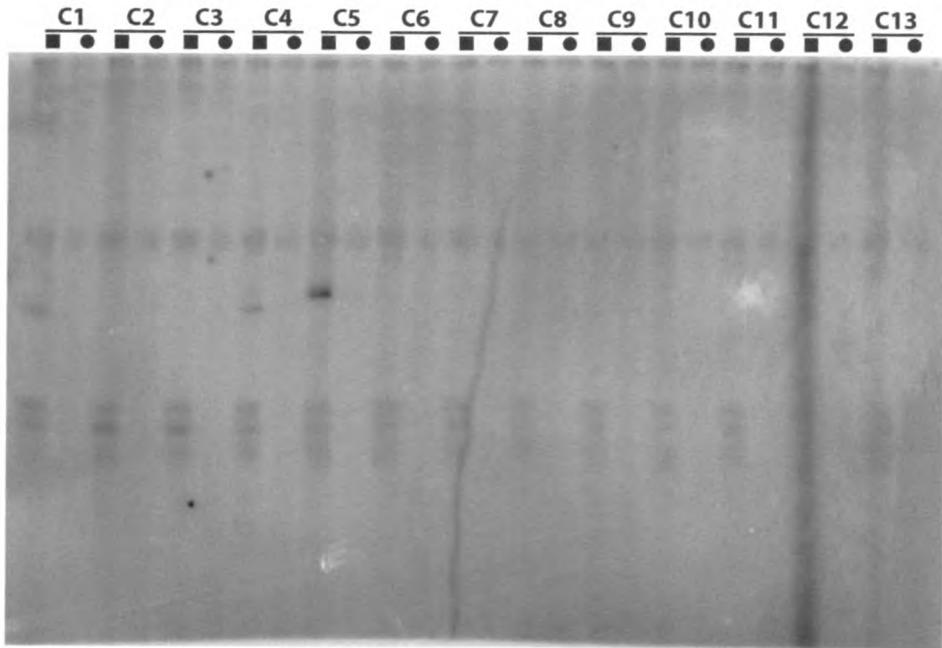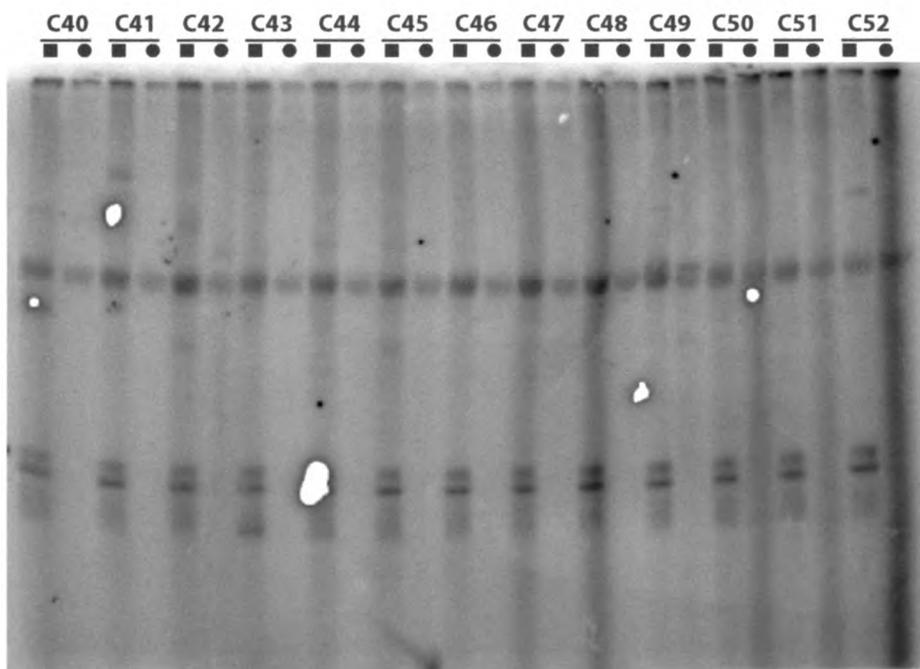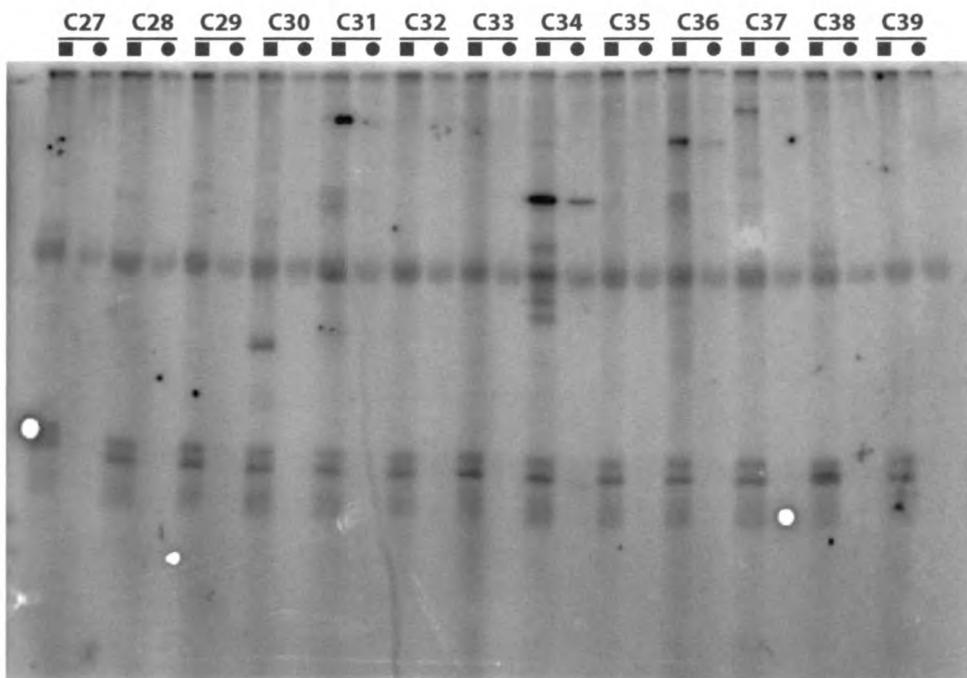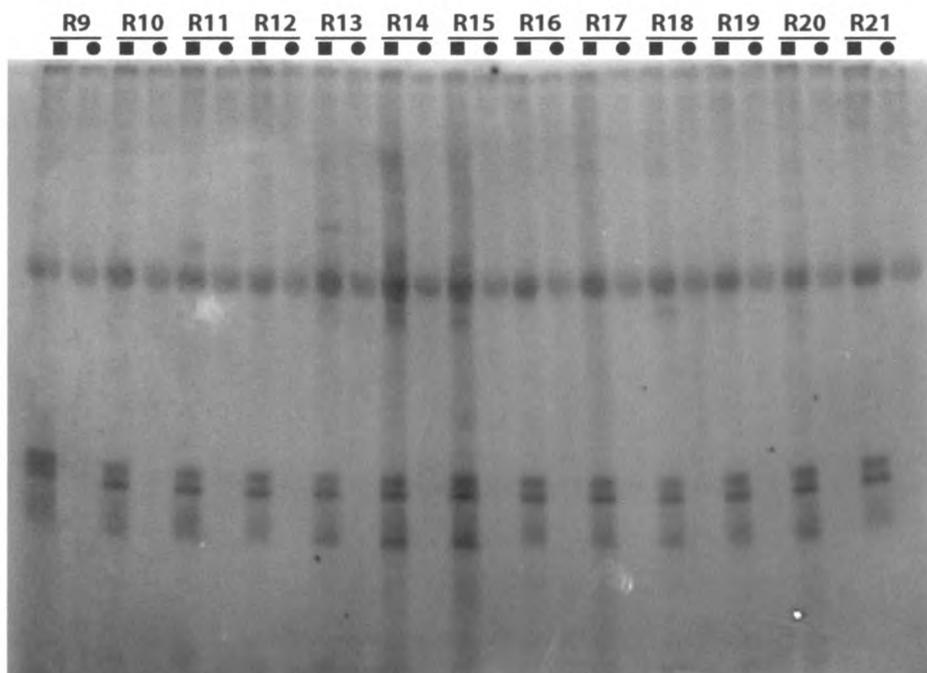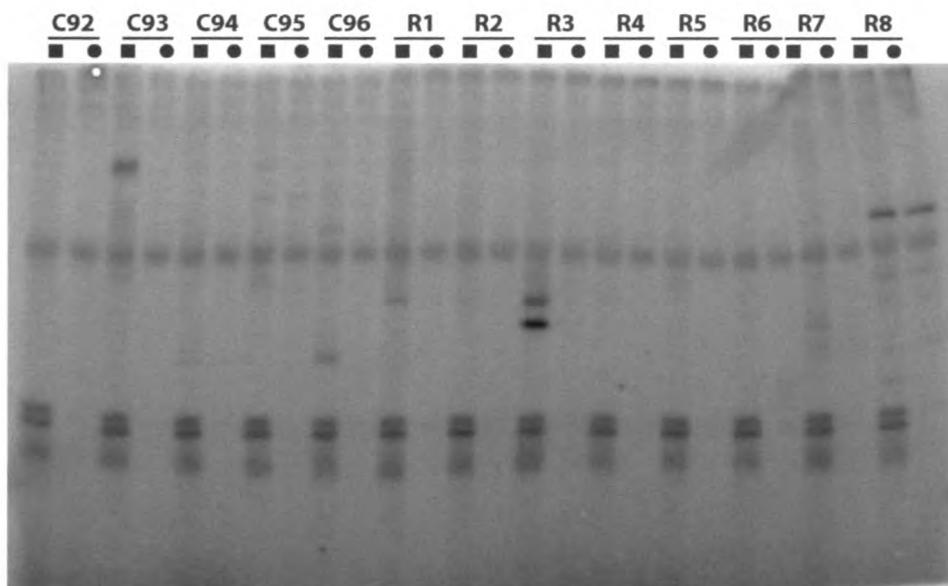
Chapter 3:

Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains

Abstract

A major challenge in the post-genomic era is the development of experimental approaches to monitor the properties of proteins on a proteome-wide level. In particular, it would be particularly desirable to assay protein subcellular localization, post-translational modification, and protein-protein interactions, both at steady-state and in response to environmental stimuli. Here we describe the construction of two collections of budding yeast strains which facilitate proteome-wide measurements of protein properties. These collections consist of strains with an epitope tag integrated at the C-terminus of essentially every ORF, one with the tandem affinity purification (TAP) tag, and one with the green fluorescent protein (GFP) tag. We demonstrate that in both of these collections we have accurately tagged a high proportion of all ORFs. Further, we demonstrate the use of the TAP collection in performing high throughput immunoprecipitation experiments.

Introduction

The complete sequencing of the *Saccharomyces cerevisiae* genome in 1996(*1*) enabled a new era of global biological analysis of this organism. Sequence analysis of the genome provides a wealth of information relevant to many aspects yeast biology, most recently in comparison to the genomes of other yeast species(*2-4*). The subsequent development of whole genome transcriptional profiling using DNA microarrays pushed this global analysis into the experimental realm by allowing for assessment of the global transcriptional profile under various conditions(*5-7*). This approach has been extremely effective in understanding many biological processes.

There are, however, limitations to microarray analysis. Though the transcriptional profile of an organism is informative, many biological processes do not produce readily-interpretable transcriptional readouts. Additionally, the proteins effectors of biological processes in the cell cannot be directly monitored through the transcriptional profile. It would be particularly useful to be able to globally monitor post-translational modifications, localization, and protein-protein interactions in order to understand how the dynamic properties of proteins allow them to carry out meaningful biological processes.

Part of what makes microarray analysis possible is the chemical similarity and stability of nucleic acids. Despite the fact that these ORFs encode proteins of diverse composition, structure and function, the associated nucleic acids have

33

virtually identical chemical properties. As a result, chemical manipulations can be done for all ORFs in parallel, simply by separating the nucleic acids spatially.

The same type of global analysis has been extremely challenging to achieve for proteins. The diversity of structure, chemical composition, and stability makes generalized manipulation difficult. Further, the behavior of these proteins in isolation is frequently not identical to their activity in the context of the cellular milieu. Mass spectrometry and two-dimensional electrophoresis have been used with some success in global protein analysis(8), but have been hampered by the complexity of the proteome, and thus far have not been capable of fine resolution analyses of a global nature. In addition, these techniques have a somewhat limited scope in the types of protein characteristics they are capable of measuring. Some groups have created purified protein libraries(9) or protein microarrays(10-12), but these approaches have suffered both from the significant effort involved in purifying the proteins and the inherent caveats of working with proteins in vitro. A further problem which has plagued all of these approaches has been difficulty in accurately assessing true coverage of the proteome.

It would therefore be extremely useful to have a system that enabled systematic high-throughput analysis of the yeast proteome, both in vivo and in vitro. One could construct such a system by fusing a constant epitope tag to all proteins, in essence making these proteins more chemically similar to each

other. Such similarity would enable systematic manipulation or analysis of these proteins by a single method.

Here we describe the design and synthesis of a set of oligonucleotide primers useful for the genomic integration of DNA coding for any C-terminal tag to every ORF in the yeast genome. Further, we describe the construction of two collections of yeast strains, one with the tandem affinity purification (TAP) tag and one with green fluorescent protein (GFP), and the high throughput technology and methodology which enabled the construction of these collections by a small team in a relatively short period of time. Lastly, we discuss the utility of these collections in the systematic execution of high throughput biochemical and microscopic assays on the yeast proteome.

Methods

Use of robotics and other high-throughput tools

In designing the oligonucleotides and the collections, as well as in developing the methodology for creation and use of these collections, we made every effort to utilize available high throughput technologies. The collection was designed in 96-well format to enable efficient oligonucleotide synthesis using a 96-well DNA synthesizer (GeneMachines Polyplex) and subsequent liquid handling by a robotic pipettor (Beckman Biomek FX). The Biomek FX was used for almost all liquid handling applications, including resuspending oligos, setting up all PCRs, loading agarose gels, transformations, inoculating cultures and

adding lysis buffer to cell pellets. In cases where the Biomek FX could not be used, we instead employed electronic multichannel pipettors, for applications such as loading SDS-PAGE gels and dispensing buffers. In minimizing the number of manual steps performed, we were able to optimize the efficiency of construction and use of the collections, enabling the entire process to be completed in-house by a relatively small team.

Oligonucleotide primer design and synthesis

The yeast genome sequence, as well as the coordinates of all ORFs, were obtained by download from the Saccharomyces Genome Database (*13*) on April 17, 2001. We removed all mitochondrial genes, as well as those encoding Ty elements. We then divided the ORFs into two categories: soluble and putative membrane proteins, reasoning that having membrane proteins as a separate group would facilitate any modifications in biochemical assays needed for these proteins. The following criteria were used to designate ORFs into the putative membrane category: 1) Any protein experimentally determined to be an integral membrane or membrane-associated protein; 2) Any ORF with homology to such an ORF; and 3) Any ORF with ≥2 putative transmembrane domains which appeared in the microarray data of polysome associated RNAs(*14*).

Within the soluble and membrane categories, we ordered genes by size, with largest genes first, and then divided ORFs into groups of 96 to facilitate subsequent manipulations in 96- and 384-well plates. Each ORF is designated

by a plate number and coordinates within that plate. All reagents and strains relevant to a given ORF occupy the same unique coordinates. This facilitated subsequent manipulations by the Biomek FX.

We used the "Promoter" program (courtesy of Joe DeRisi, publicly available at http://derisilab.ucsf.edu) to extract the last 40 nucleotides (excluding the stop codon) of each ORF, as well as 40 nucleotides of genomic sequence immediately following the stop codon of each ORF. We added the constant forward sequence from the Pringle system to the last 40 nucleotides of each ORF to create the F2 oligo sequence, and the reverse complement of the 40 nucleotides following each ORF to the constant reverse sequence to create the R1 oligo sequence. To design the sequence of the unique check primer for each ORF, we utilized Primer 3.0(15), selecting oligonucleotide primers with melting temperatures of 60°C and which hybridize between 400 and 650 nucleotides upstream of the stop codon for each ORF.

Oligonucleotides were synthesized on a GeneMachines Polyplex 96-well oligonucleotide synthesizer. This machine was modified to accommodate larger reagent bottles required for 60-mer synthesis. All reagents used were from Glen Research. We used a protocol optimized for producing full length 60-mers without a need for changing reagent bottles (involving more and longer coupling steps with less volume), enabling us to run the machine overnight, which allowed for the efficient synthesis of the 12,468 60mer and 6234 20 mer oligonucleotides required.

Synthesized oligonucleotides were cleaved off the solid synthesis support

by incubating 3 times with 100 µl of NH₄OH for ten minutes, followed by

collection into a deep well 96-well plate with a vacuum manifold (Millipore). The

oligos were then baked at 55°C for 15-24 hours, and lyophilized in a SpeedVac

AES2010 to remove the NH₄OH. Prior to use, oligonucleotides were

resuspended to a concentration of 100 µM with deionized water (typically 200-

300 µl).

Construction of collections

We performed PCR and transformation in 96-well format as follows: F2

and R1 oligos were combined to working concentrations of 5 µM each. 10 µl of

the primer combination were added to 40 µl of a PCR mix [19.5 µl H₂O, 5 µl 10X

Pwo buffer, 5 µl 20 mM MgCl₂, 5 µl 2mM dNTPs, 5 µl plasmid template DNA (~5

µg/ml), 0.5µl Expand DNA polymerase] aliquotted to 96-well plates in order to

amplify the desired tag. For PCR, we used an MJ Research Tetrad thermal

cycler with the following program [94°C 1:00, 10 x (94°C 0:15, 55°C 0:30, 72°C

2:00), 15 x (94°C 0:15, 55°C 0:30, 72°C 2:00 + 5s/cycle), 72°C 10:00]. Following

PCR, we checked for a product of correct size using 96-well agarose gels

(Amersham Ready to Run system) loaded with the Biomek FX.

These PCR products were then transformed into our base strain [ATCC

#201388: S288C, MAT a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0] . We used 15 µl of

unpurified PCR product in the following transformation recipe: 100 µl 50%PEG,

15 µl 1M Lithium acetate, 20 µl salmon sperm carrier DNA (2 mg/ml), 17µl

DMSO, and 30 µl of a suspension of lithium acetate washed log phase yeast

cells. We then performed incubations [30°C for 30 minutes followed by 42°C for

15 minutes] for transformation in the thermal cycler, and plated transformations

individually on standard yeast synthetic media plates lacking histidine (SD–HIS)

to select for genomic integrants.

After growth for three days, transformations typically yielded 5 to 100

colonies. We selected up to six individual transformants for each ORF and

streaked onto fresh selective media. After subsequent growth, we performed

whole cell PCR on each transformant to determine if the tag had integrated at the

correct locus. A small aliquot of freshly grown cells was resuspended in 5 µl of

water and boiled in 96-well format (99°C for 5 minutes in the thermal cycler). 5 µl

of boiled cells, as well as 2.5 µl of 5 µM unique "check" oligos were added to

PCR mix [13 µl H2O, 2.5 µl 10X Taq buffer, 1.5 µl 2 mM dNTPs, 0.25 µl 50 µM

"F2CHK" primer, 0.25 µl 5U/µl Taq Polymerase, 0.05 µl 10 mg/ml RNAse], and

PCR was performed [94°C 2:30, 35 x (94°C 0:45, 55°C 0:45, 72°C 1:00), 72°C

10:00]. We analyzed the results of these PCRs by 96-well agarose gel

electrophoresis, identifying correct integrants by the presence of a PCR product

of appropriate size.

For construction of the GFP collection, we used much the same method,

except individual transformants were picked and used to directly inoculate 600 µl

liquid cultures for overnight growth (SD-His medium). We centrifuged 200 µl of

these cultures in 96-well PCR plates to pellet cells, removed the supernatant, and lysed the cells in 20 μl of 0.2% SDS at 99°C for 10 minutes in the PCR machine. We then used 0.6 μl of this lysate as template for a 20 μl PCR [16.2 μl H2O, 2 μl 10X Tag buffer, 0.6 μl 2mM dNTPs, 0.2 μl each 50 μM oligonucleotide primer, 0.2 μl 5U/μl Taq polymerase, 0.04 μl 10 mg/ml RNAse] to confirm correct integration of the tag. Presence of a PCR product was again analyzed by 96-well agarose gel electrophoresis.

Assembly and growth

To assemble these collections, we selected two correct integrants (when possible) for each ORF, resuspending cells in 0.5X SD-His 15% glycerol. Plate number and coordinates were maintained for each ORF, resulting in an "A" and "B" collection for each set of epitope-tagged strains, which we froze at –80°C. For the GFP strains, we assembled the A and B collections directly from the liquid cultures used for confirmation PCR, mixing saturated cultures with 30% glycerol and freezing.

Subsequent growth was achieved by thawing the glycerol stocks and either inoculating liquid cultures with a Biomek FX robot or spotting onto YEPD plates with a 96-well pinning tool. To grow these cultures in high-throughput format, we used a GeneMachines HiGro growth chamber. Typically, cultures were 1.5-2.0 ml, and cells were grown at 30°C and 500 RPM. Under these conditions, cells grew with the same growth rate as liquid cultures in standard

flasks (data not shown). We sometimes used teflon-coated magnetic beads in each well to enhance mixing of the culture. Addition of these beads did not have an effect on growth rate (data not shown), but did prevent settling of cells which began to occur in late log phase.

Immunoblot analysis of the TAP collection

To analyze the TAP collection by immunoblot, 200 μl YEPD cultures were inoculated from a plate with a 96-well pinning tool and allowed to grow to saturation overnight. We diluted these cultures into 1.8 mL of YEPD in deep-well 96 well plates, and allowed cells to grow to logarithmic phase. Log phase cultures were analyzed by OD_{600} and centrifuged to pellet the yeast cells. We removed the media supernatant and added 50 μl hot SDS lysis buffer [50 mM Tris pH 7.5, 5% SDS, 5% glycerol, 50 mM DTT, 5 mM EDTA, Bromophenol blue, 2 μg/ml leupeptin, 2 μg/ml pepstatin A, 1 μg/ml chymostatin, 0.15 mg/ml benzamidine, 0.1 mg/ml pefabloc, 8.8 μg/ml aprotinin, 3 μg/ml anipatin], and boiled (99°C for 10 minutes in the thermal cycler). These lysates were centrifuged, and the supernatant was kept and frozen at –80°C.

We loaded 13 μl of these lysates on 26-well 4-15% gradient precast Criterion gels (Bio-Rad) with a multi-channel pipettor [Matrix technologies Impact2]. Gels were run and transferred to PVDF membranes with a Trans-blot SD semi-dry blotter (Bio-Rad). Immunoblot analysis was performed using a primary antibody mixture of an affinity purified antibody to the CBP portion of the

41

TAP tag (1:5000 dilution) and an anti-hexokinase antibody (US Biological, 1:50000) as a loading control and quantitation standard. A horseradish peroxidase (HRP) conjugated goat anti-rabbit was used as a secondary antibody, and the SuperSignal West Femto Maximum Sensitivity ECL substrate (Pierce) was used for detection. Images were collected with a CCD-based imaging system (Alpha Innotech), and analyzed with the FluoroChem FC software (Alpha Innotech).

Fluorescence microscopy on the GFP collection

We grew cells from the GFP collection to log phase in the same way, except in standard synthetic yeast liquid medium lacking histidine. Aliquots of these cultures were analyzed in 96-well glass bottom microscope slides (BD Falcon) pre-treated with concanavalin A (50 µg/ml) to ensure cell adhesion. We imaged cells using a Nikon TE200/300 inverted microscope with an oil-immersed 100X objective, and made use of scripting functions in the MetaMorph version 4.6r8 imaging software in order to automate most of this process. GFP and DAPI fluorescence images, as well as DIC images, were collected for each strain and analyzed for expression and subcellular localization.

Reorganization of TAP collection

With the TAP immunoblot data in hand, we reorganized the TAP collection strains according to abundance. Six categories of expression were created, and

within each of these categories, strains were organized by the length of the protein. This enabled more precise analysis of protein abundance, as well as multiplexed assays.

96-well growth and native extract preparation

To grow cells for extract preparation, we first inoculated overnight 600μl YEPD cultures from a YEPD plate using a 96-well pinning tool. These cultures were allowed to grow to saturation overnight on the benchtop with no agitation. The next morning we diluted these cultures into 6 deep well 96-well plates, with 1.8 ml YEPD media in each well, to an OD_{600} of between 0.1 and 0.2, and grew to log phase ($0.8 < OD_{600} < 1.0$) at 30°C and 500 rpm in a GeneMachines HiGro Shaker. When cultures had reached log phase, we centrifuged the plates at 3000 rpm for 10 minutes and aspirated the media. The cell pellets were resuspended in 150 μl cold sorbitol buffer (1.2 M sorbitol, 0.1M KPO_4 pH 7) + 2μl/ml 2-mercaptoethanol. We then combined the cell suspensions from 6 96-well plates, maintaining the coordinates, and centrifuged and aspirated again. These pellets were then resuspended in 150 μl cold sorbitol buffer with 2μl/ml 2-mercaptoethanol and 60μl/ml lyticase(16), and transferred to a 96-well PCR plate. We incubated these plates at 30°C for 15 minutes in the thermal cycler, then centrifuged gently (1000x g) for 10 minutes. After aspirating the supernatant, the pellets were washed gently in sorbitol buffer, frozen in liquid nitrogen and stored at –80°C.

To lyse the cells, we resuspended the thawed cell pellets in 100 µl hypotonic lysis buffer (50 mM Tris pH 7.5, 5 mM $MgCl_2$, 5mM EGTA, 1 mM EDTA, 0.1% Triton X-100, 1 mM 2-mercaptoethanol, 2 mM PMSF, 2.5 mM benzamidine, 1 µg/ml leupeptin, 1 µg/ml pepstatin). After incubation on ice for 10 minutes, we added 25 µl of buffer plus 0.9M NaCl, and incubated on ice for another 10 minutes. We then centrifuged at 4000 rpm in a Beckman RC-3B swinging bucket rotor fitted with 96-well plate carriers for 20 minutes to pellet cellular debris. Following centrifugation, we removed 100 µl of lysate to a new 96-well plate containing 25 µl 50% glycerol in each well, and kept a small aliquot to measure protein concentration by Bradford assay. These extracts were then frozen in liquid nitrogen and stored at –80°C. Each well contained 125 µl of total extract and typically were ~10 mg/ml.

Multiplexed immunoprecipitations

To perform high throughput immunoprecipitations, we first prepared native extracts as above, except we combined cell pellets from 6 *different* cultures into one plate. The result is that each well contains a total of 125 µl of ~10 mg/ml extract, but derived from a mixture of 6 different TAP tagged proteins. We combined strains with approximately equal expression levels of the fusion protein in order to minimize dominance of well expressed proteins or loss of minimally expressed proteins.

To perform the immunoprecipitation reactions, we first diluted extracts in 1 ml deep well 96-well plates to a total volume of 560 µl with P buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM MgCl$_2$, 5mM EGTA, 1 mM EDTA, 0.1% Triton X-100, 1 mM β-mercaptoethanol, 2 mM PMSF, 2.5 mM benzamidine, 1 µg/ml leupeptin, 1 µg/ml pepstatin). We then added 3 µg of biotin conjugated Human IgG (Jackson Immunoresearch), and incubated extracts for 30 minutes at 4°C. 40 µl of a 25% suspension of streptavidin beads (Amersham Pharmacia) was then added to each well. We incubated the plate at 4°C for another 30 minutes, vortexing *very gently* 3-4 times during this period to resuspend the beads.

We then transferred the reactions with a multi-channel pipettor and filter tips to a filter plate (Orochem catalog #OF1100). This type of plate contains 96 wells, each with a small frit above an opening at the bottom of each well. We placed this plate on a vacuum manifold; applying a vacuum allows for the removal of the supernatant but retention of the beads. Beads were washed four times with 400 µl of PDMS buffer [P buffer + 1% Triton X-100 + 300 mM NaCl]. We then centrifuged the plate briefly (1000x g for 1 minute) to remove any residual liquid remaining in or around each well. 10 µl of sample buffer was added to each well, and the plate was vortexed to resuspend the beads in the sample buffer. The plate was allowed to stand at room temperature for 5 minutes, and then centrifuged (1000x g for 2 minutes) on top of a shallow 96-well plate to collect the eluate. This process was repeated again, to generate a total elution volume of 20 µl.

To analyze the results of the multiplex immunoprecipitation, the samples were run on 26-well Criterion SDS-PAGE gels (Bio-Rad), and transferred to nitrocellulose in 20 mM NaPO$_4$ pH 6.8 buffer with a BioRad Transblot apparatus. We performed immunoblot analysis by probing with Rabbit Fc (Jackson Immunoresearch, 1:10 000 dilution of 3.8 mg/ml stock in TBST + 5% nonfat dry milk), an HRP-conjugated goat anti-rabbit Fc secondary antibody (Jackson Immunoresearch, 1:50 000 dilution of a 1 mg/ml stock in TBST + 5% milk), and the SuperSignal West Femto Maximum Sensitivity ECL substrate (Pierce). Images were collected with a CCD camera (Alpha Innotech).

For further details on any of these protocols, please refer to http://www.ucsf.edu/ekolab.

Results

Collection design and oligonucleotide synthesis

To create a collection of tagged yeast strains, we utilized oligonucleotide directed homologous recombination (Figure 1)*(17)*. Briefly, 60-mer oligonucleotide primers (the "F2" and "R1" primers) are designed which contain both a variable sequence, homologous to the gene of interest, and a constant sequence, which enables PCR amplification of sequence coding for the desired tag and a nutritional marker used to select for integrants. Homologous sequences in the oligonucleotides direct integration of the tag at the desired

46

location in the genome. Integration at the correct locus is then confirmed using PCR with a primer in the tag and one specific to the targeted ORF (Figure 2). We chose to tag the C-terminus of every ORF so that the endogenous promoter would remain intact, and to minimize the impact on the signal sequence of secreted and membrane proteins. We therefore designed and synthesized the required oligonucleotides for every ORF in the yeast genome. Though we chose to construct a collection with the TAP tag and one with the GFP tag, this oligo set could be used to construct a collection of strains with any desired tag integrated at the C-terminus of every ORF.

An important consideration in the design of the collections was the ability to take advantage of high-throughput and automation technologies. We designed the oligos in 96-well format to be efficiently synthesized in a high-throughput synthesizer, and developed methods for construction and use of the collections based around liquid handling robots and multi-channel pipettors. The use of these tools was imperative in making a project of this nature feasible. For the construction of the GFP collection, the only manual steps performed were plating transformations and picking colonies.

TAP and GFP collection construction

The first round of construction of the TAP collection was accomplished in about four months by a team of six people. We subsequently further refined the protocol for the confirmation PCR by growing individual transformants in liquid

47

culture rather than on solid medium, enabling all manipulations to be performed by the Biomek FX liquid handling robot. With this refinement, as well as others, the efficiency of the construction process was greatly enhanced: the first round of contruction of the GFP collection was carried out by two people in under three months. This oligo set and these methods could therefore be extremely useful in the construction of new collections tailored to particular experiments.

After the first round of construction of the TAP collection, we successfully tagged 97% of 6234 ORFS, as assayed by genomic PCR. For the GFP collection, we obtained PCR-positive clones for 98.7% of ORFs. We also determined that the frequency of obtaining a properly integrated strain was roughly equivalent between essential (93%) and non-essential (98%) ORFs.

Confirmation of correctly expressed fusion proteins and reconstruction

In beginning to work with these strains, we discovered some inconsistencies in expression of tagged proteins in different isolates from the same ORF. Specifically, we identified some cases in which only one of the two isolates expressed the protein of interest, despite the fact that correct integration was confirmed for both by genomic PCR. After sequence analysis of some representative isolates, we identified sequence errors in the junction between the C-terminus of the protein and the tag, or in the tag itself. The source of these errors is presumably errors in the oligonucleotides themselves or the tag amplification and transformation process. Regardless, we concluded that proper

tagging could only be reliably confirmed by analyzing expression of the tagged

protein (by immunoblot for the TAP collection, and microscopy for the GFP

collection).

We analyzed log phase cultures from the "A" and "B" isolates of both

collections for expression of the tagged proteins. For the TAP collection, we

made SDS extracts and analyzed these extracts by SDS-PAGE followed by

immunoblotting against the TAP tag. For the GFP collection, we observed the

tagged proteins by fluorescence microscopy. A more thorough discussion of

these methods is in the methods section.

After this analysis, we were able to identify which isolates of a particular

ORF were correctly expressing the tagged protein of interest. However, in cases

where expression could not be confirmed for either isolate, we were unable to

distinguish between two possibilities. First, it is possible that neither isolate

contained the correctly tagged ORF, or the protein could be correctly tagged, but

not expressed to detectable levels under our growth conditions. We therefore

compared our results from both collections to determine if there were cases in

which expression of a given ORF was detected in one collection but not the

other. Given that these collections were constructed with the same oligo set, one

would expect to obtain similar results between the two collections. Though these

means of analysis are different and therefore may have different limits of

detection, we reasoned that if a tagged protein was detected in one collection but

not the other, it was probably not tagged correctly in the collection in which it was not detected.

With this information in hand, we undertook the task of reconstructing those strains (708 for the TAP collection, 759 for GFP) for which we were unable to identify a correctly expressing isolate. In order to avoid again isolating strains which did not correctly express the tagged protein of interest, we omitted the confirmation PCR and instead analyzed individual transformants by immunoblotting or microscopy. After this process, 457 new positives were obtained for the TAP collection, and 398 new positivies were obtained for the GFP collection.

Coverage of the proteome

In the construction of these collections, we monitored our success rate at many steps in order to assess the quality and utility of these collections in performing proteome-wide studies (Table 1). While we were able to obtain an extremely high success rate by genetic analysis (96.9% for TAP, 96.7% for GFP), our subsequent expression analysis indicated that this is not the most accurate metric for coverage of the proteome. Because of the potential for errors in the integration process, and the inability of genetic analysis to uncover these errors, we feel that detection of expression of the fusion proteins is the most accurate metric of coverage, expecially considering that utility of these collections in monitoring protein characteristics ultimately rests on the ability to detect them.

By this metric, we detect 4245 proteins, or 68.1% of all 6234 ORFs, in the TAP collection, and 4156, or 66.7% of all ORFs, in the GFP collection. Given that many of the 6234 ORFs are spurious ORFs, or may not be expressed under our growth conditions, we feel that the success rates of these collections represent a very high proportion of all expressed ORFs. The overlap between the proteins detected in these collections (over 90% of ORFs detected in the GFP collection were also detected in the TAP collection) strengthens this supposition. Lastly, these numbers will undoubtedly improve as more sensitive detection methods are developed (such as immunoprecipitation followed by immunoblot), as expression in other growth conditions is examined, and as more thorough computational analysis of the genome eliminates some ORFs as spurious.

If we take our success rate for essential proteins as a proxy of our overall success rate (since essential proteins are presumably true ORFs and expressed under normal growth conditions), Our collections represent 74.6% of the proteome for the TAP collection, and 75.2% of the proteome for the GFP collection. When looking at both collections, at least one fusion protein was detected for about 80% of essential ORFs. The high percentage of essential ORFs detected also indicates that the fusion protein is indeed functional in a high proportion of cases. We conclude that these collections represent truly useful representations of the proteome for use in global analyses. In characterizing these collections, we were also able to make quantitative measurements of

protein expression*(18)* and describe cellular localization*(19)* for the majority of the yeast proteome.

"Multiplexed" immunoprecipitations

In constructing these collections, we wished to not only be able to do descriptive analyses, but also to perform experiments systematically on the entire proteome. Our hope was that standard laboratory assays typically performed on a small number of strains or proteins could be applied systematically and efficiently to the entire proteome. The fact that every strain in the TAP collection utilizes the same tag enables a generalized method to be applied to all strains to perform large scale experiments in parallel.

As a first step, we developed high-throughput methods to efficiently make extracts and immunoprecipitate proteins (Figure 3). Briefly, this involves growing 2 ml cultures to log phase in 96-well format, combining cell pellets from 6 different 96-well plates, and spheroplasting cells with lyticase. We made extracts by osmotic lysis, pelleting cellular debris and keeping the supernatant. These "multiplex" extracts contain extract from 6 different strains, and therefore 6 different TAP-tagged proteins, facilitating the parallel immunoprecipitation of 6 different proteins in each well of a 96 well plate. Immunoprecipitation in 96-well format, therefore, theoretically enables the simultaneous pulldown of 576 proteins.

Importantly, because of the reorganization of the TAP collection according to abundance and size, the six TAP-tagged proteins in a given well are all of approximately equal abundance ensuring that extremely abundant proteins will not out-compete proteins of lesser abundance for binding to beads. Further, the 6 proteins in a given well represent the maximal possible size distribution within the abundance category, facilitating the resolution of the individual proteins in subsequent analysis by SDS-PAGE.

To test the feasibility of high-throughput immunoprecipitations, we prepared multiplex extracts and pulled down TAP tagged proteins as described in the methods. The last column of our plates are left empty, in order to provide space for controls necessary for any subsequent assay, so the theoretical maximum number of proteins pulled down in this reconstruction is 528 proteins. As shown in Figure 3, we were successfully able to pull down a large number of proteins with this procedure.

To quantify the efficiency of the multiplex pulldown, we counted the number of distinct individual bands in each lane, and summed the total for this gel and the entire plate. We cannot discount the possibility that some of these bands may represent breakdown products or modified proteins; nevertheless, due to the distinctness of the bands, we believe that a high proportion represent full length proteins. The gel shown contains 22 lanes, or one quarter of the plate, so the theoretical maximum number of immunoprecipitated proteins is 122. We are able to detect 111 distinct bands in this gel, or 84.1% of possible proteins. In

53

immunoblots for the entire plate, we were able to detect 452 distinct bands, representing 85.6% of the 528 proteins possible. In other studies we have observed that this efficiency typically ranges from 70-90%(data not shown).

Discussion

In this paper, we have described the construction of two collections of yeast strains, both with C-terminal fusions of almost every ORF in the yeast genome, one with the TAP tag and one with the GFP tag. The oligonucleotide primer set and the methods discussed could be used to efficiently construct a new collection, with any desired C-terminal tag and in any desired genetic background. With our final refinements of these methods, a new library could be constructed by a small team in a matter of months. Furthermore, many of the methods described could be easily modified for use with other model organisms which support efficient homologous recombination, such as *Schizosaccharomyces pombe.*

We have confirmed that these collections do indeed represent a significant portion of the proteome, as we have confirmed expression of the ORF-tag fusions in individual transformants, either by immunoblot analysis for the TAP collection or by microscopic analysis of the GFP collection. Importantly, all fusions are under control of the native promoter, and so are truly representative of the proper cellular context of the protein. We therefore believe that these collections will be useful tools in performing large scale proteomic experiments.

Indeed, in the course of confirming expression of the fusion proteins in these collections, we have been able to obtain valuable information about the absolute abundance of proteins (TAP collection) as well as their cellular localization (GFP collection).

We wished to take proteomic analysis beyond being descriptive, so we developed methods in order to apply standard biochemical experiments in a parallel manner. We found that multiplexed immunoprecipitation could be performed efficiently from extracts derived from yeast cultures as small as 2 ml. This procedure should be easily modifiable in order to perform almost any extract based assay in a parallel manner on the proteome.

These collections represent exciting possibilities for the future in a number of respects. First, it could be interesting to apply the descriptive methods outlined in this paper to experimental situations. One way in which cells can rapidly respond to environmental stimuli is to alter the localization, abundance, or post-translational modification of proteins, sometimes without any change in transcriptional state. One could use the methods described in this paper to monitor these aspects of many proteins under various growth conditions or in response to environmental insults, either with the entire collections or by examining a specific subset or family of proteins.

Second, it will be exciting to see what types of biochemical and/or microscopic assays are applied in this high-throughput parallel manner. The ability to perform immunoprecipitations in a "multiplex" format allows for the

efficient screen of the entire proteome with only a handful of experiments. Any number of post-translational modifications could be examined with minor modifications to this assay; it will be especially exciting to examine the dynamic nature of these modifications in response to environmental stimuli.

Further, the ability to systematically cross yeast strains through the use of the "magic marker"(20) would enable either of these collections to be efficiently crossed to any desired genetic background. This would enable the examination of how different mutants impact global localization, abundance, or perhaps post-translational modifications, or could be used to sensitize the strain background to various chemical or environmental stimuli. A tantalizing specific approach would be to cross the TAP collection with a strain or strains containing an ORF tagged with a different epitope, thereby enabling high-throughput IP western experiments to identify protein-protein interactions.

Worthy of mention is the ease and accessibility to high-throughput experimentation which these collections provide. Though we employed the use of robotics and other high-throughput equipment to construct these collections, it is possible to efficiently utilize these reagents with simply a handful of multi-channel pipettors and a 96-well pinning tool. We therefore hope that these collections will open the door to systematic proteomic analysis by all laboratories, including those with more limited resources.

Acknowledgements

Literature Cited

1. The Yeast Genome Directory. *Nature* **387**, 3-105 (1997).

2. Brachat, S. et al. Reinvestigation of the Saccharomyces cerevisiae genome annotation by comparison to the genome of a related fungus: Ashbya gossypii. *Genome Biol* **4**, R45 (2003).

3. Cliften, P. et al. Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* **301**, 71-76 (2003).

4. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E.S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241-254 (2003).

5. DeRisi, J.L., Iyer, V.R. & Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680-686 (1997).

6. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470 (1995).

7. Lockhart, D.J. et al. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* **14**, 1675-1680 (1996).

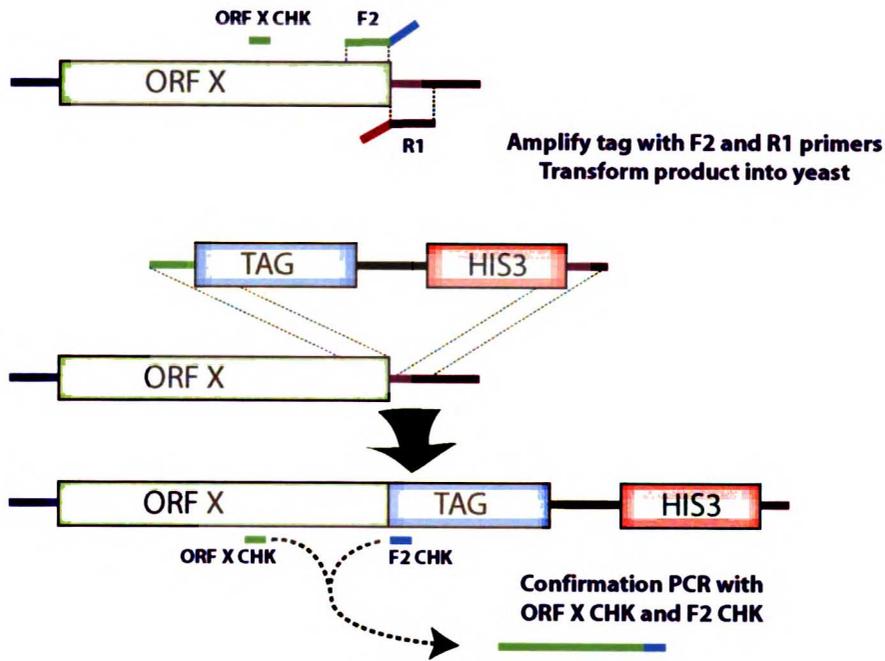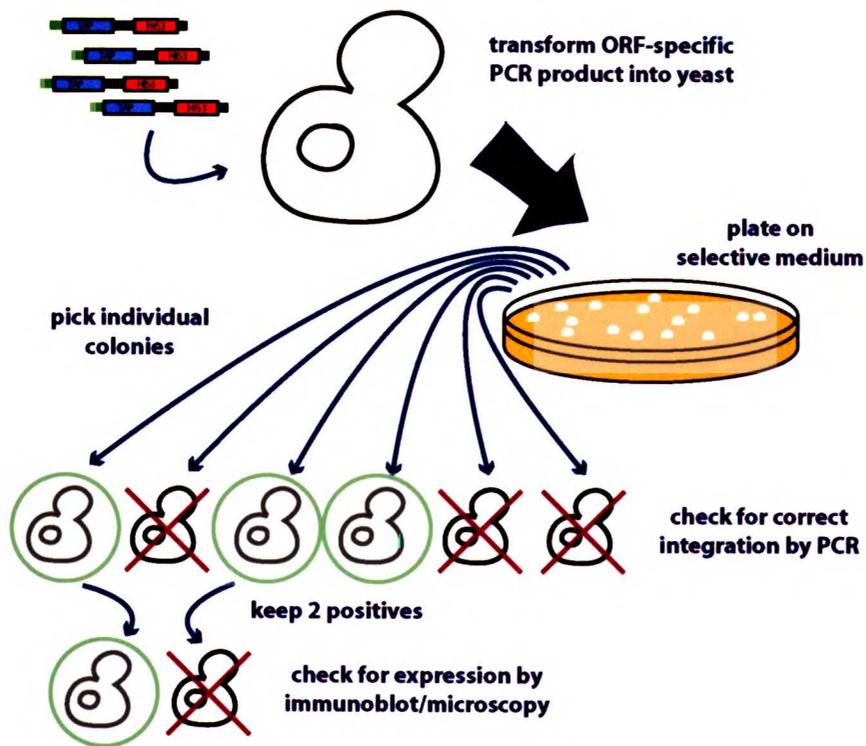8. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).

9. Martzen, M.R. et al. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153-1155 (1999).

10. Phizicky, E., Bastiaens, P.I., Zhu, H., Snyder, M. & Fields, S. Protein analysis on a proteomic scale. *Nature* **422**, 208-215 (2003).

11. MacBeath, G. & Schreiber, S.L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760-1763 (2000).

12. Zhu, H. et al. Global analysis of protein activities using proteome chips. *Science* **293**, 2101-2105 (2001).

13. Dolinski, K., Balakrishnan, R., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hirschman, J. E., Hong, E. L., Issel-Tarver, L., Sethuraman, A., Theesfeld, C. L., Binkley, G., Lane, C., Schroeder, M., Dong, S., Weng, S., Andrada, R., Botstein, D., and Cherry, J. M. Saccharomyces Genome Database. *ftp://genome-ftp.stanford.edu/pub/yeast/SacchDB*.

14. Diehn, M., Eisen, M.B., Botstein, D. & Brown, P.O. Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nat Genet* **25**, 58-62 (2000).

15. Rozen, S. & Skaletsky, H.J. Primer3. *Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html* (1998).

16. Haswell, E.S. & O'Shea, E.K. An in vitro system recapitulates chromatin remodeling at the PHO5 promoter. *Mol Cell Biol* **19**, 2817-2827 (1999).

17. Longtine, M.S. et al. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**, 953-961 (1998).

18. Ghaemmaghami, S. et al. Defining the Yeast Proteome. *Nature* **in press** (2003).

19. Huh, W.-K. et al. Global Analysis of Protein Localization in Budding Yeast. *Nature* **in press** (2003).

20. Tong, A.H. et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364-2368 (2001).

Figure 1:Schematic of construction and verification process of collections. (A) PCR-mediated homologous recombination. Three oligonucleotide primers are designed and synthesized for each ORF: The F2 and R1 oligos, which contain regions of homology to the ORF of interest as well as sequences used to amplify the desired tag, and a CHK oligo, used to verify integration at the correct locus. The F2 and R1 oligos are used to amplify the desired tag and a selectable marker; this PCR product then integrates into the genome at the C-terminus of the ORF of interest. The CHK oligo, as well as an oligo within the tag, is used to verify integration at the desired location. (B) Generalized construction process. We transformed the PCR amplified tag into yeast and plated on selective medium. We then picked individual colonies and verified correct integration of the tag. Correct integrants were then screened for expression of the protein fusion by immunoblot or microscopy and a correctly expressing isolate was selected.

A

ORF X CHK   F2

ORF X

R1

Amplify tag with F2 and R1 primers
Transform product into yeast

TAG   HIS3

ORF X

ORF X   TAG   HIS3

ORF X CHK   F2 CHK

Confirmation PCR with
ORF X CHK and F2 CHK

B

transform ORF-specific
PCR product into yeast

plate on
selective medium

pick individual
colonies

check for correct
integration by PCR

keep 2 positives

check for expression by
immunoblot/microscopy

Figure 2:Examples of verification of integration and expression.(A) 96-well

agarose gel used to analyze check PCRs to confirm proper integration of the tag.

Presence of band indicates proper integration, absence indicates improper

integration. (B) Immunoblot analysis of TAP collection isolates. Arrow indicates

the hexokinase band used as a loading control and quantitation standard. (C)

Fluorescence microscopy of GFP collection isolates. Shown is an example in

which expression from the A collection did not match expression from the B

collection.

A

correct integrant

incorrect integrant

B

←HXK

C

	DIC	GFP	DAPI
YJR117W "A"			
YJR117W "B"			
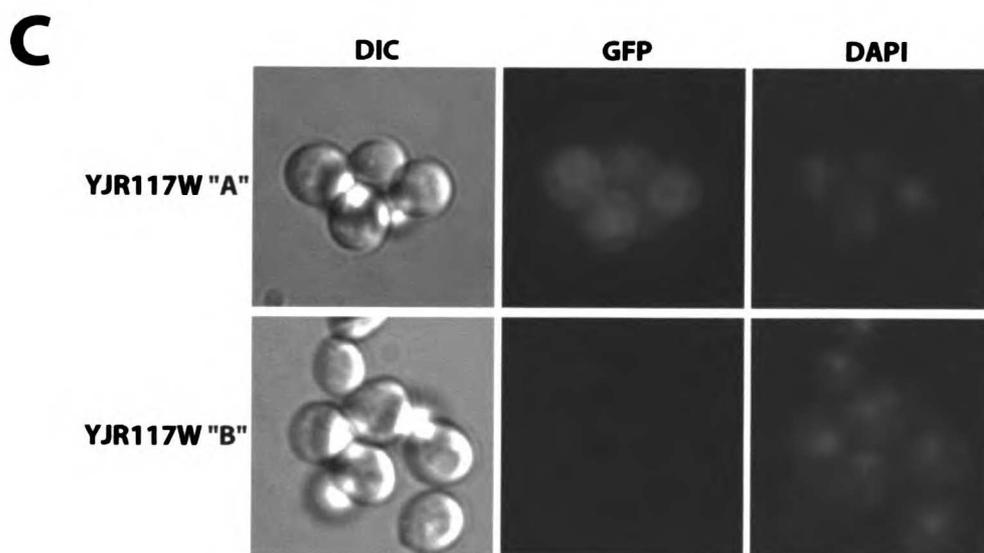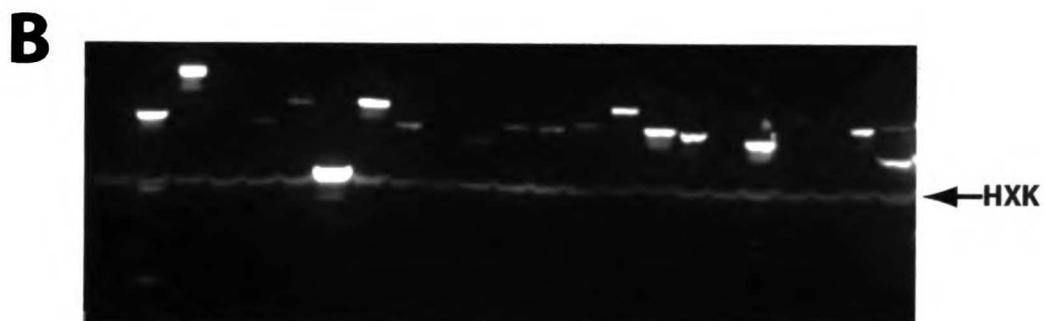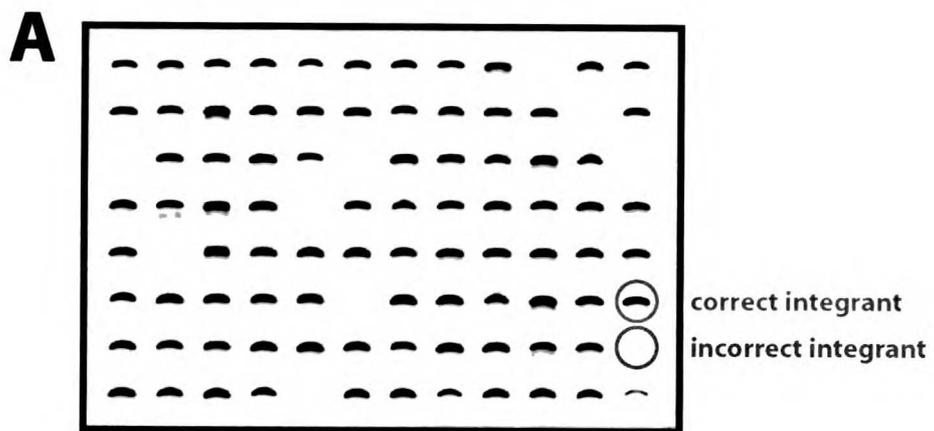
Table 1: Success rates at various stages of construction. Details of success rate

at various stages of construction for the TAP and GFP collections. Success rates

for all ORFs and essential ORFs of each collection are displayed, as well as the

associated percentage. ND=not determined.

Table 1

| | TAP collection | | | | GFP collection | | | |
| | all ORFs | | essential ORFs | | all ORFs | | essential ORFs | |
	#	%	#	%	#	%	#	%
Total ORFs	6234	100.0	1100	100.0	6234	100.0	1100	100.0
Tag amplified with ORF specific PCR	6211	99.6	1096	99.6	ND		ND	
≥1 Transformants obtained	6047	97.0	1014	92.2	6151	98.7	1018	92.5
Positive by PCR	6040	96.9	1003	91.2	6029	96.7	953	86.6
Positive by expression (first round)	3811	61.1	723	65.7	3758	60.3	712	64.7
Positive by expression (after reconstruction)	4268	68.5	821	74.6	4156	66.7	827	75.2

Figure 3: High throughput "multiplex" immunoprecipitations. (A) Schematic of process. Cultures from 6 2 mL 96-well plates are pelleted, combined, and used to prepare extracts. IgG biotin and streptavidin beads are added and allowed to bind, followed by transfer to a filter plate which allows for retention of beads. Beads are washed, proteins are eluted and analyzed by SDS-PAGE and immunoblot. (B) Example immunoblot of results of multiplex immunoprecipitation.
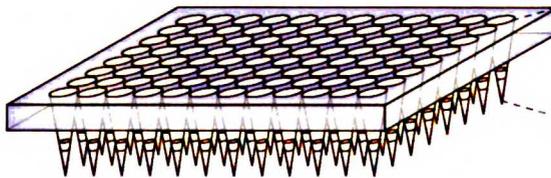
A

2 mL 96-well culture plates
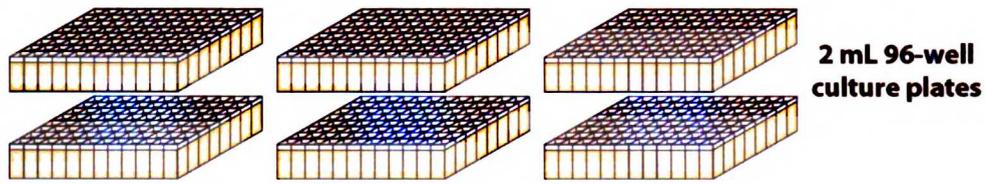
Combine 6 plates into 1 and prepare extracts

1 mL deep well plate
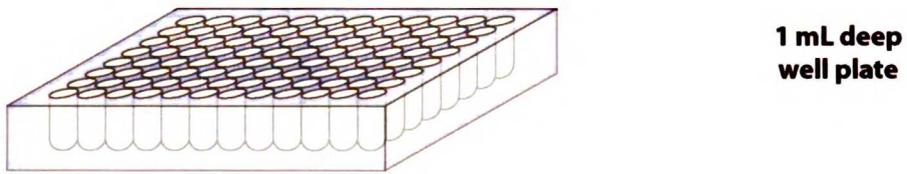
Add IgG biotin and streptavidin beads; transfer to filter plate

Filter plate: beads are retained while liquid flows through

wash and elute; analyze by SDS-PAGE

B
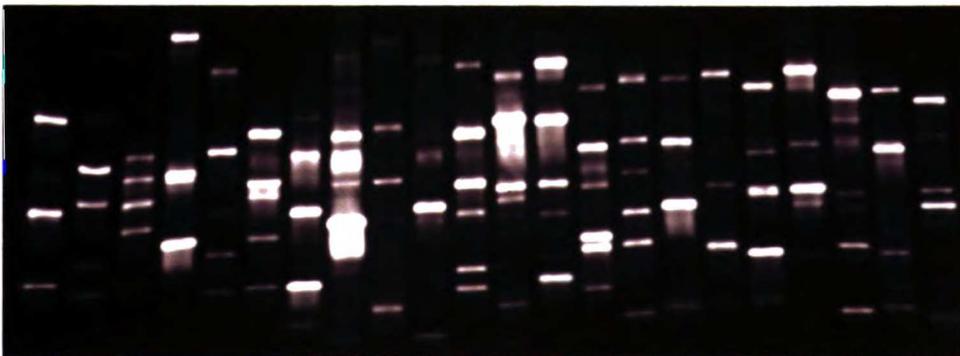
Chapter 4:

High throughput kinase substrate screening with

the TAP collection

Introduction

The budding yeast cyclin dependent kinase (CDK) Pho85 is one of seven CDKs in yeast*(1)*, and is the only one which associates with multiple cyclins and is non-essential. As such, it represents an opportunity to investigate the function and regulation of cyclin dependent kinases, and an extensive amount of genetic work has been done on Pho85 and its ten associated cyclins (Pho85 associated CycLins, or PCLs)*(2-4)*. *PHO85* has a pleiotropic phenotype when disrupted, and has been implicated in a number of cellular processes, including phosphate and glycogen metabolism, response to DNA damage, cell wall integrity and morphogenesis, non-fermentable carbon source utilization, and G1 progression*(5)*. In some cases, it has been possible to ascribe these functions to a group of PCLs, but functional overlap of the PCLs has made interpretation of this data difficult.

Much of the compelling genetic data has come from synthetic lethal analysis. It has been known that *pho85Δ* is synthetic lethal with disruption of the G1 cyclins *CLN1* and *CLN2(6, 7)*. The synthetic lethality of *pho85Δcln1Δcln2Δ* is recapitulated by disruption of *PCL1* and *PCL2*; the *pcl1Δpcl2Δcln1Δcln2Δ* strain is inviable, implicating these two Pcls in G1 progression. Two synthetic lethal screens have also uncovered other genetic interactions with *PHO85*. In the first, Lenburg and O'Shea identified several factors involved in regulation of cellular morphogenesis*(3)*. A subsequent systematic screen greatly expanded the number of known synthetic lethal interactions*(4)*. This screen identified genetic interactions with genes representing a wide variety of cellular processes, but

70

notably many involved in polarized growth and cell wall maintenance. Though this and other strong evidence implicates Pho85 in many important cellular processes, its molecular function in most of these processes is not well understood. Few substrates of Pho85 kinases have been identified, and even fewer have been shown to be physiologically relevant (see Chapter 1).

In order to understand, at a molecular level, how Pho85 is involved in these many processes, we seek to identify the substrates of other Pho85 kinases. Identification of substrates for this kinase (and others) has proved problematic – functional overlap, as well as difficulty in maintaining proper specificity *in vitro* have plagued these studies (for example, see chapter 2). Highlighting this is the fact that every Pcl/Pho85 kinase purified thus far has been able to phosphorylate Pho4 *in vitro(8-10)*, though it is known that Pho80/Pho85 is the only physiologically relevant kinase for this substrate; disruption of *PHO80* is sufficient to render Pho4 completely unphosphorylated *in vivo(11)*.

Of primary importance, then, in identifying new substrates, is ensuring that the proper physiological kinase substrate specificity is recapitulated in whatever approach is chosen. For this reason, we have developed a new kinase assay which is highly specific and robust, and confirmed that it recapitulates the physiological specificity of Pho85 kinases. To search for new substrates of Pho85, we have adapted this assay to efficiently screen the entire yeast proteome. In screening part of the proteome with the Pho80/Pho85 and Pcl1/Pho85 kinases we identify putative novel substrates for both, and perform preliminary characterization of those for Pcl1/Pho85.

Methods

Purification of Pho85 kinases

Purification of Pho85 kinases was performed essentially as described for Pho80/Pho85(8). Briefly, expression plasmids for Pho85-6His (EB1164 for the wild type kinase, EB1375 for the F82G allele) and the desired cyclin (EB1076 for Pho80, EB1348 for Pcl1, and EB1495 for Pcl7) are cotransformed into BL21 DE3 E. coli. Fresh transformants are grown to early logarithmic phase, induced with 200 nM IPTG for 6 hours, pelleted and frozen. Lysates from these cell pellets are loaded on a Ni^{2+} charged chelating column (Pharmacia), and the kinase is eluted with imidazole.

To ensure that purified kinases were indeed heterodimeric Pho85 kinases and not simply Pho85 monomer, we also loaded kinases on a Superose 12 size exclusion column (Pharmacia), calibrated to correctly separate dimeric kinase from Pho85 monomer. The portion of dimeric kinase varied with the cyclin; we kept only dimeric kinase fractions for subsequent studies. Concentrations of purified kinase were measured using UV absorption, and sometimes normalized to kinase activity on a peptide.

The F82G allele of Pho85 complements a pho85Δ in vivo, and displays the same catalytic activity as wild-type Pho85 in vitro(12).

72

Purified components kinase assays

Purified components assays with full length Pho4 and the Pho4 NLS peptide were performed as described(8).

Preparation of radiolabelled N-6-Benzyl ATP

Radiolabelled N-6-Benzyl ATP was prepared from unlabelled N-6-Benzyl ADP using recombinant 12His tagged yeast nucleotide diphosphate kinase (NDPK) purified from *E. coli*. Purified NDPK was applied to Co^{2+} charged IDA sepharose in a disposable chromatography column (BioRad BioSpin column #732-6008). We washed the columns three times with HBS + 5 mM $MgCl_2$, then applied $\gamma-^{32}P$ ATP (ICN #35020). We washed again three times, then applied a small volume of N-6-Benzyl ADP (gift of K. Shokat), eluting three times with HBS +5 mM $MgCl_2$. We collected all washes and elutions and determined the yield by scintillation counting relative to the load.

The in extract kinase assay

The basic kinase assay contains three parts: the extract (prepared as described for native extracts, chapter 3), 3X reaction mix containing ^{32}P radiolabelled N-6-Benzyl ATP [10% glycerol, 120 mM phosphoenolpyruvate, 160 mM β-glycerophosphate, 200 mM Tris pH 7.5, 30 mM NaF, 3 mM ATP, 150 mM NaCl, 10 mM EGTA, 0.3 mg/ml BSA, 25 mM MgCl2, 300 nM calyculin A, 5µCi γ-^{32}P labeled N-6-Benzyl ATP / 200 µg extract], and 3X kinase mix (containing the analog sensitive kinase of interest and pyruvate kinase for ATP regeneration)[30

nM Pho80/Pho85[F82G] or Pcl1/ Pho85[F82G], 0.65 mg/mL pyruvate kinase (Roche), 10% glycerol, 50 mM Tris pH 7.5, 150 mM NaCl, 0.05% NP-40, 1 mM β-mercaptoethanol, 2 mM PMSF, 2.5 mM benzamidine, 1 μg/ml leupeptin, 1 μg/ml pepstatin]. These components are mixed in equal parts at room temperature, and the reaction is allowed to proceed for 15-20 minutes. We subsequently immunoprecipitate the candidate substrate, washing away the extract, kinase, and label. The elutions are run on SDS-PAGE gels, transferred to nitrocellulose, and analyzed by immunoblot and phosphorimager analysis.

High throughput application of the in extract assay

To perform this assay in 96-well format, multiplex extracts were prepared (as in Chapter 3) and frozen in 1 ml deep well 96-well plates. These extracts were thawed, and 125 μl 3X reaction mix, as well as 125 μl 3X kinase mix (3X = 30 nM for Pho80/Pho85[F82G]), were added with a multi-channel pipettor. The reaction proceeded for 15 minutes; we then added 200 μl cold PDMS buffer (see chapter 3) containing biotin-conjugated Human IgG (Jackson Immunoresearch, 2 μl of 1.5 mg/ml stock), and incubated at 4°C for 30 minutes. From this point, immunoprecipitation followed the multiplex protocol discussed in Chapter 3.

Samples were analyzed by immunoblot (as in Chapter 3), and then blots were exposed to a phosphor screen (Molecular Dynamics). These phosphor screens were imaged with a Storm 860 phosphorimager (Molecular Dynamics), and the immunoblot and phosphorimages were compared to analyze the results of the assay.

74

TEV cleavage of identified proteins

To perform TEV protease cleavage on the substrates from the Pcl1 screen, we performed the in extract assay on the individual TAP collection strains as described above except after washing the beads, we split the beads into two tubes (+TEV and −TEV) and resuspended in 25 µl buffer. We then added 2 µl TEV protease (10 U/µl) to the +TEV tubes, and allowed both reactions to sit at room temperature for 30 minutes. After cleavage, we kept the supernatant from the +TEV reactions and added 5µl 6X sample buffer. We removed the supernatant from the −TEV reactions and eluted from the beads with 30 µl sample buffer. These samples were then analyzed by SDS-PAGE and phosphorimaging and compared.

Results

Substrate specificity of Pho85 kinases

To investigate the source of substrate specificity for Pho85 kinases, we compared the abilities of Pho85 kinases to phosphorylate non-physiological substrates under two different regimes. It is known that Pho80/Pho85 phosphorylates the transcription factor Pho4 *in vivo*, and deletion of *PHO80* is sufficient to render Pho4 completely unphosphorylated so it is the only physiological kinase of Pho4. Using recombinant Pho80/Pho85 and Pcl1/Pho85, we investigated the ability of these kinases to phosphorylate full length Pho4 and a peptide derived from Pho4 *in vitro*.

Though Pcl1/Pho85 cannot phosphorylate Pho4 *in vivo*, it is able to do so *in vitro*, albeit less efficiently than Pho80/Pho85 (Figure 1B). In contrast, the Pho4 peptide is phosphorylated approximately equally by both kinases (Figure 1A), and with a much higher $K_m(8)$, suggesting that local protein sequence is not a source of substrate specificity for these kinases.

The progressive loss of specificity observed in less physiological regimes makes it difficult to identify true substrates of these kinases (and others) in an efficient manner. We therefore sought to bridge the gap between specificity and efficiency by developing high throughput methods which more accurately identify protein kinase substrates.

The in extract assay

To identify bona fide *in vivo* kinase substrates, we developed an assay which resembles, to the extent possible, the cellular milieu in which physiologically relevant phosphorylation takes place. This assay is based around mutation of a conserved residue in the kinase ATP binding pocket which confers the ability to utilize a bulky derivative of ATP as the phosphate donor (Figure 2A). Like a wild-type kinase, the mutant kinase retains the ability to utilize normal ATP, but only the mutant can make use of the bulky ATP derivative. The F82G allele of Pho85, which allows use of the ATP analog, complements a *pho85Δ*, and displays the same enzyme kinetics as wild-type Pho85 *in vitro(12)*.

Use of the mutant kinase enables its use in a cellular context containing other kinases and ATP-using enzymes. Though N-6-Benzyl ATP is not cell

76

permeable, and so cannot be used *in vivo*, we sought to replicate this condition

to the extent possible, developing a kinase assay which takes place in an extract

(Figure 2B). The mutant kinase and radiolabelled N-6-Benzyl ATP are added to

an extract, and a kinase reaction takes place. During this reaction, all cellular

ATPases are active, though only the mutant kinase is able to use the

radiolabelled ATP analog, so only the substrates of the mutant kinase will

become radiolabelled. To analyze a candidate substrate, it is immunoprecipitated

from the extract after the reaction and analyzed by SDS-PAGE and

phosphorimaging.

To test the quality of this assay, we compared the ability of two mutant

kinases to phosphorylate a protein which is the physiological substrate of only

one of the kinases. We analyzed phosphorylation of Pho4, a transcription factor

which controls the response to phosphate starvation, which is a known substrate

of the cyclin-CDK Pho80/85. We compared the ability of Pho80/Pho85[F82G] to

that of Pcl7/Pho85[F82G] to phosphorylate Pho4 in this assay. Pcl7 is the Pho85

associated cyclin most closely related to Pho80, but cannot phosphorylate Pho4

in association with Pho85 *in vivo*, as Pho4 is completely unphosphorylated in a

pho80Δ strain. Furthermore, Pcl7/Pho85 is able to phosphorylate Pho4 *in vitro*,

though not as well as Pho80/85 (Figure 3B). It therefore serves as a stringent

test of the specificity of this assay, as its true substrate specificity is not

maintained in an assay with purified components.

To compare the activities of these two kinases on Pho4 in this assay, we

titrated the recombinant kinases and monitored Pho4 phosphorylation. These

77

kinases display the same activity on a Pho4 peptide, so the concentrations are directly comparable. As shown in Figure 3A, Pho80/Pho85[F82G] is able to phosphorylate Pho4 efficiently, and detectable phosphorylation is observed at concentrations as low as 0.5 nM. In marked contrast, Pcl7/Pho85[F82G] is not able to detectably phosphorylate Pho4 except at a concentration of 50 nM, and even this phosphorylation is not efficient.

While it is qualitatively clear that this assay is very specific, we also quantified the specificity by dividing the degree of Pho4 phosphorylation from Pho80/Pho85[F82G] by that from Pcl7/Pho85[F82G] at each concentration. As shown in Figure 3C, the in extract assay shows a significantly higher degree of specificity than an *in vitro* assay with purified components. In fact, these numbers do not do justice to the specificity of this assay, as the phosphorylation of Pho4 by Pcl7/Pho85[F82G] at 0.5 and 5 nM is not significantly above background, and Pho80/Pho85[F82G] has probably completely phosphorylated Pho4 at 50 nM.

Application of the in extract assay to high throughput analysis

Though we have shown that the in extract assay is robust and specific, it requires a candidate substrate. In order to identify new Pho85 kinase substrates, we chose to consider the entire yeast proteome as candidate substrates, and developed methodology to efficiently perform this assay on every yeast protein.

We had previously constructed a collection of yeast strains in which almost every ORF is C-terminally fused to the TAP (Tandem Affinity Purification) tag (Chapter 3). Further, we have shown that the TAP tagged proteins can be

efficiently immunoprecipitated in a 'multiplex' format. Briefly, this involves making extracts from six different strains in each well of a 96 well plate, followed by immunoprecipitation and washing in a 96 well filter plate.

To adapt the in extract assay to the multiplex format, we simply performed the kinase reaction in 96 well format after making extracts, then proceeded with the immunoprecipitation procedure as described (Figure 4). We chose 10 nM as the kinase concentration at which to screen, as a compromise between maximizing signal strength and maintaining specificity.

Screening the TAP collection with Pho80/Pho85 and Pcl1/Pho85

Preliminary work in the lab identified rough expression levels of fusion proteins in the TAP collection (S. Ghaemmaghami, personal communication). Based on these results, we reorganized the collection into 5 distinct expression categories (DB0-DB4, DB4 = highest expression), maintaining organization by size within each category. The last column of each 96-well plate was left empty for controls. We first chose to screen the soluble proteins with DB score ≥ 2, as expression of these proteins has been verified, and they would be the most likely to be purified correctly. When combining cultures for multiplex immunoprecipitation, we always combined cultures from strains with the same expression score, in order to minimize effects of competition for binding to beads. We screened about 2400 proteins with Pho80/Pho85 and Pcl1/Pho85, using 10 nM of each kinase for all reactions (Figure 4).

The first observation one can make from looking at the overall data is that these two highly related kinases yield strikingly different profiles of phosphorylated proteins identified in this screen. Though both screens identified phosphoproteins, there is no overlap of phosphoproteins between these two kinases. This reiterates the specificity inherent in the assay, and adds to the confidence in the identified proteins as *bona fide* substrates. Another rough observation is that Pho80/Pho85 identifies far fewer hits than Pcl1/Pho85. This could mean that Pcl1/Pho85 has a broader spectrum of targets, which would be consistent with genetic data demonstrating its involvement in G1 progression and cell wall integrity.

We identified 25 phosphorylated bands in the phosphorimages from the Pcl1/Pho85 screen (Figure 4, red circles), representing 140 possible TAP fusion proteins associated with these bands. For Pho80/Pho85 we identified 3 phosphoproteins (Figure 4, red circles), representing 12 possible TAP strains. Though we did not identify Pho4 as a substrate of Pho80/Pho85, it was not among the proteins screened, so unfortunately does not serve as a positive control for these screens.

Deconvolution of Pcl1/Pho85 substrates

We examined the results of the Pcl1/Pho85 screen and identified lanes containing phosphorylated bands. Each of these 25 bands could represent any of the proteins pulled down in this reaction, so it was necessary to deconvolve the positives, and identify with which TAP collection strain each phosphorylated band

is associated. To do so, we selected each of the 140 possible individual strains from all of the hits (detailed in Table 1), assembled into 96-well format, and repeated the assay on the individual strains (Figure 5). Instead of combining different strains in the extract preparation stage, we instead combined six identical culture plates, for an effective culture volume of 12 ml for each strain. This also has the added benefit of increasing the signal over background, as the background phosphorylation observed is independent of the presence of a TAP tagged protein or a mutant kinase.

Deconvolution identified 19 individual strains which yielded a phosphoprotein in this assay. These 19 phosphoproteins are of identical mobility to the corresponding phosphoproteins identified in the screen.

Identifying true substrates of Pcl1/Pho85

In examining the proteins identified in the Pcl1/Pho85 screen, we noticed that many of them have been previously identified as being in complexes with each other. Because immunoprecipitation reactions in the screen were performed under native conditions, two possibilities exist for each phosphoprotein identified as a substrate of Pcl1/Pho85. First, the phosphoprotein could be the TAP tagged protein itself. Alternatively, the phosphoprotein could simply be an associated protein which is immunoprecipitated along with the TAP tagged protein. To distinguish between these two possibilities, we took advantage of the fact that the TAP tag contains a TEV protease cleavage site, which allows for the removal of the two Z domains from fusion proteins. We

reasoned that if the TAP tagged protein itself is phosphorylated, then cleavage with TEV protease should result in a shift of apparent molecular weight on an SDS-PAGE gel corresponding to loss of the ZZ portion of the tag (Figure 6A). On the other hand, if the phosphoprotein is simply a protein associated with the TAP tagged protein, no such shift should be observed.

We tested 18 of the individual strains by this assay (Figure 6B, results summarized in Table 2), and found that 8 displayed a mobility shift upon TEV protease cleavage, confirming that the TAP fusion protein itself is phosphorylated. 8 other phosphoproteins did not shift, indicating that some other associated protein is in fact phosphorylated, and 2 were not detectable after TEV treatment, perhaps indicating degradation or inaccessibility to the protease site.

Discussion

Though an extensive amount of genetic work on Pho85 has implicated it in a number of cellular processes, how it functions in these processes is not well understood. At the molecular level, only two substrates of Pho85 have been identified and shown to have a physiological role. We seek to understand the molecular functions of the Pcl/Pho85 kinases by identifying their substrates. To do so, we have developed a new kinase assay which replicates the physiological specificity of the different Pho85 kinases, and adapted this assay to high throughput analysis of the yeast proteome. In screening part of the proteome with the Pho80/Pho85 and Pcl1/Pho85 kinases, we observe distinct patterns of phosphoprotein hits, further validating the specificity of this assay.

Success of the assay and the screen

Several observations highlight the potential of this high-throughput kinase assay. First, Pho80/Pho85 and Pcl1/Pho85 exhibit unique sets of *non-overlapping* putative hits in these screens. Second, in the Pcl1/Pho85 screen, we identified several cases in which we pulled down multiple members of a complex in which one component is phosphorylated. In addition to the added confidence from identifying the same substrate multiple times, this observation informs us that proteins exist in their native complexes in this assay, which may be important in maintaining proper specificity. Furthermore, though many members of these complexes contain SP or TP dipeptides capable of being phosphorylated, they are not, again underscoring the specificity of the assay.

83

More important than the number of putative substrates we identified is the large number of negatives observed in these screens. Despite screening almost 2400 proteins with each kinase, we only observed a handful of positives. Perhaps the most compelling finding in evaluating the success of this screen, though, is the particular group of putative substrates identified for Pcl1/Pho85 (discussed below), as their functions are in many cases consistent with previously defined roles of Pcl1/Pho85.

Another potential source of verification comes from phosphopeptides identified by a mass spectrometry approach. Of 216 phosphopeptides identified in this unbiased study, peptides containing proline directed phosphoserines were recovered from two putative Pcl1/Pho85 substrates, Ssd1 and Npl3*(13)*. Though there is no link to Pho85, this does demonstrate that both of these proteins are phosphorylated under physiological conditions at residues which Pho85 is potentially capable of phosphorylating. These sites are likely candidates to pursue in identifying which residues on these proteins are phosphorylated in the extract assay.

The presumed specificity of the assay can be attributed to two factors. First, the native context of the extract maintains much of the physiological conditions present in the cell. Substrates must compete with all other cellular proteins in the extract for recognition by the kinase. Though not all possible sources of specificity, such as subcellular localization, are present, proteins remain in a chemically native context, probably in physiological complexes, as can be inferred from the fact that phosphorylated proteins copurify with multiple

members of observed protein complexes. These complexes may also be important for proper kinase substrate recognition.

Second, the absolute concentrations of kinase and substrates in this assay are close to their cellular concentrations. The TAP fusions are under the control of the endogenous promoter, and an appropriate concentration of recombinant kinase can be added. Thus, the true physiological enzymatic environment of the kinase is replicated, and only physiological substrates are phosphorylated. Because candidate substrates are subsequently immunoprecipitated, effectively concentrating the reaction after it has occurred, low (physiological) kinase concentrations can be used due to the enhanced sensitivity in detecting phosphoproteins. This stands in stark contrast to the concentration of kinase (100 nM) needed for a robust signal when screening the GST library.

True validation of the specificity of this assay, though, will entail the confirmation of these substrates as being physiologically relevant (discussed below).

Potential caveats of the assay

Though this assay appears to have been successful in identifying novel substrates of Pho85, several potential caveats to this approach exist. First, though the conditions of the assay seek to replicate the cellular context of physiological phosphorylation, it is still inherently an *in vitro* approach. Some contributing factors to kinase substrate specificity may not still exist in this

context, most notably subcellular localization. Phosphorylation may also only occur under specific physiological conditions (i.e. in response to environmental stimuli), and inhibitors of phosphorylation may exist in the extract, preventing observation in this assay.

Second, the TAP fusion protein may not be expressed under the growth conditions used. This would clearly make it impossible to identify phosphorylation of this class of substrates. One specific instance of this is the transcription factor Gcn4. Pho85 is known to mediate the stability of this protein, probably with Pcl5*(14)*. Gcn4, however, is only expressed under conditions of amino acid starvation, so this substrate would probably elude this assay. The screen could possibly be further refined to examine other environmental conditions simply by growing cells under these conditions, so perhaps could identify substrates specific to different physiological regimes.

Lastly, the substrate may already exist in a fully phosphorylated state during the reaction, making further phosphorylation and resultant radiolabelling impossible. The positive control Pho4 potentially addresses this concern: it is fully phosphorylated under normal growth conditions, but is presumably dephosphorylated in the process of making extracts. The lability of phosphorylation, however, may vary greatly from protein to protein. The addition of a phosphatase treatment to the extracts prior to the reaction could potentially alleviate this concern.

Putative Pho80/Pho85 substrates

Though we did not deconvolve the hits for Pho80/Pho85, upon examining the potential individual strains associated with these hits, two interesting possibilities emerge. Both Pho81-TAP and Glc8-TAP are among these possibilities, and the phosphoprotein observed matches the predicted size of these proteins in both cases. Pho81 is a CDK inhibitor which binds to Pho80/Pho85 in a ternary complex, and regulates its activity on Pho4. Though it is required for proper expression of phosphate responsive genes, it is not known how Pho81 itself is regulated. Pho81 is not phosphorylated under phosphate starved conditions (Dennis Wykoff, personal communication), but it is possible that it is negatively regulated by phosphorylation under phosphate replete conditions.

Glc8 is a regulator of the phosphatase Glc7*(15)*. There is some evidence that Glc7 is the phosphatase which acts on Pho4 (J. Raser, personal communication), so phosphorylation of Glc8 could be a potential source of regulation of this process. Glc8 has also been identified as a substrate of Pcl7/Pho85, both by a purified components approach within the lab (N. Dephoure, personal communication) and by another group*(16)*. As Pcl7 is the most closely related cyclin to Pho80, it is possible that proper specificity was not maintained in these approaches or in ours, or it is possible that both kinases act on Glc8. Supporting this, both kinases are able to phosphorylate Glc8 in the in extract assay, and do so with equal efficiency (data not shown).

Putative Pcl1/Pho85 substrates

We have further investigated the phosphoprotein positives from the Pcl1/Pho85 screen, identifying the individual TAP fusion strains associated with these phosphoproteins, and performing TEV protease cleavage on the positives to determine whether the TAP fusion itself is the phosphorylated protein. The putative substrates identified for Pcl1/Pho85 in the screen provide a compelling list of possible functions of this kinase. The eight proteins which we have identified fall into three distinct categories, with various connections to known functions of Pho85

The first group of putative substrates consists of four proteins (Rom2, Sap185, Ssd1, and Vip1), all of which are consistent with the previously characterized role of Pcl1 in polarized growth and morphogenesis. Rom2 is the guanine nucleotide exchange factor for Rho1 in the PKC pathway*(17)*, to which Pho85 is intimately linked genetically. Also, a *rom2Δpho85Δ* strain is inviable*(4)*. Sap185 associates with the Sit4 phosphatase, and has a role in G1 progression*(18)*. Ssd1 functions in maintenance of cell wall integrity, and displays a strong genetic interaction with Sit4*(19)*. Lastly, Vip1 functions in the cortical actin cytoskeleton*(20)*.

A second group of substrates consists of Npl3 and Rpg1. At first glance, these two proteins, an mRNA binding protein and a translation initiation factor, do not seem to relate to known functions of Pho85. Npl3 is responsible for mRNA binding and export and seems to have a role in localizing mRNA to the bud tip*(21)*. Rpg1 associates and colocalizes with Sla2, a structural constituent of the

actin cytoskeleton, at cortical actin patches and at the incipient bud site(22).

Perhaps these proteins function in promoting localized translation at the

emerging bud, and are regulated by Pcl1/Pho85 in doing so. Though speculative,

this model is again consistent with previously characterized roles of Pcl1, and

also possibly the requirement of Pho85 for proper localization of Ash1

protein(23).

A third class of substrates has a more distant connection to Pho85. Cdc19

is a pyruvate kinase, is required for START in the cell cycle, and influences

carbon flux during fermentation(24). Ugp1 is involved in UDP-glucose

metabolism(25), and is induced upon loss of Pho85 function(12, 26). Both of

these proteins function in carbon source utilization, in which Pho85 is involved,

albeit not necessarily through Pcl1.

The remaining proteins which were either not cleaved or not resolved in

the TEV experiment await further characterization. They seem to be involved in a

variety of processes, including transcription (Rpb2, Rpb7, Rpb11), septin

organization (Kcc4, Cdc10), response to DNA damage (Hrr25), G1 progression

(Cka2), and carbon source utilization (Pfk1). Their connection to Pho85 remains

to be seen.

Possibilities for the future

Validation of substrates

An important next step in the pursuit of this screen, and the approach in

general, is to validate the physiological relevance of the identified substrates. As

can be seen from the studies using the GST library (Chapter 2), it is difficult to interpret the meaning of any proteins identified by this approach without some evidence that these interactions are relevant *in vivo*. One approach to validating these substrates is to metabolically label cells with ^{32}P orthophosphate and compare the degree of phosphorylation in strains which vary in the activity of Pho85 (compare wild-type to *pho85Δ*, or *PHO85(F82G)* with and without inhibitor present). Though this approach has been successful in the past, it is laborious and can be difficult to accurately quantify phosphorylation. Another similar approach takes advantage of a newly developed protein stain which specifically binds phosphate groups*(27)*. One can immunoprecipitate a candidate substrate from conditions varying in Pho85 activity, and quantify degree of phosphorylation relative to amount of protein (by quantitative total protein stain). This approach has the advantage of being easier to implement on a larger scale (i.e. with many candidate substrates), and can detect phosphorylation quantitatively, enabling detection of reduced phosphorylation in cases where other sites on the same protein are phosphorylated by a different kinase (as could be the case with Npl3 and Sky1).

One potential barrier to validation in this manner is if any residues on a protein are phosphorylated by more than one kinase in a redundant manner. It is not immediately clear how to circumvent this issue, so it is important to keep this possibility in mind when investigating the physiological relevance of these candidate substrates.

Another important step in validating these substrates is to map the residues which are phosphorylated by Pho85 in the candidate substrates, and verify that these sites are the same as those phosphorylated *in vivo*. The most conclusive proof of physiological relevance would be to show that phosphorylation of one or more of these individual residues is dependent on Pho85.

Validation of the approach by identifying previously characterized substrates

One other way of assessing the success of this approach is to ask how complete the screen is, and whether known *bona fide* substrates are identified. In other words, it is important to know not only that the substrates identified in the screen are correct, but also that other true substrates are not missed. This gives some power to the negative results from the screen, as it is informative to know which proteins are *not* phosphorylated by a given kinase. Though the screen cannot be perfect, in that not all proteins are represented by the collection and the existence of variability in pulldown efficiency inherent in the technique, this information still remains useful in that it tells us about the fraction of the proteome which is impacted by a given kinase.

For this reason, it will be important to complete the screen with both kinases on all strains in the TAP collection. Most importantly, Pho4 must be identified as a substrate of Pho80/Pho85; this will confirm the utility of this approach, especially compared with the GST library approach, for which Pho4 was not identified as a substrate of Pho80/Pho85. In various reconstruction

experiments Pho4 has successfully been identified as a substrate, but we will also confirm that it is identified in the current incarnation of the assay and the high-throughput screening format.

Preliminary results from the continuation of the screen have identified Rvs167 as a substrate of Pcl1/Pho85 (Noah Dephoure, personal communication). Though Rvs167 has not been characterized as extensively as Pho4, it has been shown to be phosphorylated *in vivo*, and this phosphorylation is dependent on the Pcl1,2 subfamily of Pho85 cyclins*(9)*. Its identification, then, lends further confidence in the accuracy and completeness of this approach.

Identification of other substrates of Pho80/Pho85 and Pcl1/Pho85

Completing the screen of the TAP collection with these kinases will undoubtedly yield additional candidate substrates, especially for Pcl1/Pho85. Some of these maybe the true substrates from TAP proteins already identified as being associated with a phosphoprotein but which were not cleaved in the TEV protease experiment. This, however, may not be the case; it is possible that a TAP fusion strain was never successfully constructed, or that the TAP fusion itself is either nonfunctional or cannot be efficiently immunoprecipitated. In this instance, it will be necessary to seek alternate means to identify the proper phosphoprotein. A possible approach would be to purify the TAP complex and identify phosphopeptides by mass spectrometry.

One substrate which has eluded identification thus far is of particular interest. We identified three subunits of RNA polymerase II (Rpb2,Rpb7, and

Rpb11) which yielded a phosphoprotein in the extract assay but which were not

cleave in the TEV protease experiment. Because of the large size of this

phosphoprotein, it is tempting to speculate that the true substrate in these

complexes is Rpo21, the core subunit of RNA polymerase II. There is evidence

that 4 other CDKs in yeast (Kin28, Srb10, Ctk1, and Bur2) are capable of

phosphorylating the C-terminal domain of this protein, and they seem to have

distinct functional roles in doing so(28). It would be quite interesting (and curious)

if Pho85 were also capable of phosphorylating this protein; one can speculatively

imagine several functional reasons for this. Clearly this could also be an artifact

of the approach, and it will be important to keep this in mind along with other

assessments of the specificity and validity of the approach.

Other putative Pho85 substrates

Several other proteins have been proposed to be substrates of Pho85 in

the literature. For some of these (notably Pho4, Gsy2, and Rvs167), there is

evidence of physiological relevance, but others are lacking in this respect (such

as Sic1(29), Swi5(30), and Glc8(16)). It will be interesting to see if some of these

are identified by the screen, especially the putative Pcl1/Pho85 substrate Sic1. It

will also be informative to test these candidate substrates individually in the

extract assay with their respective proposed Pho85 kinases, as well as the other

Pho85 kinases. We may be able to assess the verity of these substrates, and

could gain further valuable information about the specificity of this assay.

It will also be interesting to screen for substrates of other Pho85 kinases besides Pho80/Pho85 and Pcl1/Pho85. Identification of these downstream effectors will hopefully enable further understanding of Pho85 function. In addition, knowledge of many substrates of different Pcl/Pho85 kinases may ultimately aid in an understanding of how cyclin binding confers substrate specificity to CDKs, a subject about which little is known(31).

Establishing functional biological relevance of the substrates

A daunting, but necessary, challenge of future work will be to link whatever substrates are identified to the biology of Pho85. Some tempting leads exist, but there is no clear path of how to accomplish this. Initial efforts will undoubtedly focus on candidates which have been previously characterized and for which some known function already exists. Rom2 is an obvious example of this, as it is known to be the guanine nucleotide exchange factor for Rho1, and is involved in the PKC pathway which is linked by synthetic lethal interactions to Pho85. We are lucky in that the Pcl1/Pho85 substrates identified thus far all have been characterized to some extent, but there is some variation in the degree of characterization, and future candidate substrates may have little to no characterization. In these cases, a starting point will be to mutate the sites of phosphorylation in the substrates and look for a phenotype consistent with functions of Pho85.

These and other unforeseen barriers will undoubtedly complicate the understanding of the molecular functions of Pho85. The establishment of the

94

activities of Pho85 substrates, and how and why Pho85 regulates the processes in which they are involved, is the ultimate aim of this course of study. Hopefully, we will be able to further establish the validity of the substrates identified by this approach so that more detailed investigation will at least rest on confidence in their physiological relevance.

Literature cited

1. Dolinski, K., Balakrishnan, R., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hirschman, J. E., Hong, E. L., Issel-Tarver, L., Sethuraman, A., Theesfeld, C. L., Binkley, G., Lane, C., Schroeder, M., Dong, S., Weng, S., Andrada, R., Botstein, D., and Cherry, J. M. Saccharomyces Genome Database. *http://genome-ftp.stanford.edu/pub/yeast/SacchDB*.

2. Measday, V. et al. A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. *Mol Cell Biol* **17**, 1212-1223 (1997).

3. Lenburg, M.E. & O'Shea, E.K. Genetic evidence for a morphogenetic function of the Saccharomyces cerevisiae Pho85 cyclin-dependent kinase. *Genetics* **157**, 39-51 (2001).

4. Huang, D., Moffat, J. & Andrews, B. 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 (2002).

5. Carroll, A.S. & O'Shea, E.K. Pho85 and signaling environmental conditions. *Trends Biochem Sci* **27**, 87-93 (2002).

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

7. Measday, V., Moore, L., Ogas, J., Tyers, M. & Andrews, B. The PCL2 (ORFD)-PHO85 cyclin-dependent kinase complex: a cell cycle regulator in yeast. *Science* **266**, 1391-1395 (1994).

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

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

10. Huang, D. et al. 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 (1998).

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

12. Carroll, A.S., Bishop, A.C., DeRisi, J.L., Shokat, K.M. & O'Shea, E.K. Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proc Natl Acad Sci U S A* **98**, 12578-12583 (2001).

13. Ficarro, S.B. et al. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat Biotechnol* **20**, 301-305 (2002).

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

15. Cannon, J.F., Pringle, J.R., Fiechter, A. & Khalil, M. Characterization of glycogen-deficient glc mutants of Saccharomyces cerevisiae. *Genetics* **136**, 485-503 (1994).

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

17. Ozaki, K. et al. Rom1p and Rom2p are GDP/GTP exchange proteins (GEPs) for the Rho1p small GTP binding protein in Saccharomyces cerevisiae. *Embo J* **15**, 2196-2207 (1996).

18. Luke, M.M. et al. The SAP, a new family of proteins, associate and function positively with the SIT4 phosphatase. *Mol Cell Biol* **16**, 2744-2755 (1996).

19. Kaeberlein, M. & Guarente, L. Saccharomyces cerevisiae MPT5 and SSD1 function in parallel pathways to promote cell wall integrity. *Genetics* **160**, 83-95 (2002).

20. Feoktistova, A., McCollum, D., Ohi, R. & Gould, K.L. Identification and characterization of Schizosaccharomyces pombe asp1(+), a gene that interacts with mutations in the Arp2/3 complex and actin. *Genetics* **152**, 895-908 (1999).

21. Gilbert, W., Siebel, C.W. & Guthrie, C. Phosphorylation by Sky1p promotes Npl3p shuttling and mRNA dissociation. *Rna* **7**, 302-313 (2001).

98

22. Palecek, J., Hasek, J. & Ruis, H. Rpg1p/Tif32p, a subunit of translation initiation factor 3, interacts with actin-associated protein Sla2p. *Biochem Biophys Res Commun* **282**, 1244-1250 (2001).

23. McBride, H.J. et al. The protein kinase Pho85 is required for asymmetric accumulation of the Ash1 protein in Saccharomyces cerevisiae. *Mol Microbiol* **42**, 345-353 (2001).

24. Pearce, A.K. et al. Pyruvate kinase (Pyk1) levels influence both the rate and direction of carbon flux in yeast under fermentative conditions. *Microbiology* **147**, 391-401 (2001).

25. Daran, J.M., Dallies, N., Thines-Sempoux, D., Paquet, V. & Francois, J. Genetic and biochemical characterization of the UGP1 gene encoding the UDP-glucose pyrophosphorylase from Saccharomyces cerevisiae. *Eur J Biochem* **233**, 520-530 (1995).

26. Nishizawa, M., Tanabe, M., Yabuki, N., Kitada, K. & Toh, E.A. Pho85 kinase, a yeast cyclin-dependent kinase, regulates the expression of UGP1 encoding UDP-glucose pyrophosphorylase. *Yeast* **18**, 239-249 (2001).

27. Schulenberg, B., Aggeler, R., Beechem, J.M., Capaldi, R.A. & Patton, W.F. Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. *J Biol Chem* **278**, 27251-27255 (2003).

28. Murray, S., Udupa, R., Yao, S., Hartzog, G. & Prelich, G. Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase. *Mol Cell Biol* **21**, 4089-4096 (2001).

29. Nishizawa, M., Kawasumi, M., Fujino, M. & Toh-e, A. 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 (1998).

30. Measday, V., McBride, H., Moffat, J., Stillman, D. & Andrews, B. Interactions between Pho85 cyclin-dependent kinase complexes and the Swi5 transcription factor in budding yeast. *Mol Microbiol* **35**, 825-834 (2000).

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

Figure 1: Phosphorylation of Pho4 and a Pho4 peptide by Pho80/Pho85 and Pcl1/Pho85. (A) Purified Pho80/Pho85 and Pcl1/Pho85 were used to phosphorylate a peptide derived from the Pho4 nuclear localization sequence (NLS). Peptides were phosphorylated *in vitro*, bound to nitrocellulose filters, and degree of phosphorylation was quantified by scintillation counting. (B) Purified Pho80/Pho85 and Pcl1/Pho85 were used to phosphorylate full length recombinant Pho4 purified from *E. coli in vitro*. Samples were analyzed by SDS-PAGE and phosphorimaging.

A

Pho4 NLS Peptide Phosphorylation

Y-axis: Arbitrary scintillation counter units
X-axis: Time (minutes)

Legend:
- ♦ Pcl1/Pho85
- ■ Pho80/85

B

Pcl1/Pho85 (10 nM)   Pho80/85 (2 nM)

1'   4'   7'   10'   1'   4'   7'   10'

Phosphorylation of Pho4

Y-axis: Arbitrary phosphorimager units
X-axis: Time (minutes)

Legend:
- ■ Pho80/85
- ♦ Pcl1/Pho85

102
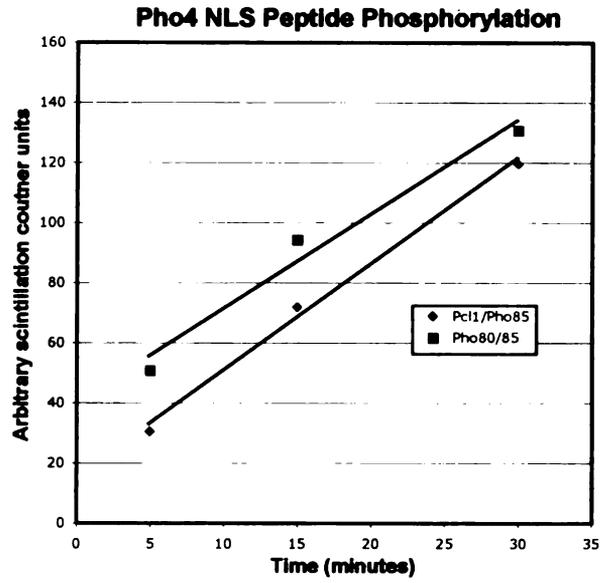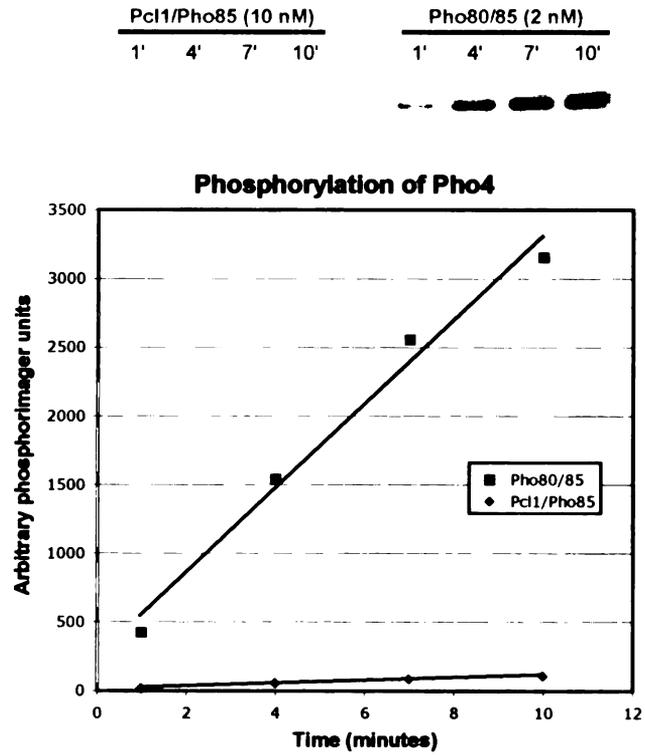
Figure 2: The analog sensitive kinase and the in extract assay.(A) A conserved hydrophobic residue in the ATP binding pocket of all kinases is mutated to a smaller residue. In the case of Pho85 (bottom row), phenylalanine 82 is mutated to glycine. Like the wild-type kinase, the mutant retains the ability to use ATP (middle column), but can also use N-6-Benzyl ATP which is unusable by wild-type kinases (right column). (B) For the assay, purified mutant kinase and radiolabelled N-6-Benzyl ATP are added to an extract. The reaction takes place in the extract, and the candidate substrate of interest is immunoprecipitated. Phosphorylation is analyzed by SDS-PAGE and phosphorimaging.

A

wild-type kinase

mutant kinase

B

Add mutant kinase and radiolabelled N-6-Benzyl ATP to extract

reaction in extract

IP candidate substrate

Analyze by SDS-PAGE and phophorimager
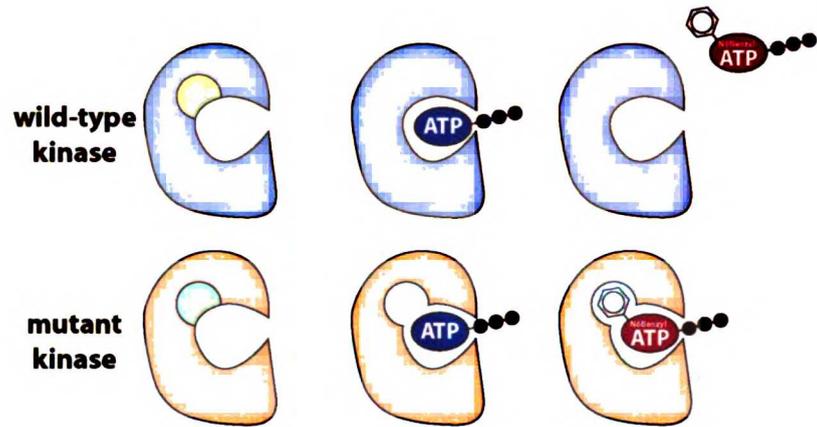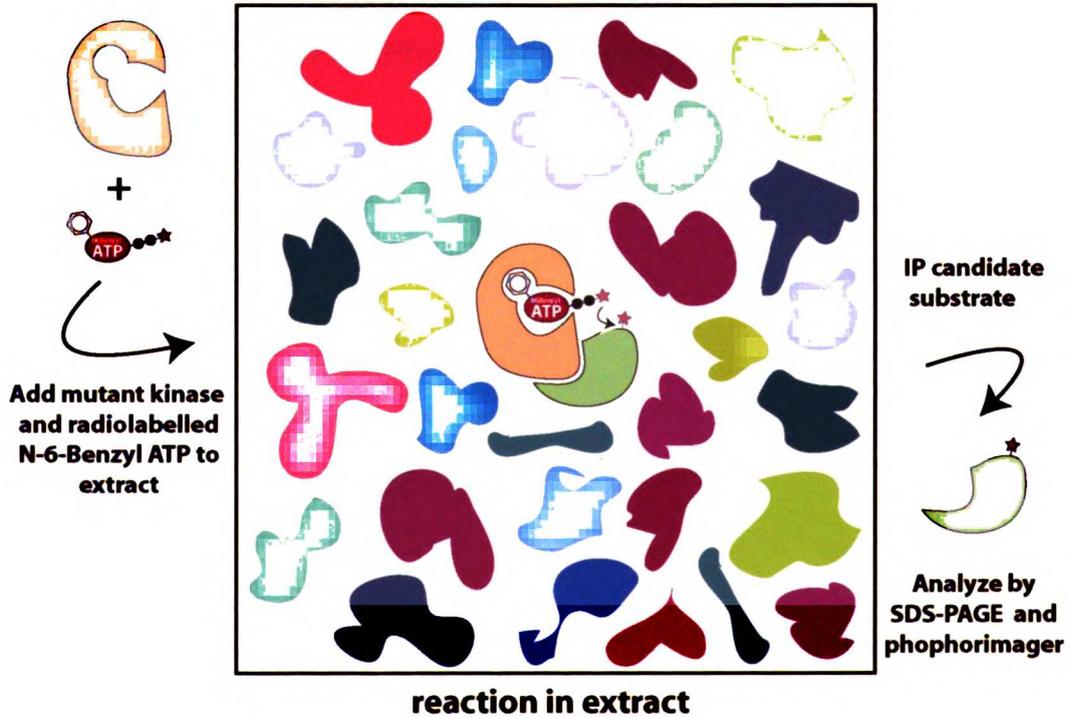
Figure 3: Specificity of the in extract assay. (A) Pho80/Pho85[F82G] and

Pcl7/Pho85[F82G] were used to phosphorylate Pho4 with the in extract assay.

The assay was performed with each kinase at three concentrations, and results

were analyzed by SDS-PAGE and phosphorimaging. (B) Pho80/Pho85 and

Pcl7/Pho85 were used to phosphorylate Pho4 with an *in vitro* purified

components assay. A time course with each kinase was performed, and results

were analyzed by SDS-PAGE and phosphorimaging. (C) Comparison of fold

specificity with the two assays. Fold specificity for the in extract assay was

calculated by dividing the degree of phosphorylation of Pho4 with Pho80/Pho85

by that with Pcl7/Pho85 at 5 nM kinase concentration. Fold specificity for the

purified components assay was calculated by dividing the slope in B for

Pho80/Pho85 by that for Pcl7/Pho85, then multiplying by 5 to account for the

difference in concentration.

A

Pho4-ZZ phosphorylation

B

Pho4 phosphorylation

C

Fold Specificity Pho80/Pcl7
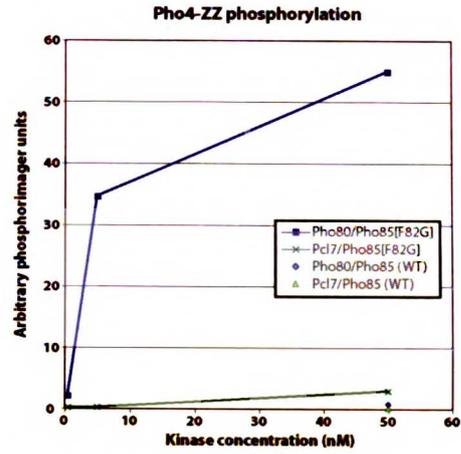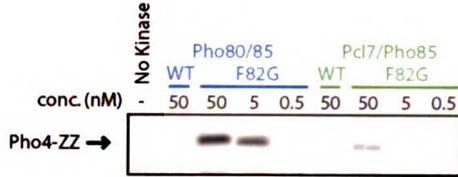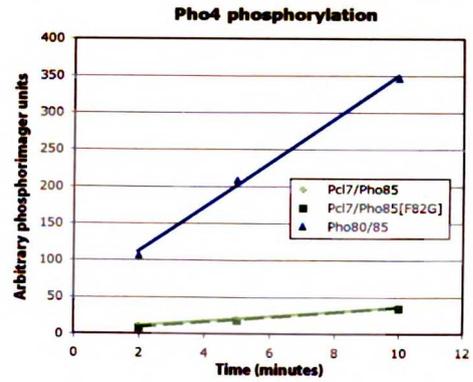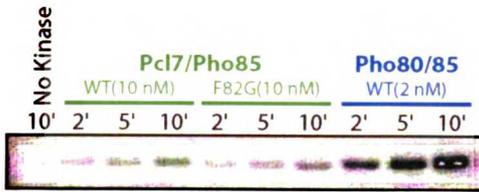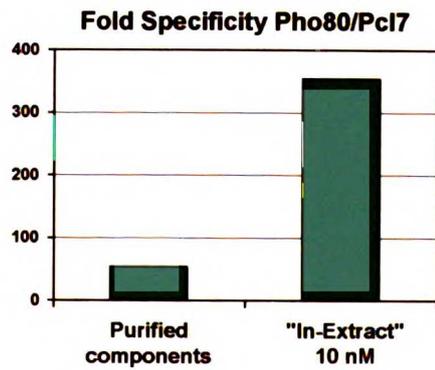
Figure 4: Screening the TAP collection for substrates of Pho80/Pho85 and Pcl1/Pho85. The in extract assay was performed in high throughput format with Pho80/Pho85[F82G] at a concentration of 10 nM. For each gel, a phosphorimage and immunoblot are shown. The first number and letter denote the multiplex plate which was screened. Each multiplex plate is analyzed on four gels. The last number indicates the gel number, with gel 1 representing rows A and B, gel 2 representing rows C and D, gel 3 representing rows E and F, and gel 4 representing rows G and H. Positives are circled in red.
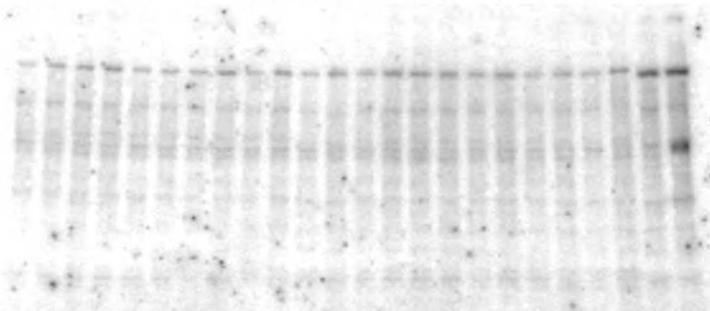
phosphorimage

**Pho80
2A1**

immunoblot
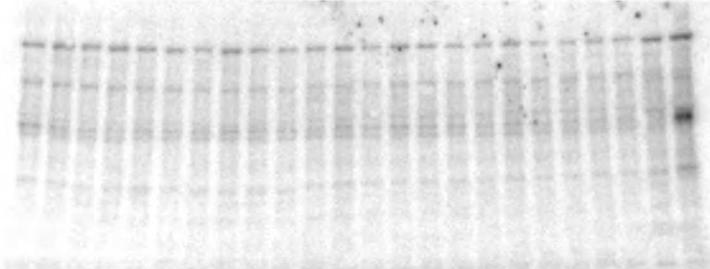
phosphorimage

**Pho80
2A2**

immunoblot
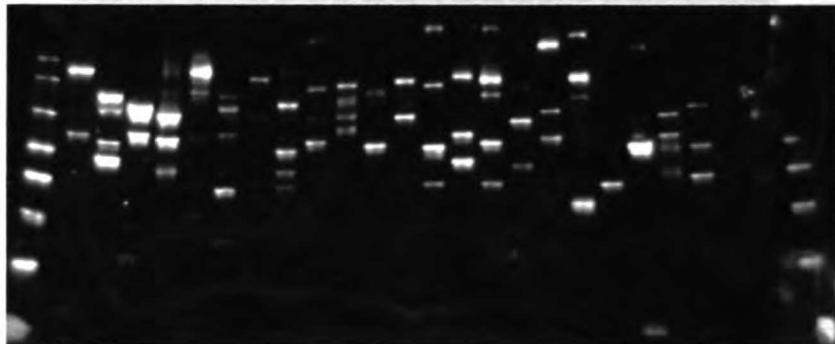
phosphorimage

**Pho80
2A3**

immunoblot
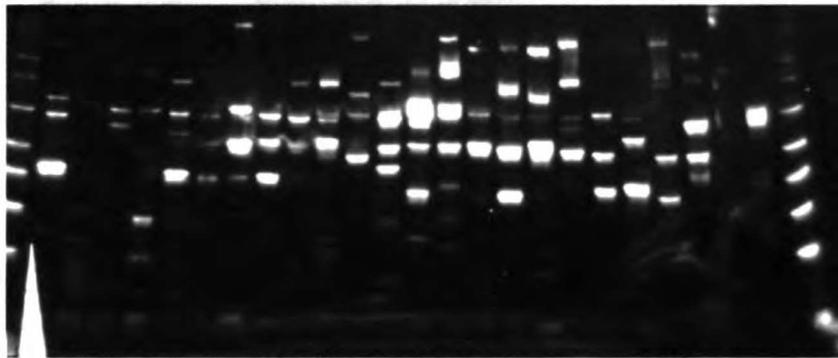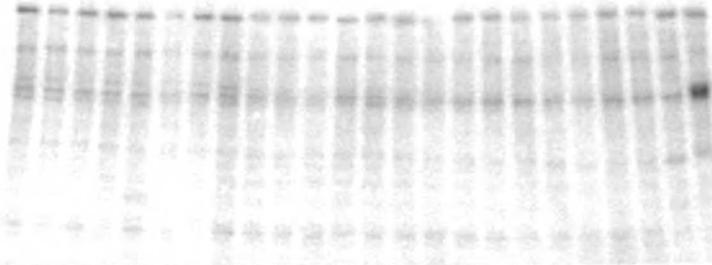
phosphorimage

**Pho80
2A4**

immunoblot

phosphorimage

**Pho80
2B1**

immunoblot

phosphorimage

**Pho80
2B2**

immunoblot

Pho80
2B3

phosphorimage

immunoblot

Pho80
2B4

phosphorimage

immunoblot
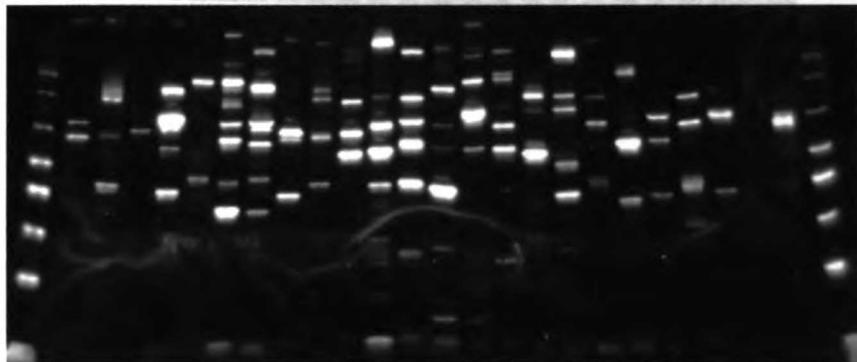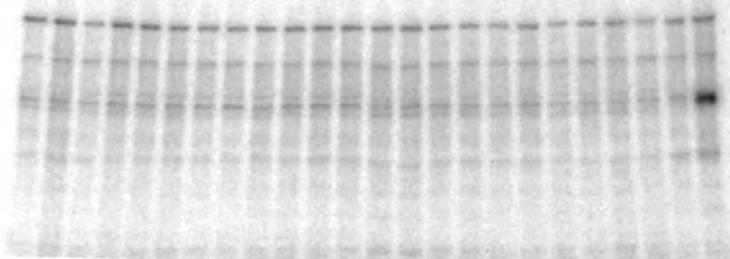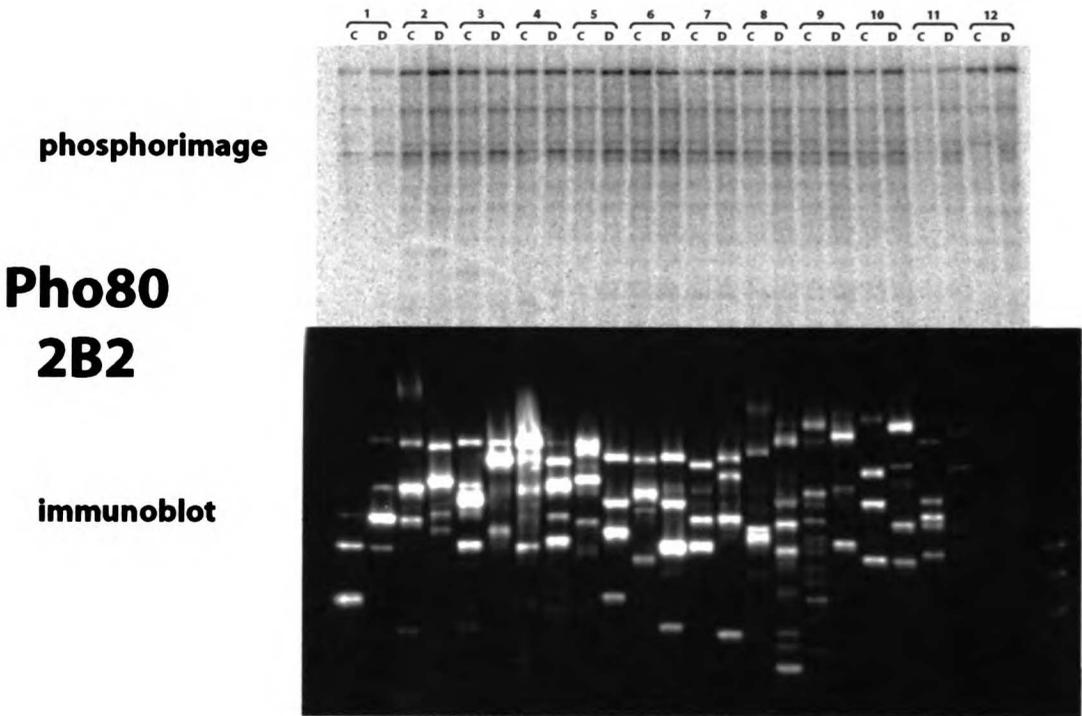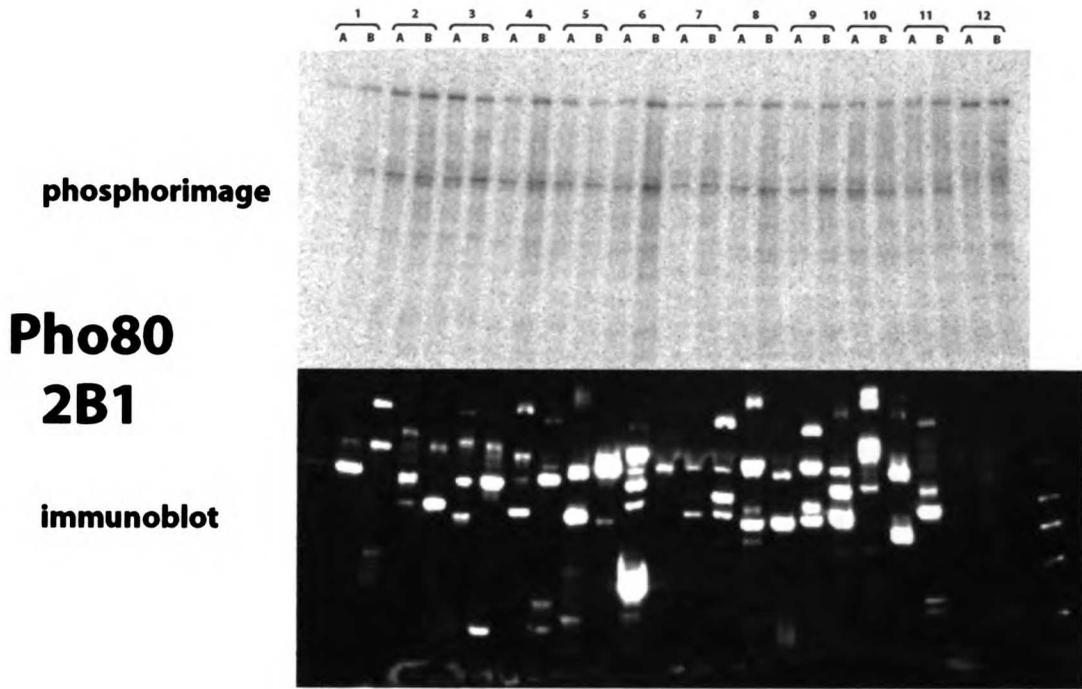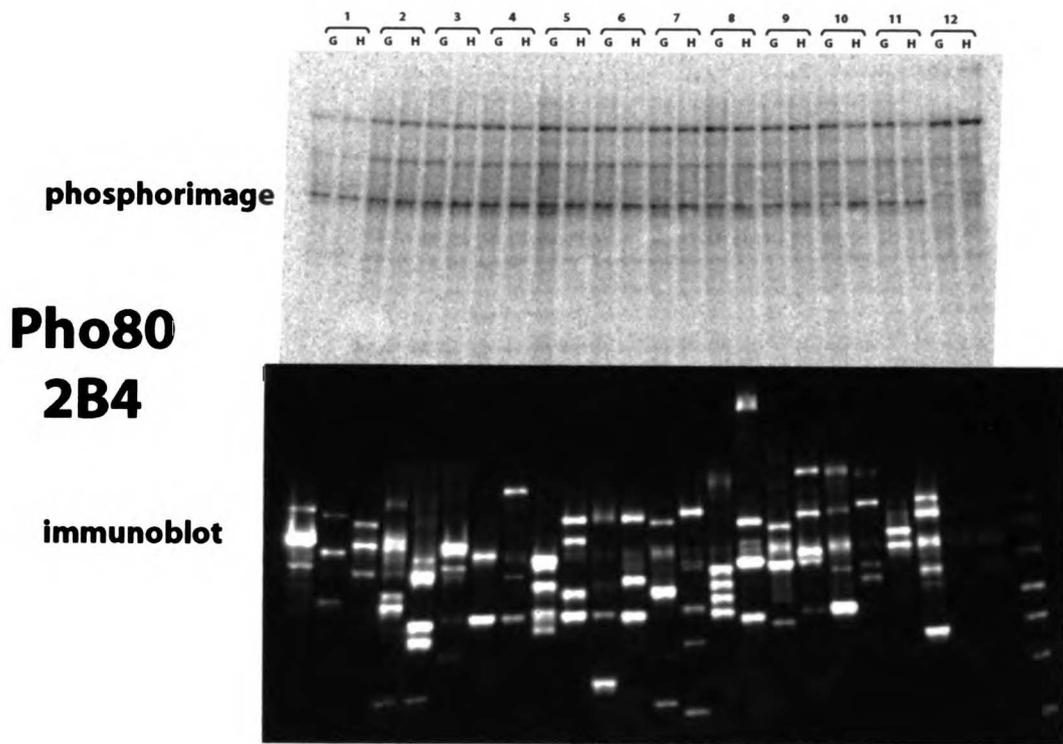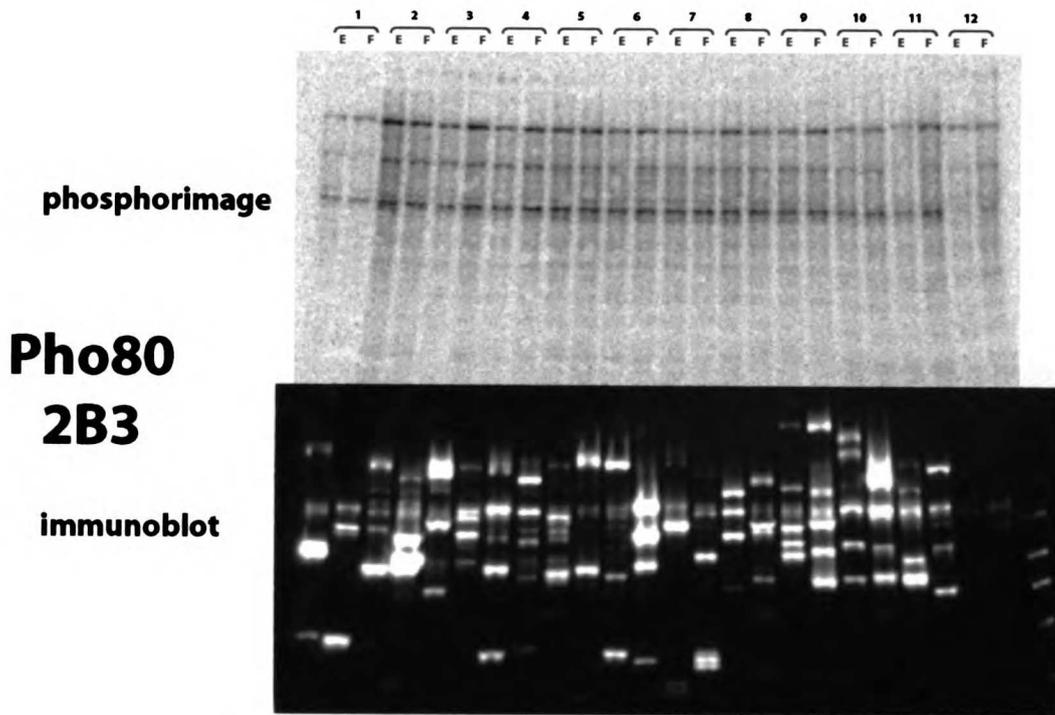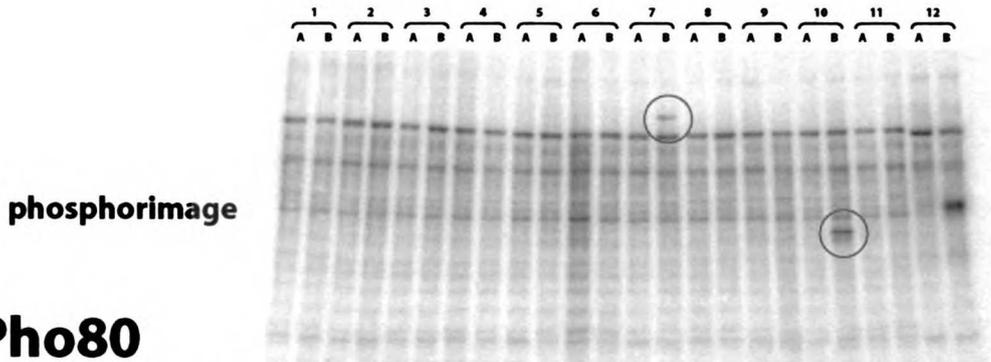
phosphorimage

**Pho80
3A1**

immunoblot
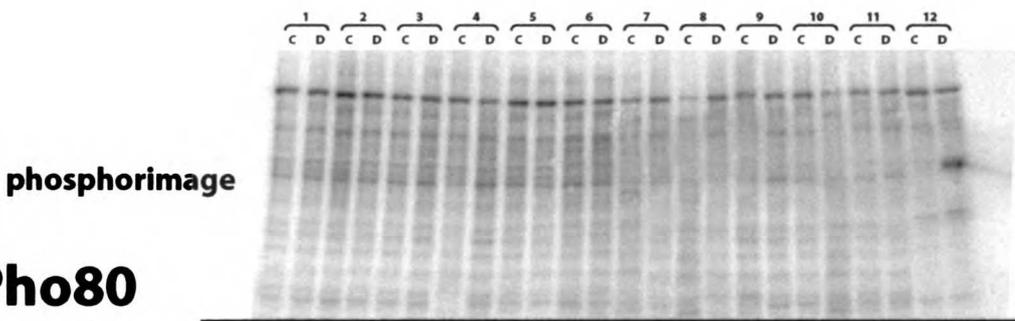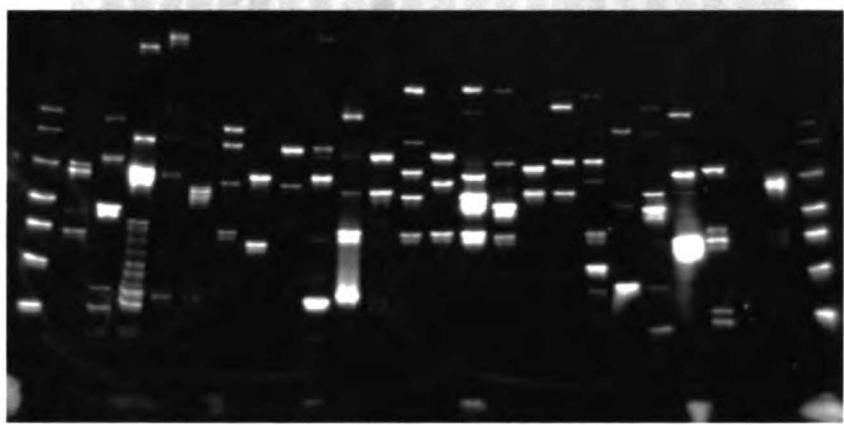
phosphorimage

**Pho80
3A2**

immunoblot
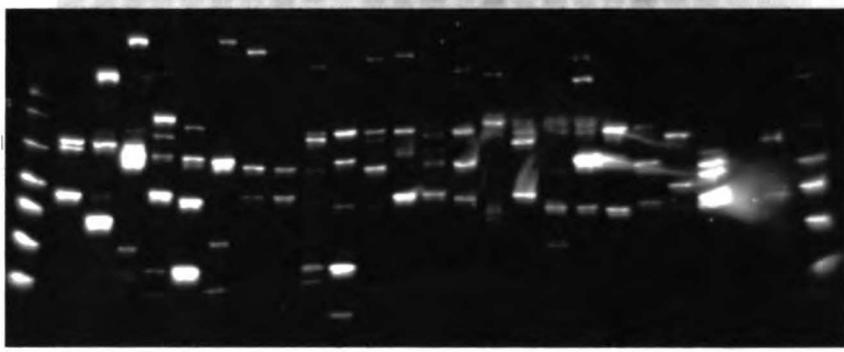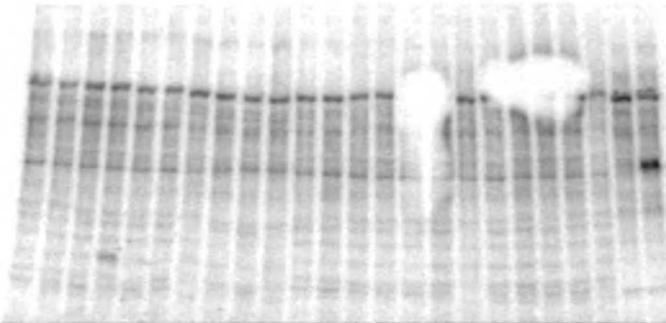
Pho80
3A3

phosphorimage

immunoblot

Pho80
3A4

phosphorimage

immunoblot

Pho80 3B1

phosphorimage

immunoblot

Pho80 3B2

phosphorimage

immunoblot

Pho80
3B3

phosphorimage

immunoblot

Pho80
3B4

phosphorimage

immunoblot

Pho80
4A1

phosphorimage

immunoblot

Pho80
4A2

phosphorimage

immunoblot

phosphorimage

Pho80
4A3

immunoblot

phosphorimage

Pho80
4A4

immunoblot

117
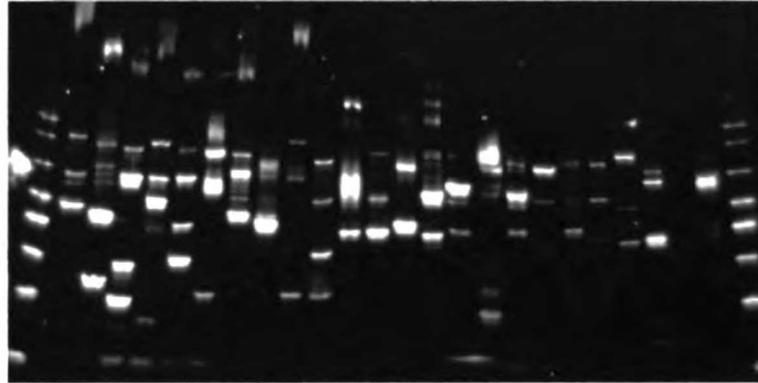
phosphorimage

Pcl1 2A1

immunoblot
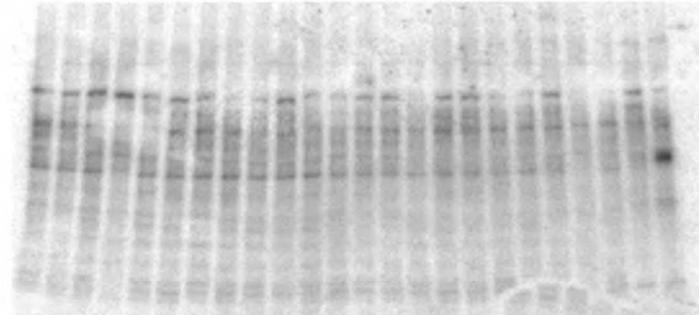
phosphorimage

Pcl1 2A2
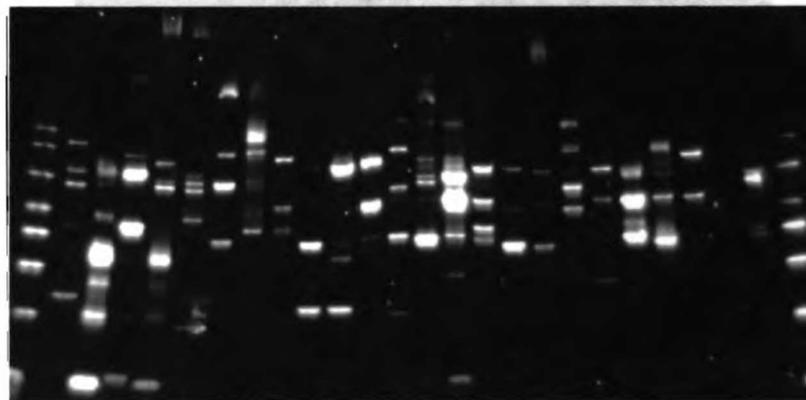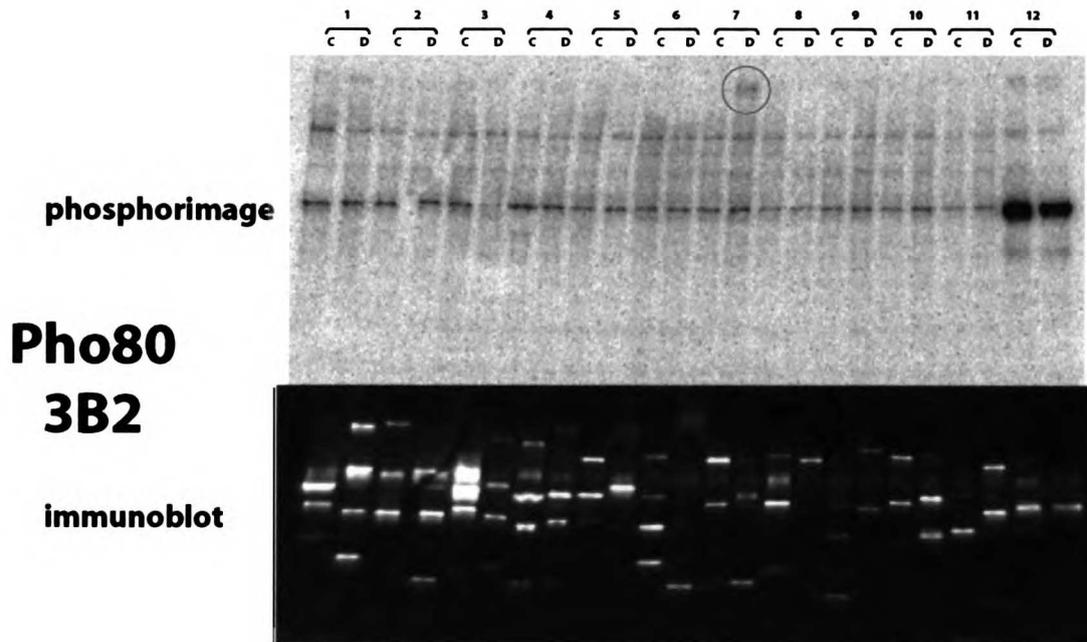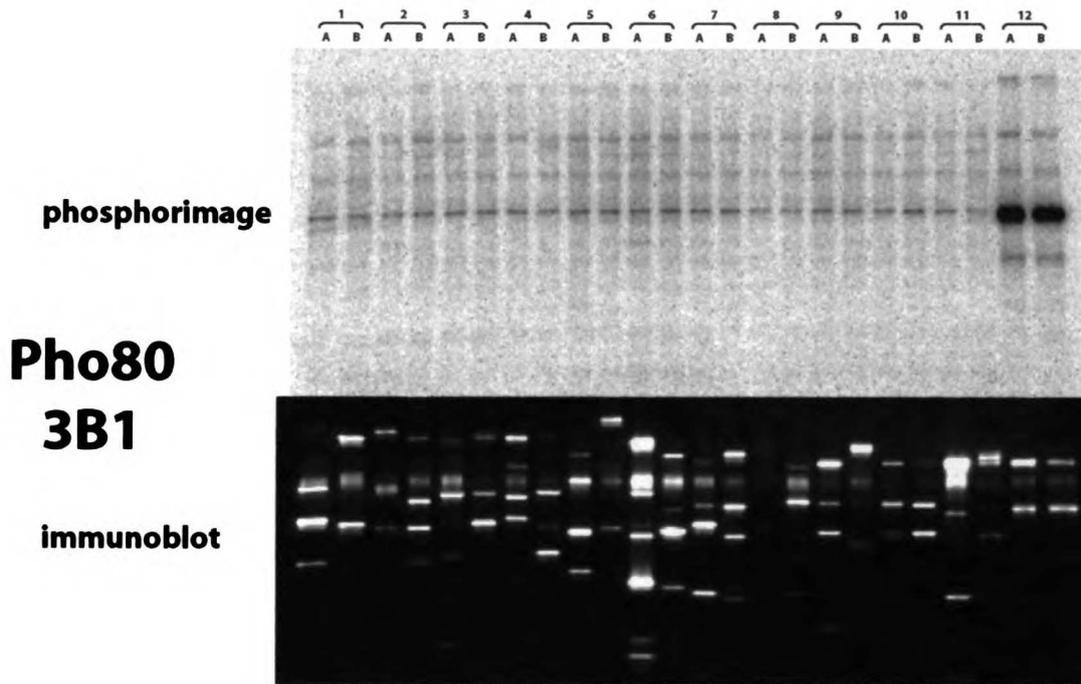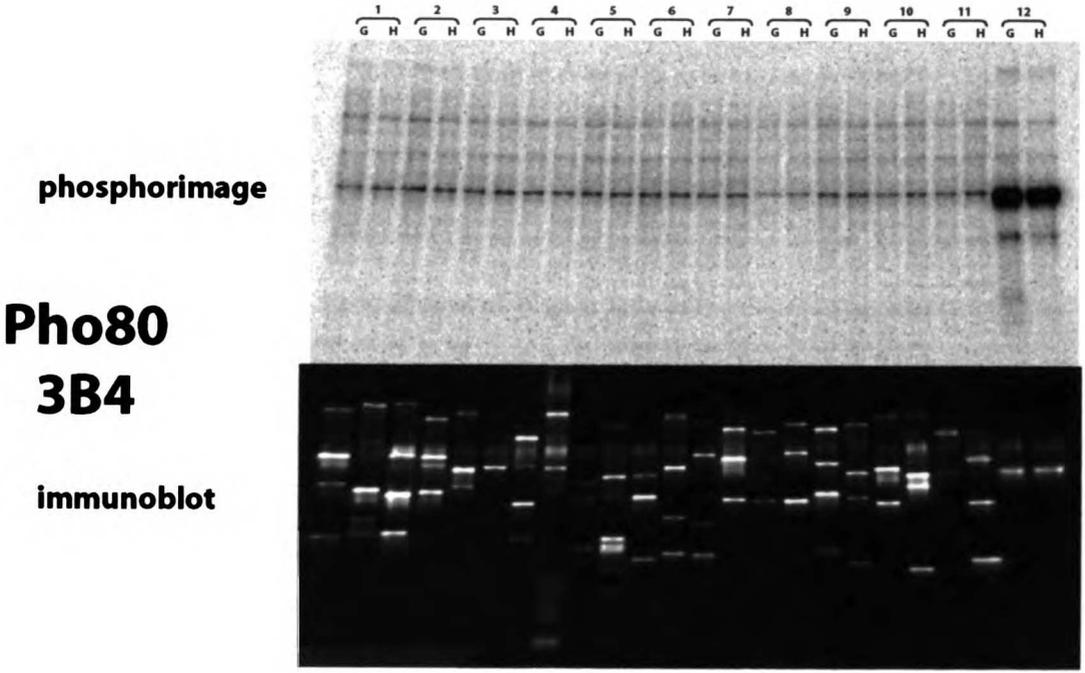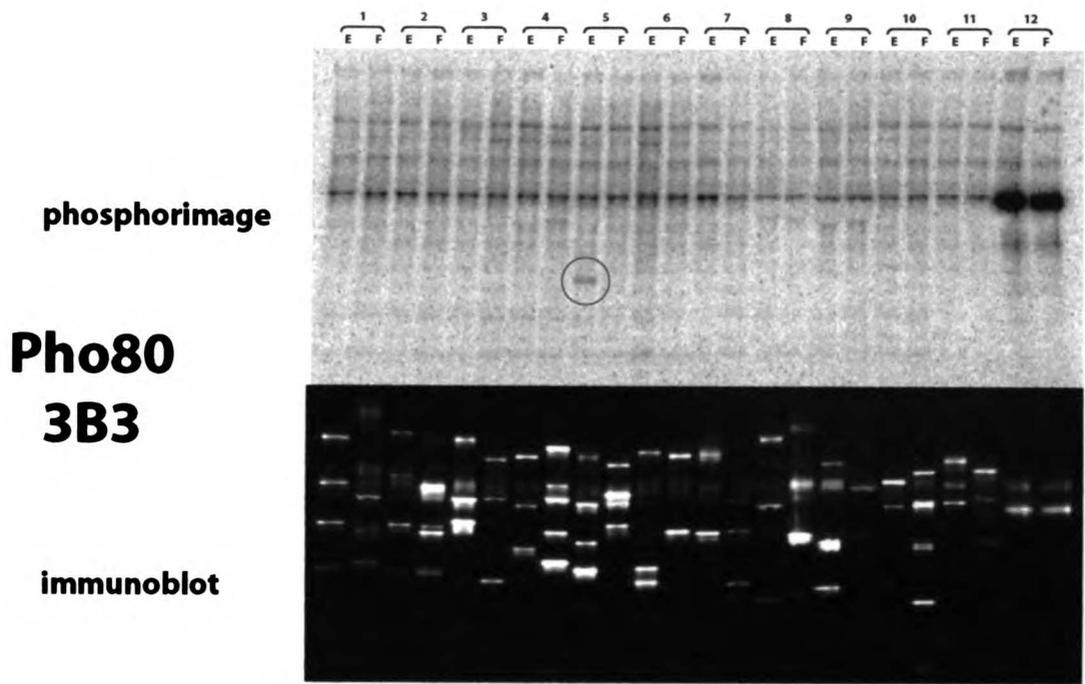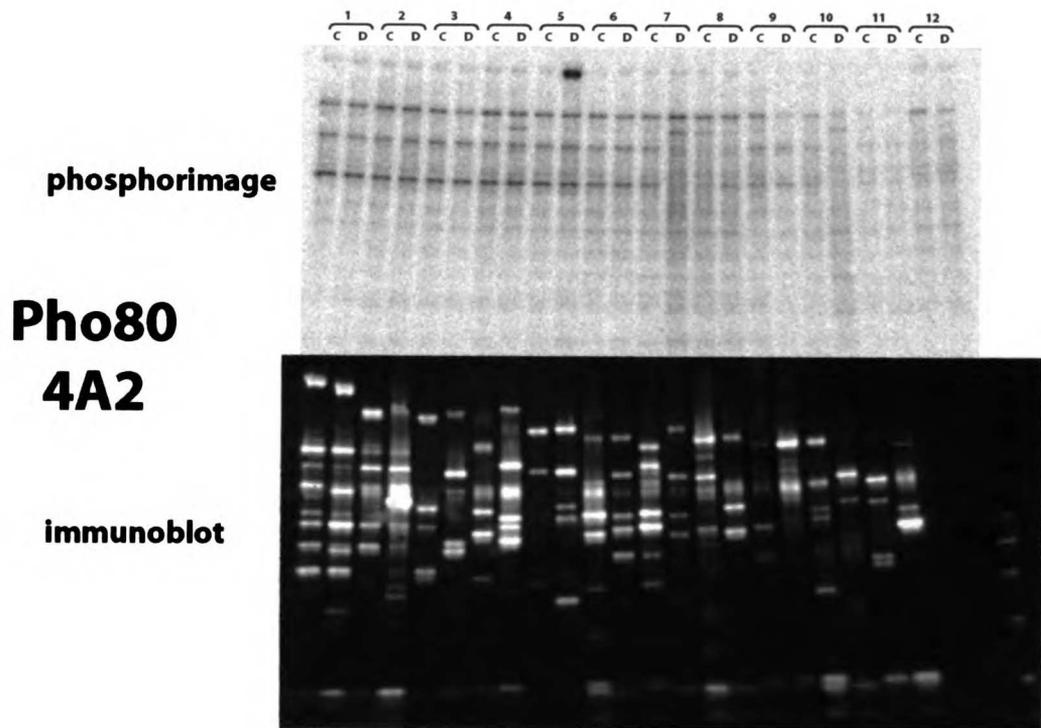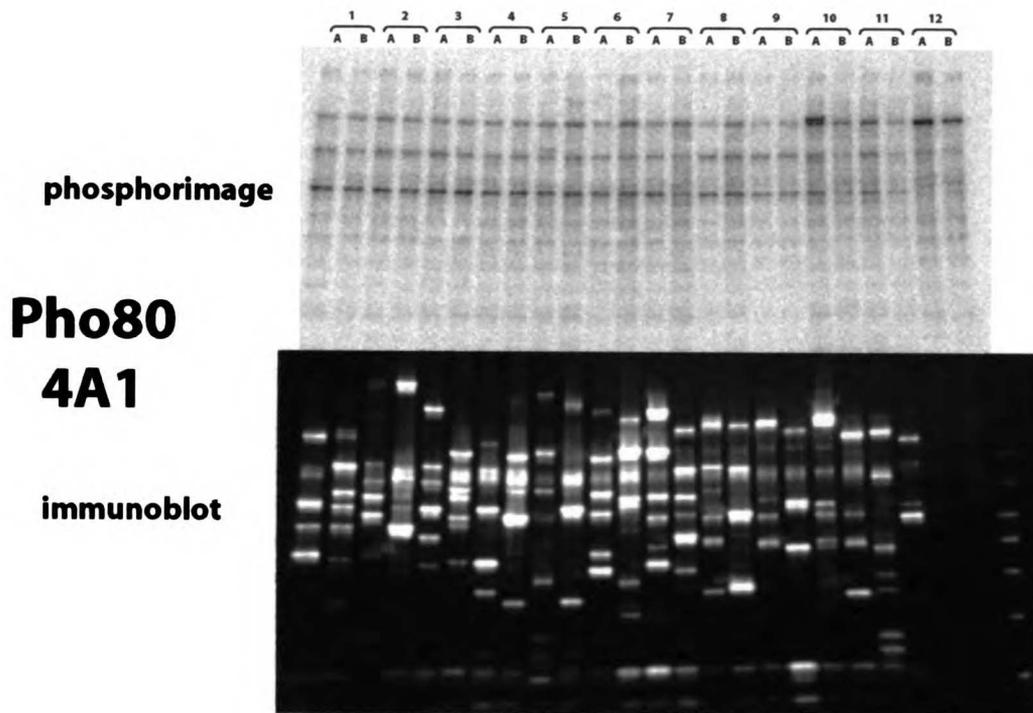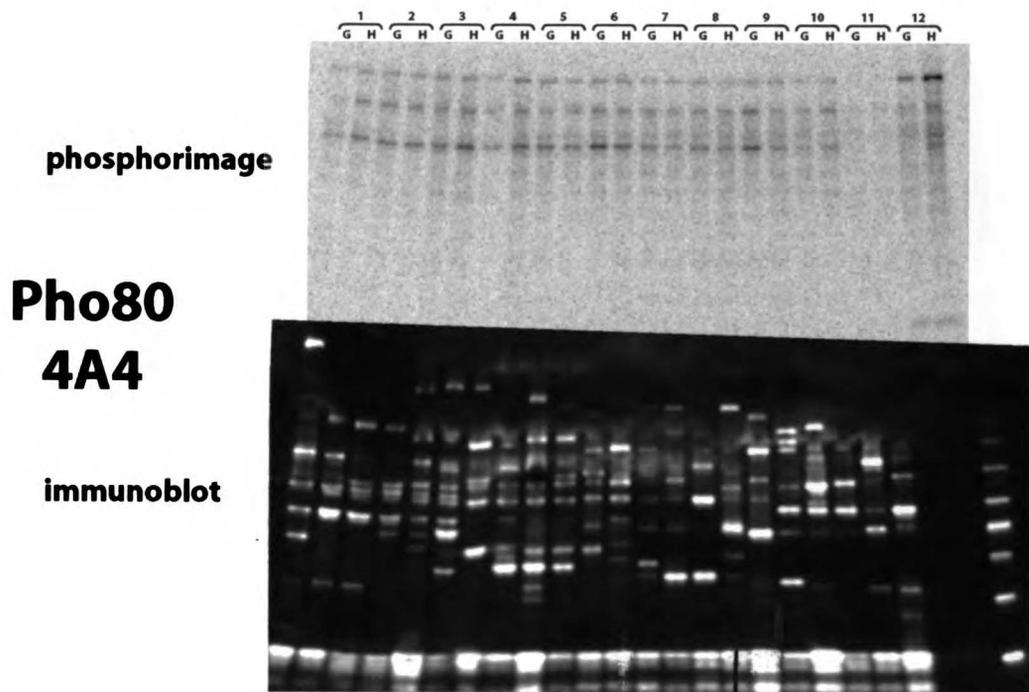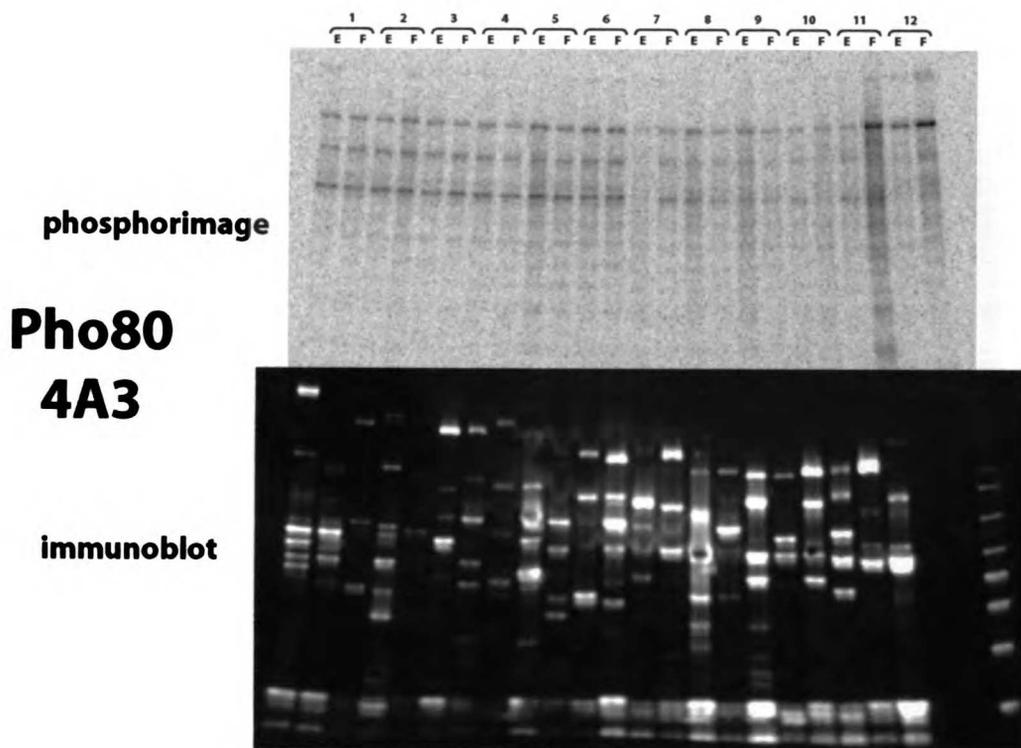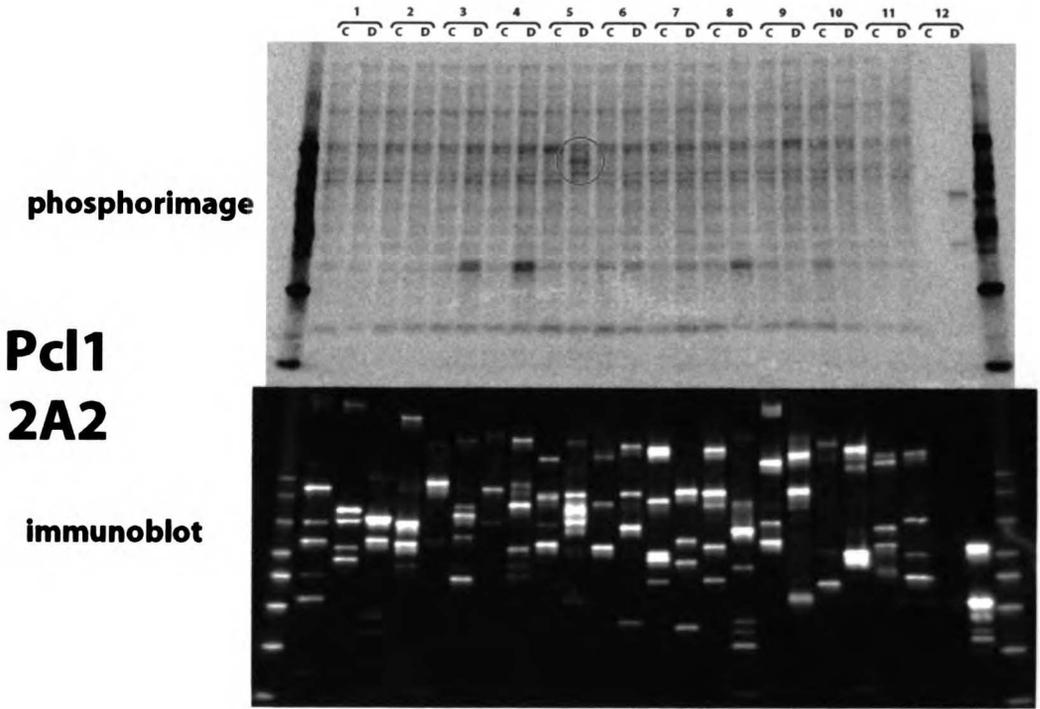
immunoblot

phosphorimage

Pcl1
2A3

immunoblot

phosphorimage

Pcl1
2A4

immunoblot

119

Pcl1 3A1

phosphorimage

immunoblot

Pcl1 3A2

phosphorimage

immunoblot

phosphorimage

Pcl1
3A3

immunoblot

phosphorimage

Pcl1
3A4

immunoblot

phosphorimage

**Pcl1
3B2A**

immunoblot

phosphorimage

**Pcl1
3B2B**

immunoblot

phosphorimage
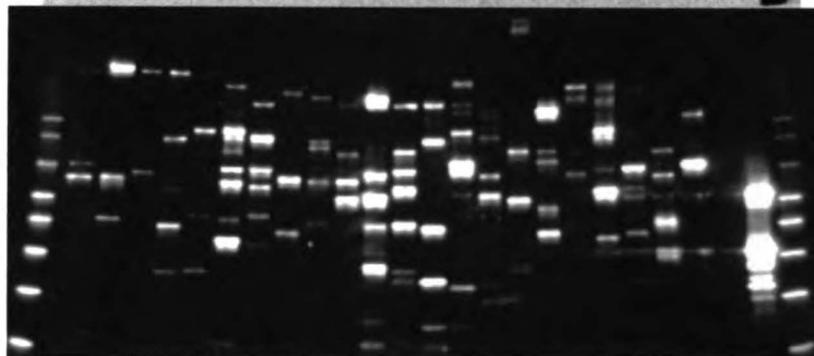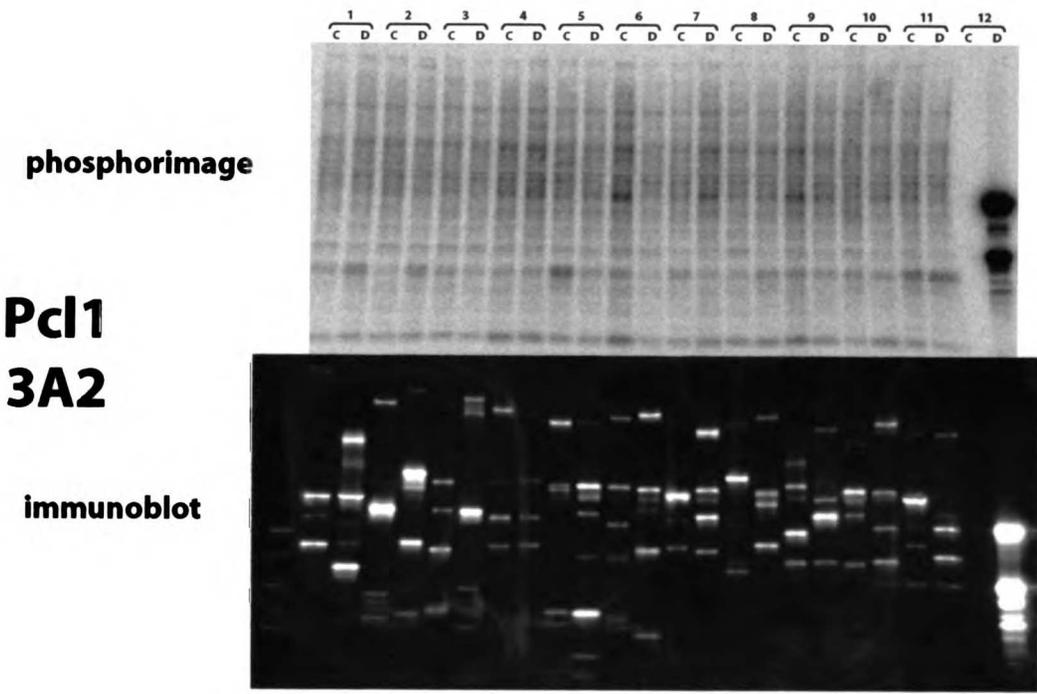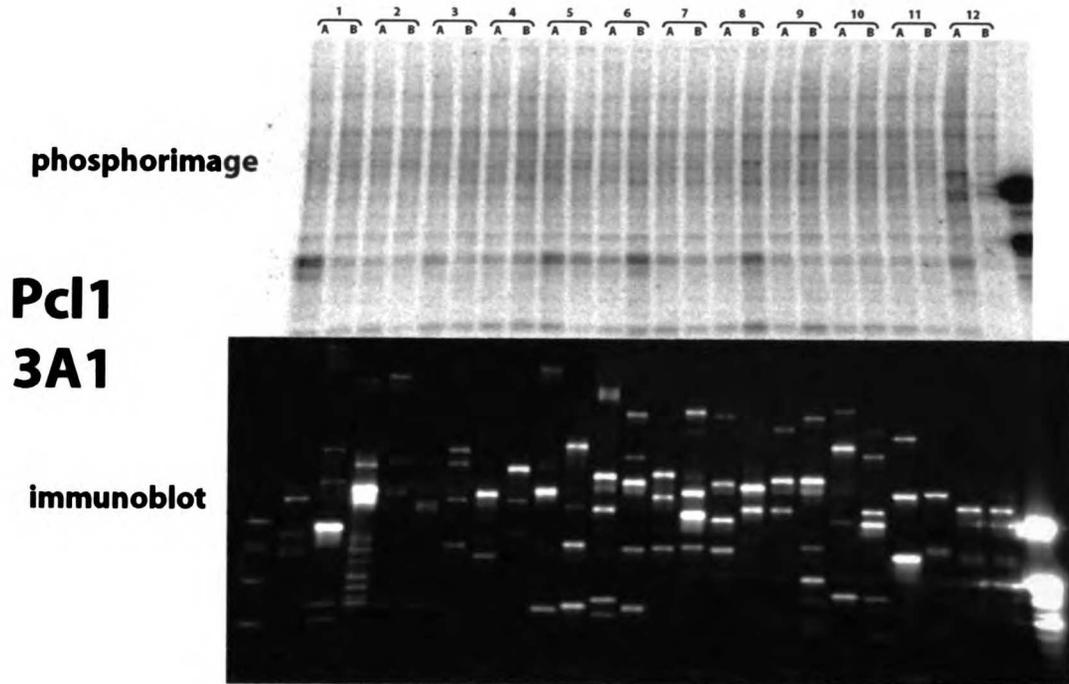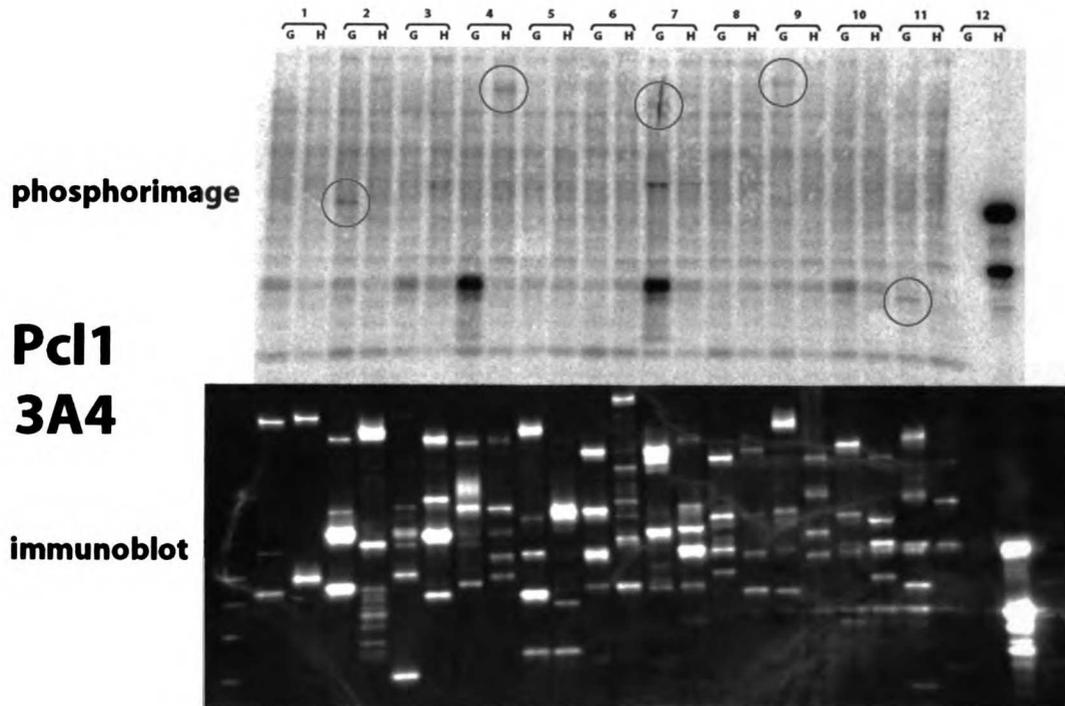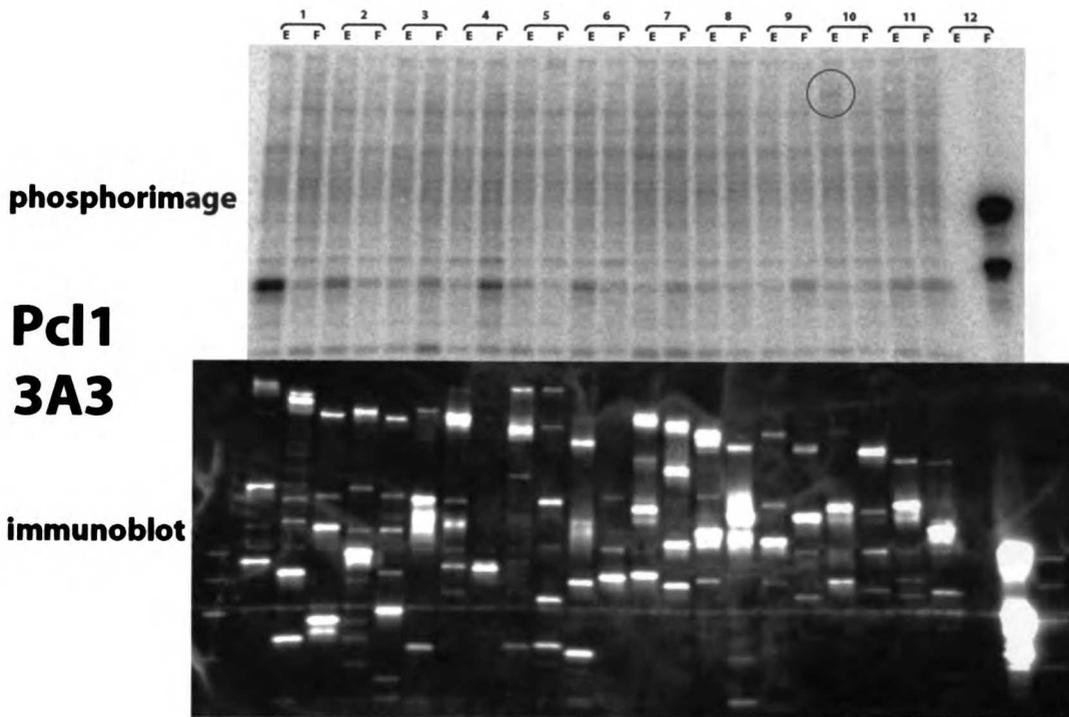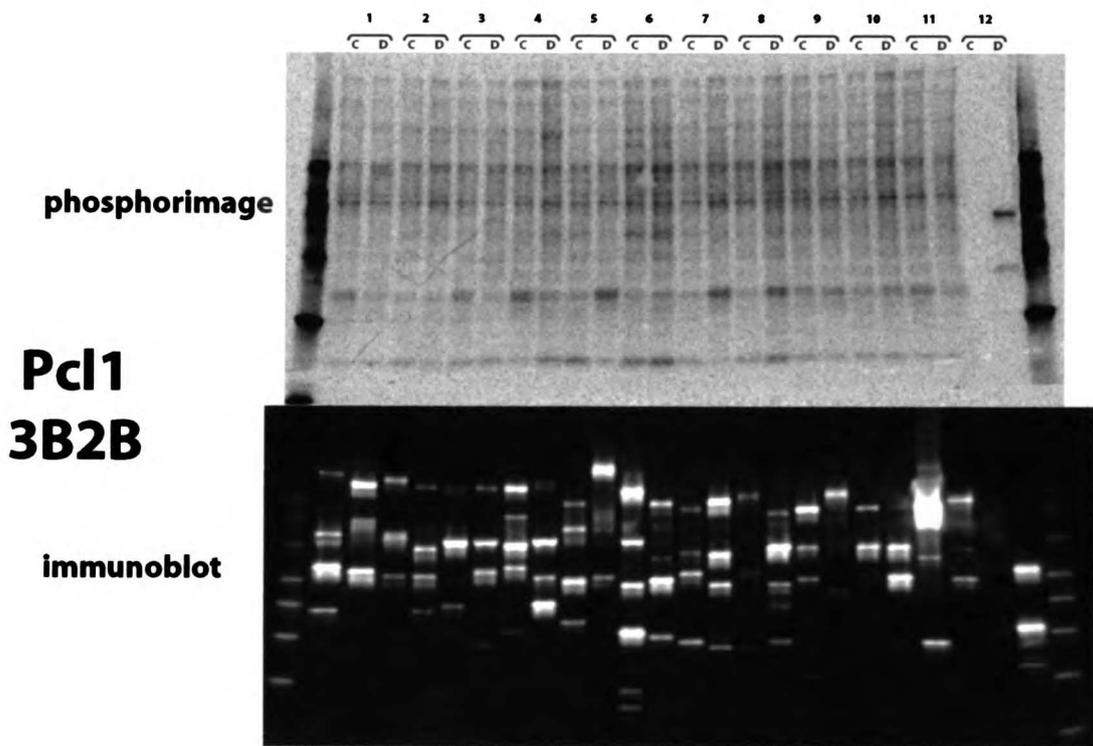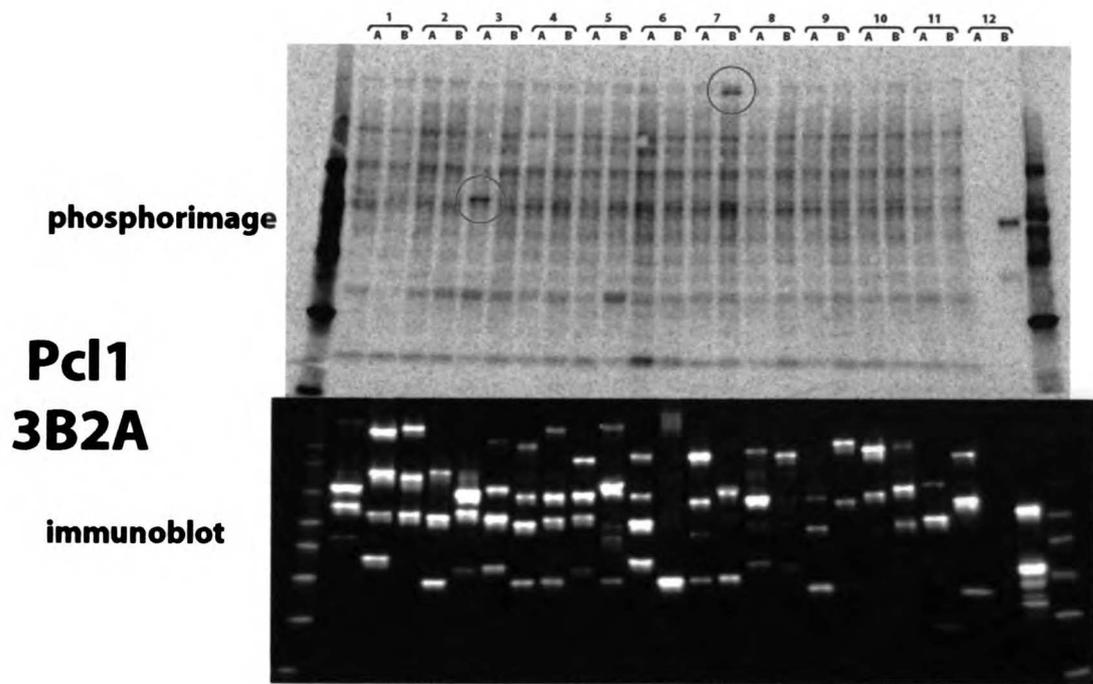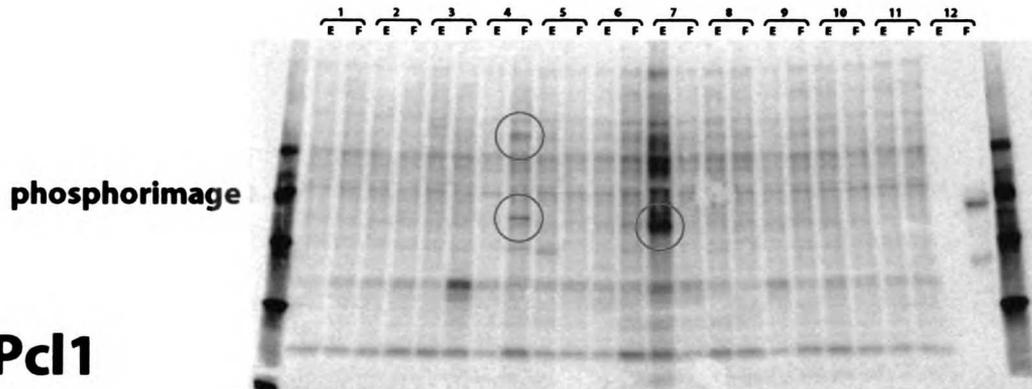
Pcl1 3B3
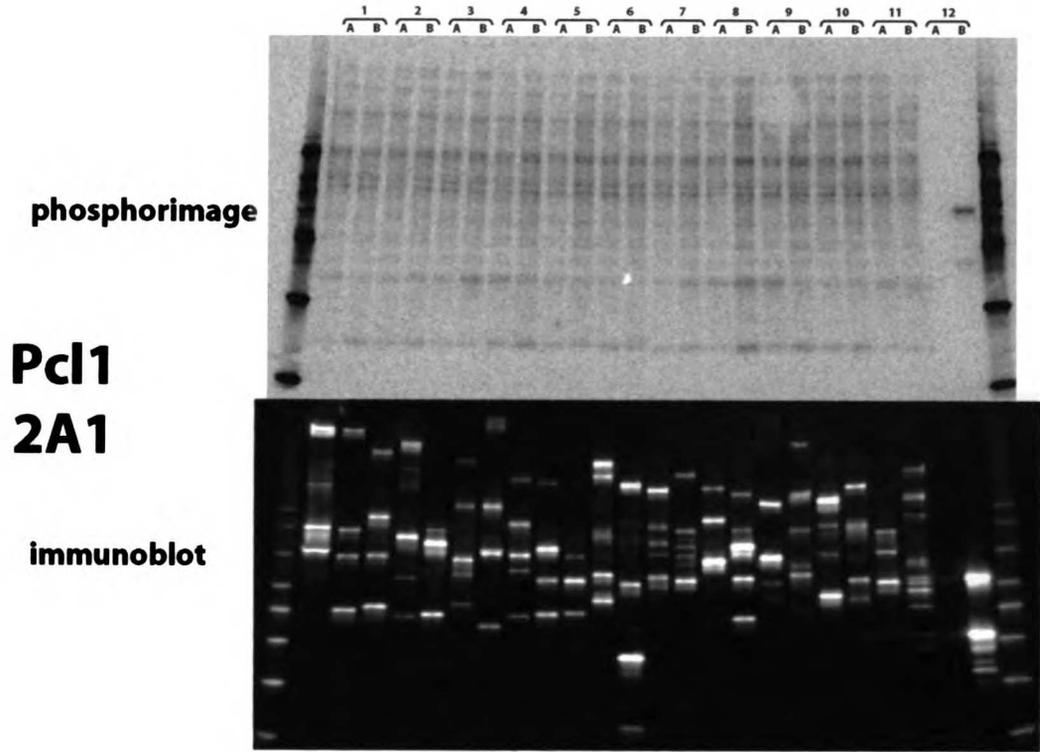
immunoblot

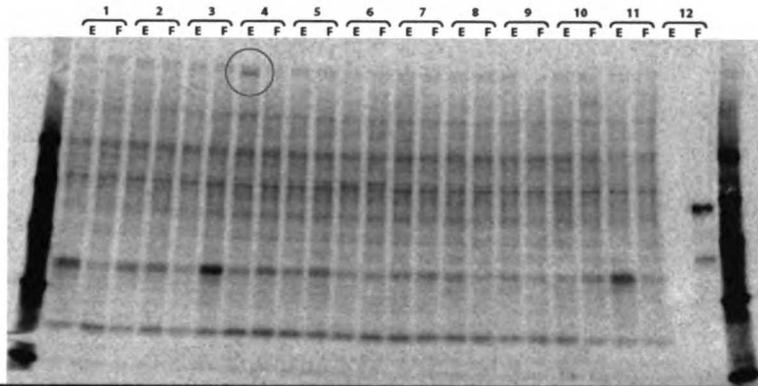phosphorimage

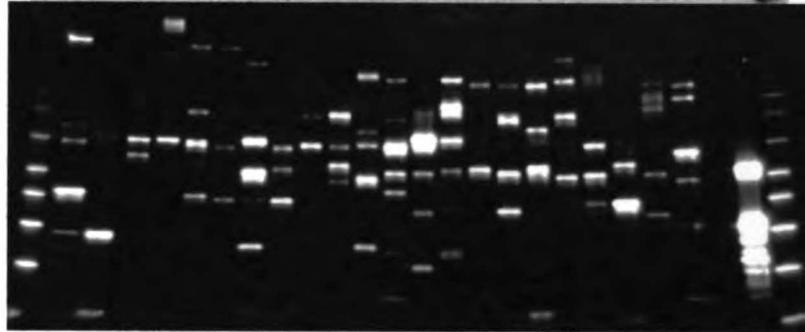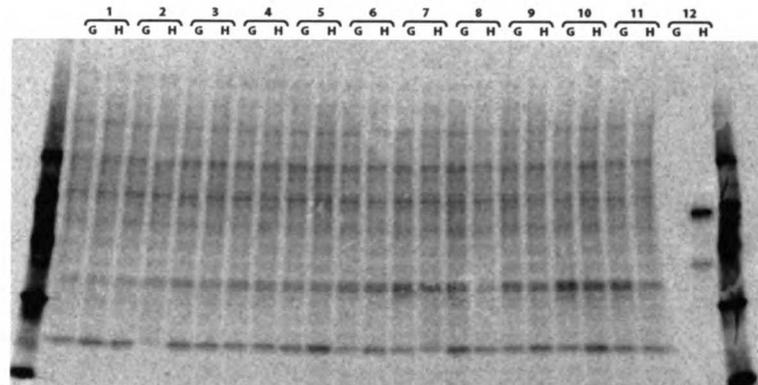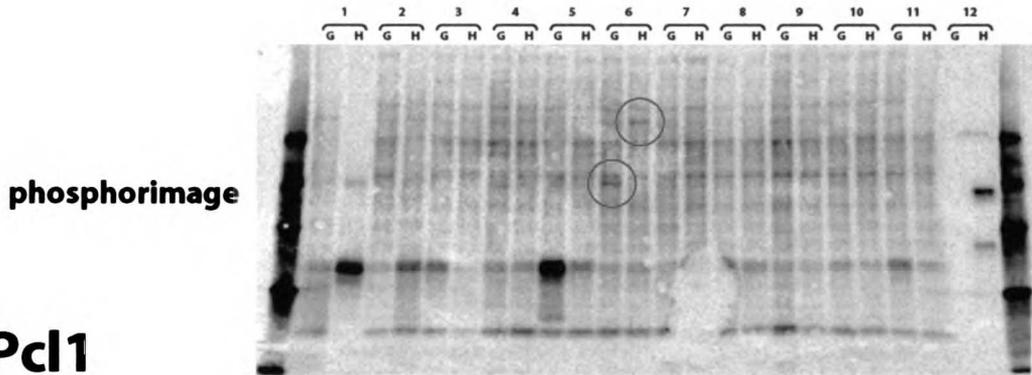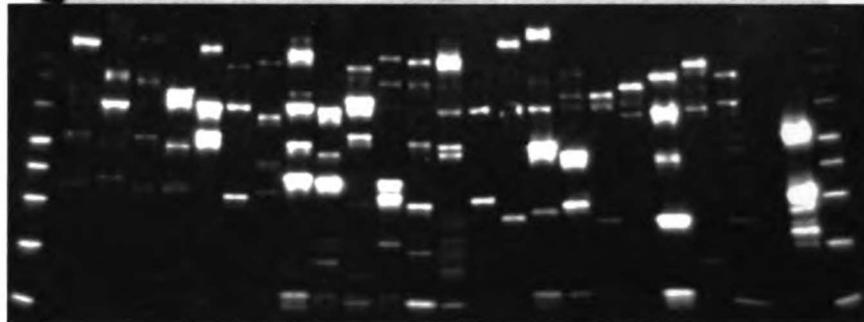**Pcl1 3B4**

immunoblot

phosphorimage

**Pcl1
4A1**

immunoblot

phosphorimage

**Pcl1
4A2**

immunoblot

Pcl1 4A3

phosphorimage

immunoblot

Pcl1 4A4

phosphorimage
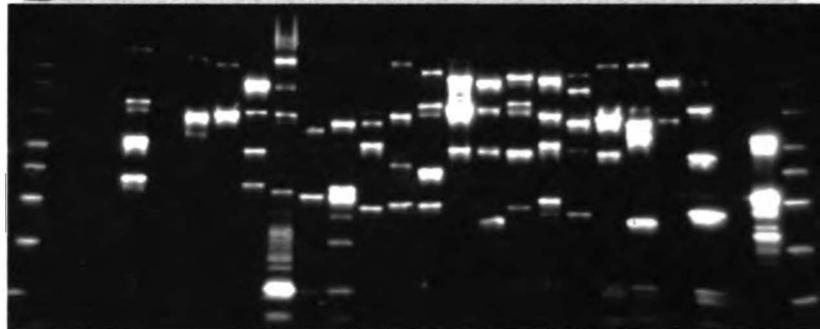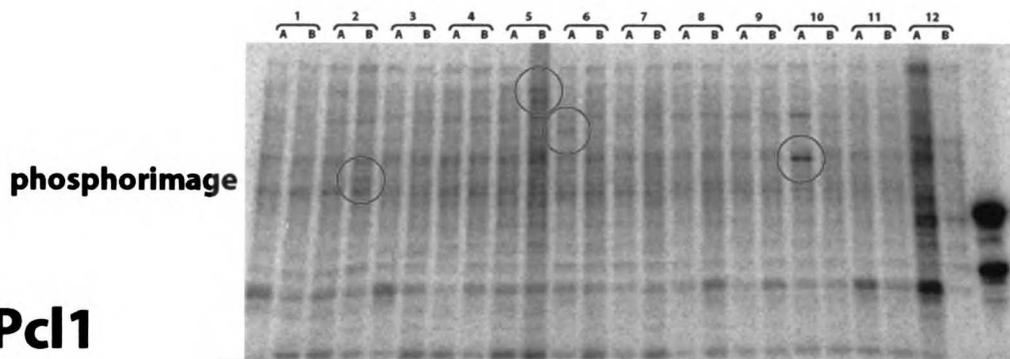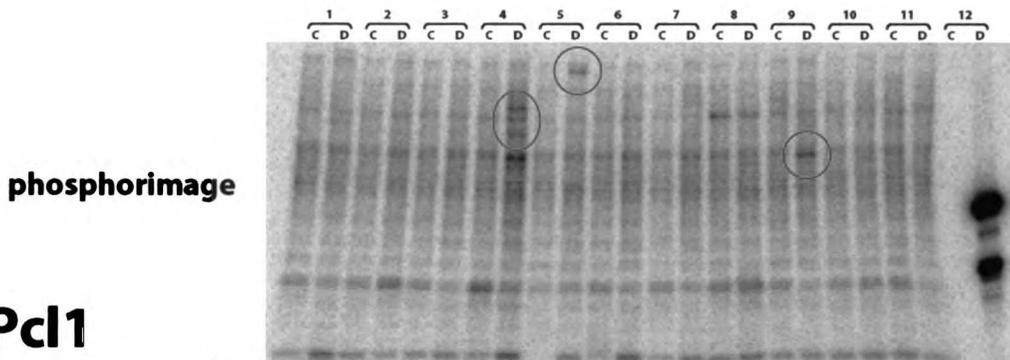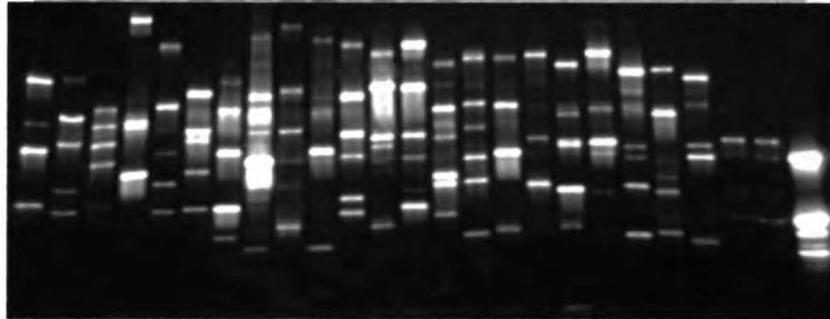
immunoblot

Table 1: Possible identities of Pcl1/Pho85 substrate individual strains. Each

positive from the Pcl1/Pho85 screen is detailed, as well as the possible individual

strains involved. The source plate and position are noted, as well as the plate

and coordinates after rearrangement for further testing. Those strains which

retested are shown in bold blue type with a green background.

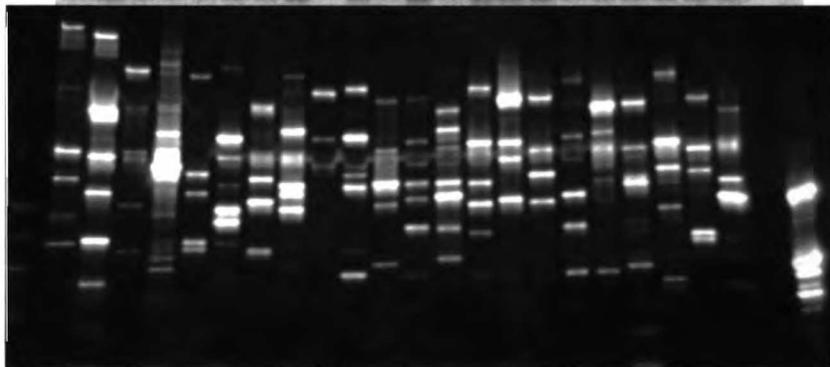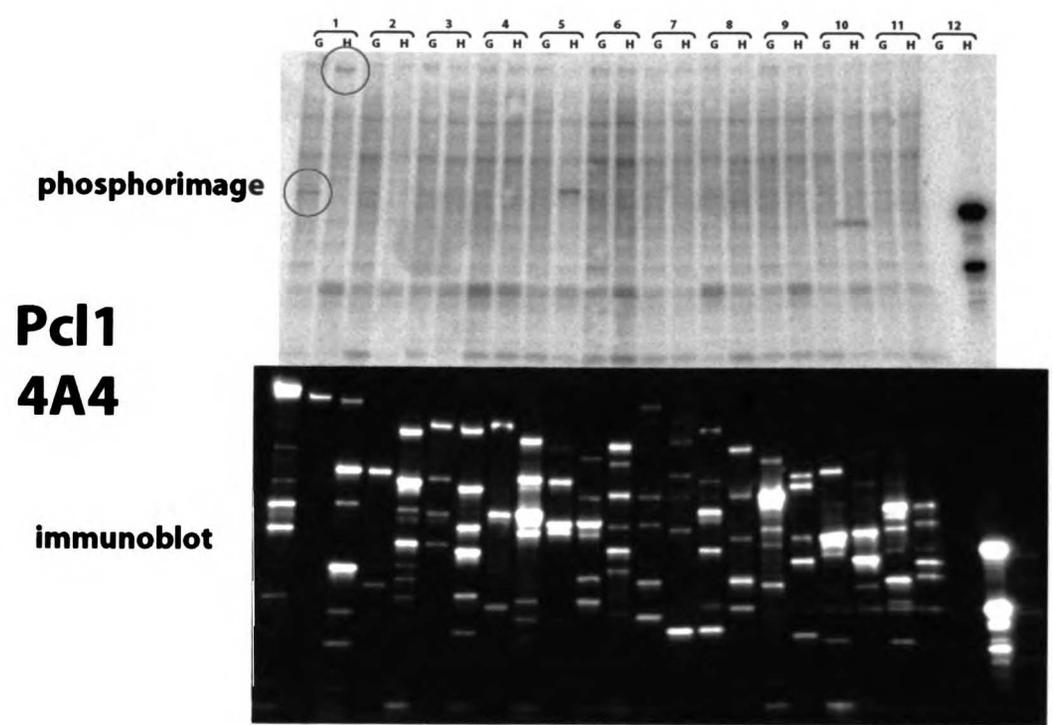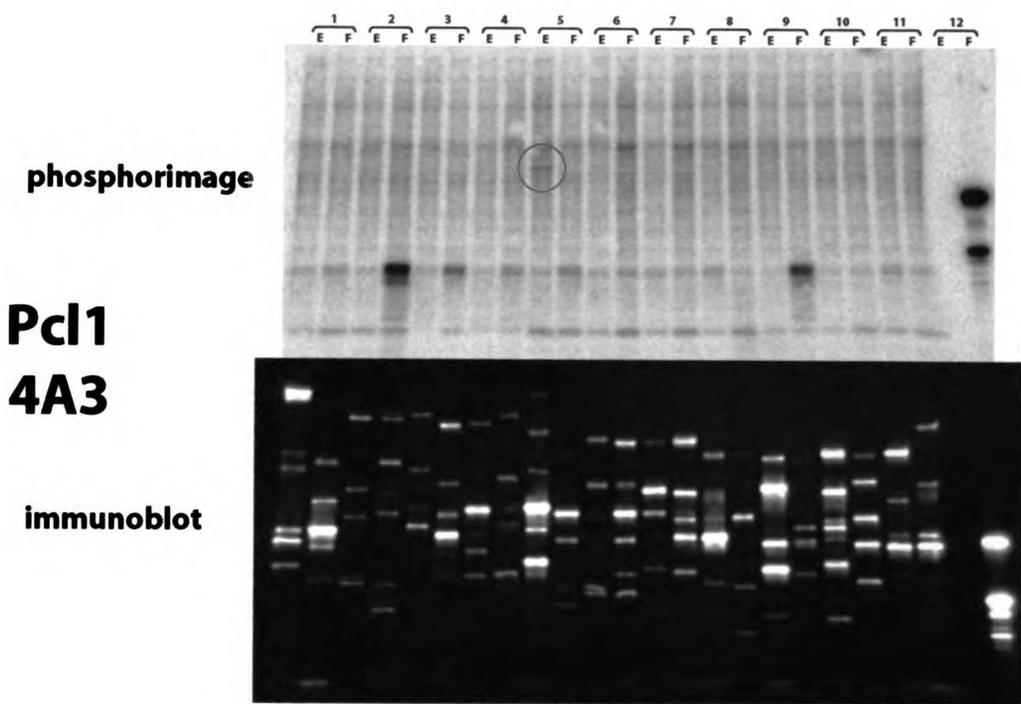| Screen Hit				Possible Positives				
Plate	Gel	Lane	Pos.	Source Location		ORF	Gene	New Location
2A	2	10	D5	01DB2	D05	YKL038W	RGT1	PHA A01
				03DB2	D05	YDR295C		PHA B01
				05DB2	**D05**	**YPL204W**	**HRR25**	**PHA A02**
				07DB2	D05	YHR040W		PHA B02
				09DB2	D05	YDR162C	NBP2	PHA A03
				11DB2	D05	YOR189W		PHA B03
2A	3	7	E4	01DB2	E04	YPL110C		PHA A04
				03DB2	E04	YPR007C		PHA B04
				05DB2	E04	YBR056W		PHA A05
				07DB2	E04	YIL094C	LYS12	PHA B05
				09DB2	E04	YNL129W		PHA A06
				11DB2	**E04**	**YOL005C**	**RPB11**	**PHA B06**
3A	3	19	E10	**01DB3**	**E10**	**YCL024W**	**KCC4**	**PHA A07**
				03DB3	E10	YDR479C		PHA B07
				05DB3	E10	YGL126W	SCS3	PHA A08
				07DB3	E10	YDL213C		PHA B08
				09DB3	E10	YLR396C	VPS33	PHA A09
3A	4	3	G2	01DB3	G02	YPR184W		PHA B09
				03DB3	G02	YJR130C		PHA A10
				05DB3	G02	YGR285C	ZUO1	PHA B10
				07DB3	G02	YKR043C		PHA A11
				09DB3	G02	YJL156W-A		PHA B11
3A	4	8	H4	**01DB3**	**H04**	**YLR371W**	**ROM2**	**PHA C01**
				03DB3	H04	YKR084C	HBS1	PHA D01
				05DB3	H04	YOR245C		PHA C02
				07DB3	H04	YCL005W		PHA D02
				09DB3	H04	YBL079W	NUP170	PHA C03
3A	4	13	G7	**01DB3**	**G07**	**YLR410W**	**VIP1**	**PHA D03**
				03DB3	G07	YMR282C	AEP2	PHA C04
				05DB3	G07	YJL164C	SRA3	PHA D04
				07DB3	G07	YLR301W		PHA C05
				09DB3	G07	YBR229C	ROT2	PHA D05
3A	4	17	G9	**01DB3**	**G09**	**YJL098W**	**SAP185**	**PHA C06**
				03DB3	G09	YDL042C	SIR2	PHA D06
				05DB3	G09	YHR129C	ARP1	PHA C07
				07DB3	G09	YER004W		PHA D07
				09DB3	G09	YBL011W	SCT1	PHA C08
3A	4	21	G11	01DB3	G11	YNL118C	PSU1	PHA D08
				03DB3	G11	YML115C	VAN1	PHA C09
				05DB3	G11	YIL035C	CKA1	PHA D09
				07DB3	G11	YDR487C	RIB3	PHA C10
				09DB3	G11	YKL217W	JEN1	PHA D10
3B	2A	5	A3/C3	02DB3	A03	YJL070C		PHA C11
				04DB3	A03	YMR300C	ADE4	PHA D11
				06DB3	A03	YOR049C		PHA C12
				08DB3	A03	YGR049W	SCM4	PHA D12
				02DB3	C03	YJL078C	PRY3	PHA E01
				04DB3	C03	YGR155W	CYS4	PHA F01

Screen Hit				Possible Positives					
Plate	Gel	Lane	Pos.	Source Location		ORF	Gene	New Location	
				06DB3	C03	YNL231C	PDR16	PHA	E02
				08DB3	C03	YLR065C		PHA	F02
3B	2A	14	B7/D7	02DB3	B07	YHR197W		PHA	E03
				04DB3	B07	YER081W		PHA	F03
				06DB3	B07	YNR034W	SOL1	PHA	E04
				08DB3	B07	YNL147W	LSM7	PHA	F04
				02DB3	D07	YPL158C		PHA	E05
				04DB3	D07	YJL146W	IDS2	PHA	F05
				06DB3	D07	YNL215W		PHA	E06
				08DB3	D07	YOR224C	RPB8	PHA	F06
3B	3	8	F4	**02DB3**	**F04**	**YMR309C**	**NIP1**	**PHA**	**E07**
				04DB3	F04	YDR257C	RMS1	PHA	F07
				06DB3	**F04**	**YOR061W**	**CKA2**	**PHA**	**E08**
				08DB3	F04	YKL056C		PHA	F08
3B	3	13	E7	02DB3	E07	YOR014W	RTS1	PHA	E09
				04DB3	E07	YEL040W	UTR2	PHA	F09
				06DB3	E07	YOR143C	THI80	PHA	E10
				08DB3	E07	YBL025W	RRN10	PHA	F10
3B	4	11	G6	02DB3	G06	YNL224C		PHA	E11
				04DB3	G06	YMR223W	UBP8	PHA	F11
				06DB3	**G06**	**YCR002C**	**CDC10**	**PHA**	**E12**
				08DB3	G06	YGR203W		PHA	F12
3B	4	12	H6	02DB3	H06	YDR490C	PKH1	PHA	G01
				04DB3	H06	YMR239C	RNT1	PHA	H01
				06DB3	H06	YIL043C	CBR1	PHA	G02
				08DB3	H06	YPL144W		PHA	H02
4A	1	4	B2	01DB4	B02	YOR341W	RPA190	PHA	G03
				02DB4	B02	YML034W		PHA	H03
				03DB4	B02	YOR128C	ADE2	PHA	G04
				04DB4	**B02**	**YDR432W**	**NPL3**	**PHA**	**H04**
				05DB4	B02	YKL216W	URA1	PHA	G05
				06DB4	B02	YMR121C	RPL15B	PHA	H05
4A	1	10	B5	**01DB4**	**B05**	**YDR293C**	**SSD1**	**PHA**	**G06**
				02DB4	B05	YGL205W	POX1	PHA	H06
				03DB4	B05	YOR310C	NOP58	PHA	G07
				04DB4	B05	YEL046C	GLY1	PHA	H07
				05DB4	B05	YNL162W	RPL42A	PHA	G08
				06DB4	B05	YGL030W	RPL30	PHA	H08
4A	1	11	A6	01DB4	A06	YOR086C		PHA	G09
				02DB4	A06	YLR258W	GSY2	PHA	H09
				03DB4	A06	YBR126C	TPS1	PHA	G10
				04DB4	A06	YLR291C	GCD7	PHA	H10
				05DB4	A06	YFR050C	PRE4	PHA	G11
				06DB4	A06	YKL013C	ARC19	PHA	H11
4A	1	19	A10	**01DB4**	**A10**	**YBR079C**	**RPG1**	**PHB**	**A01**
				02DB4	A10	YAL016W	TPD3	PHB	B01
				03DB4	A10	YER178W	PDA1	PHB	A02
				04DB4	A10	YNR035C	ARC35	PHB	B02

Screen Hit				Possible Positives					
Plate	**Gel**	**Lane**	**Pos.**	**Source Location**		**ORF**	**Gene**	**New Location**	
				05DB4	A10	YDR196C		PHB	A03
				06DB4	A10	YGL189C	RPS26A	PHB	B03
4A	2	8	D4	01DB4	D04	YOR296W		PHB	A04
				02DB4	**D04**	**YOR361C**	**PRT1**	**PHB**	**B04**
				03DB4	D04	YKL081W	TEF4	PHB	A05
				04DB4	D04	YBR025C		PHB	B05
				05DB4	**D04**	**YBR011C**	**IPP1**	**PHB**	**A06**
				06DB4	D04	YDL165W	CDC36	PHB	B06
4A	2	10	D5	**01DB4**	**D05**	**YOR151C**	**RPB2**	**PHB**	**A07**
				02DB4	D05	YIL078W	THS1	PHB	B07
				03DB4	D05	YEL052W	AFG1	PHB	A08
				04DB4	D05	YJR072C		PHB	B08
				05DB4	D05	YIL069C	RPS24B	PHB	A09
				06DB4	D05	YBR082C	UBC4	PHB	B09
4A	2	18	D9	**01DB4**	**D09**	**YGR240C**	**PFK1**	**PHB**	**A10**
				02DB4	D09	YDR398W		PHB	B10
				03DB4	D09	YKR042W	UTH1	PHB	A11
				04DB4	D09	YIL154C	IMP2	PHB	B11
				05DB4	D09	YGR034W	RPL26B	PHB	A12
				06DB4	D09	YKL138C	MRPL31	PHB	B12
4A	3	9	E5	01DB4	E05	YMR109W	MYO5	PHB	C01
				02DB4	E05	YKR018C		PHB	D01
				03DB4	**E05**	**YAL038W**	**CDC19**	**PHB**	**C02**
				04DB4	E05	YDR502C	SAM2	PHB	D02
				05DB4	E05	YIL053W	RHR2	PHB	C03
				06DB4	E05	YPL249C-A	RPL36B	PHB	D03
4A	4	1	G1	01DB4	G01	YPL231W	FAS2	PHB	C04
				02DB4	G01	YOR141C	ARP8	PHB	D04
				03DB4	G01	YGR124W	ASN2	PHB	C05
				04DB4	G01	YIL033C	SRA1	PHB	D05
				05DB4	G01	YBL027W	RPL19B	PHB	C06
				06DB4	G01	YLR185W	RPL37A	PHB	D06
4A	4	2	H1	01DB4	H01	YMR229C	RRP5	PHB	C07
				02DB4	H01	YLR386W		PHB	D07
				03DB4	H01	YHR064C	PDR13	PHB	C08
				04DB4	H01	YNL248C	RPA49	PHB	D08
				05DB4	H01	YMR310C		PHB	C09
				06DB4	H01	YLR287C-A	RPS30A	PHB	D09
4A	4	10	H5	01DB4	H05	YKR096W		PHB	C10
				02DB4	H05	YFR015C	GSY1	PHB	D10
				03DB4	**H05**	**YKL035W**	**UGP1**	**PHB**	**C11**
				04DB4	H05	YBL032W		PHB	D11
				05DB4	H05	YML024W	RPS17A	PHB	C12
				06DB4	**H05**	**YDR404C**	**RPB7**	**PHB**	**D12**

*Figure 5: Deconvolution of Pcl1/Pho85 substrates.* To identify the individual strain associated with each Pcl1/Pho85 substrate hit, we picked all possible strains and reorganized them into 96 well format (detailed in Table 1). The in extract assay was performed on each individual strain, and results analyzed by immunoblot and phosphorimaging. Phosphoproteins which retested are circled in red.
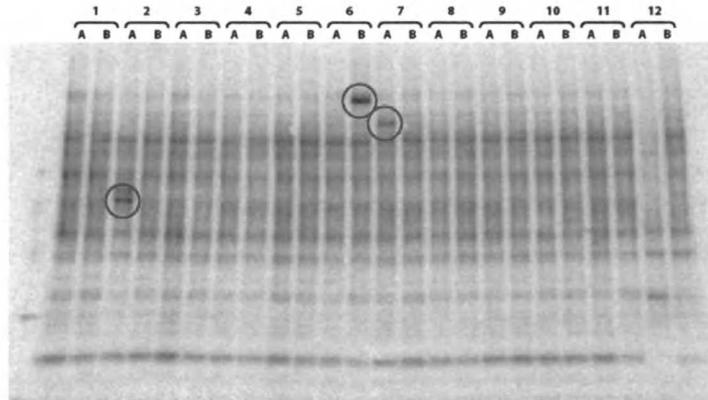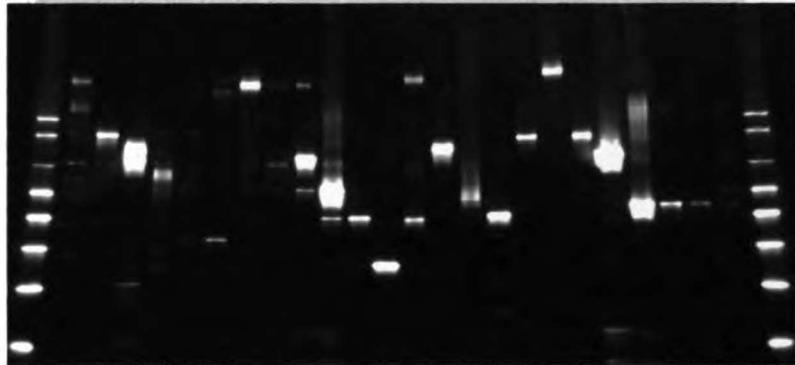
PHA1

PHA2

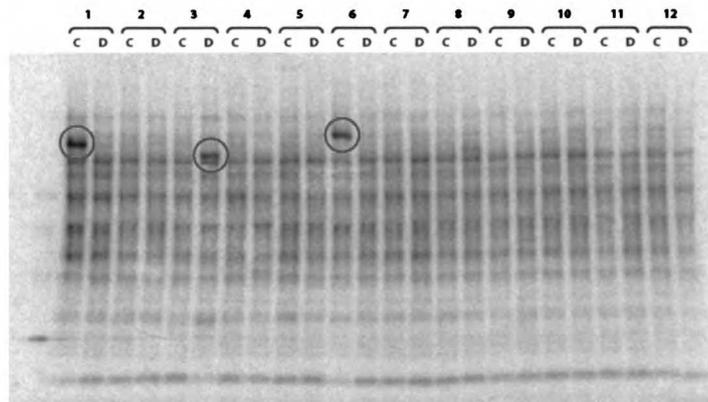PHA3

phosphorimage

immunoblot

PHA4

phosphorimage
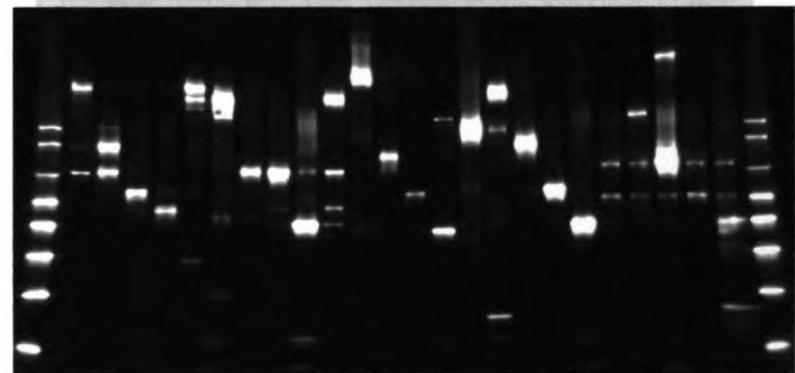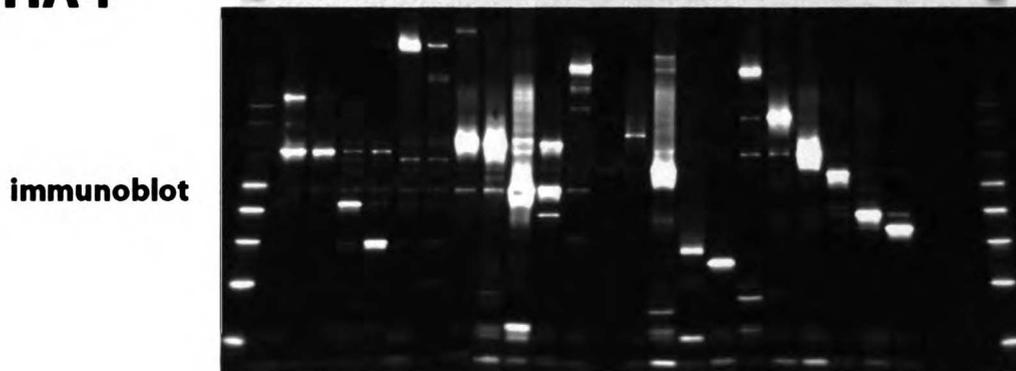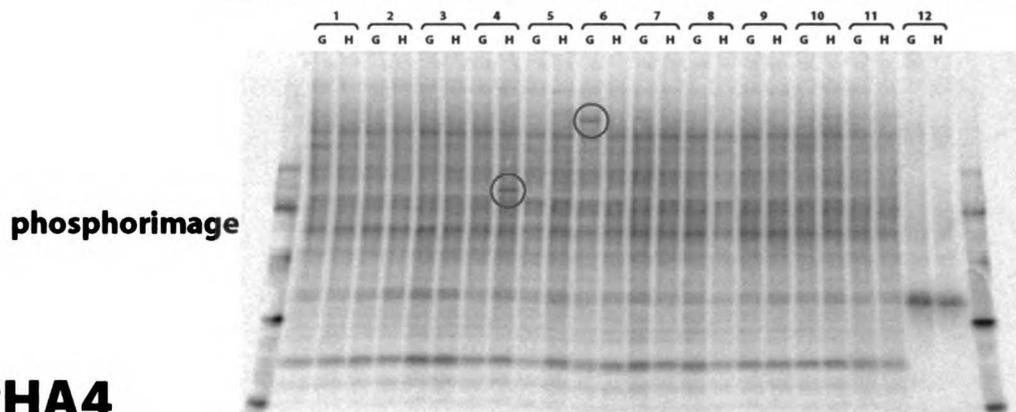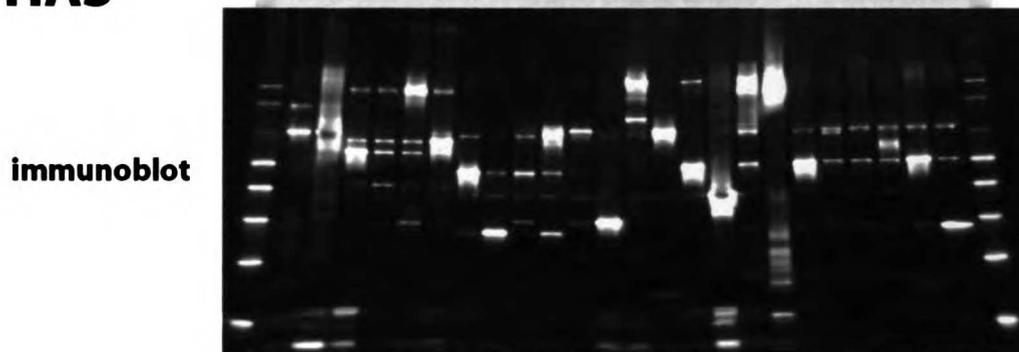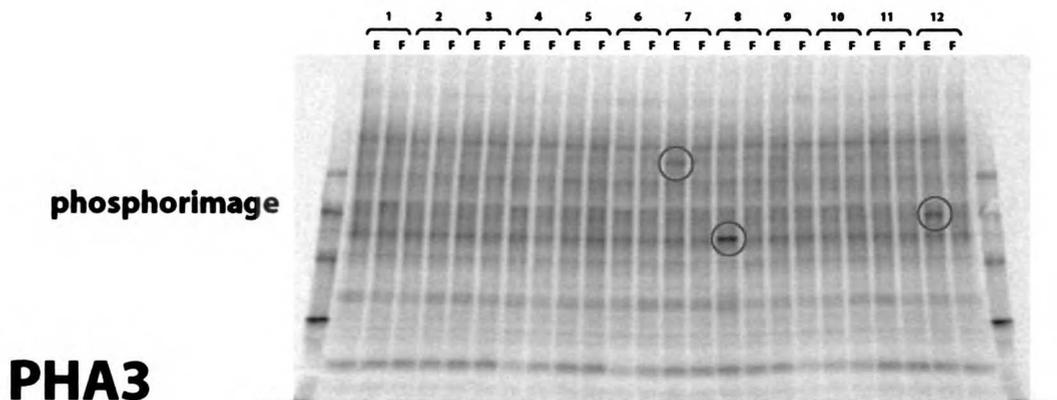
immunoblot
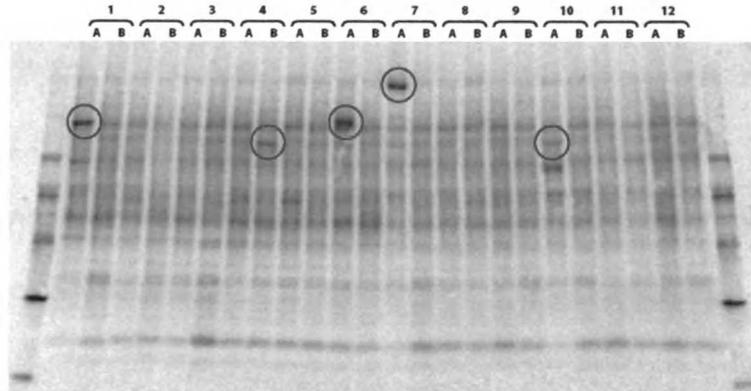
PHB1

phosphorimage
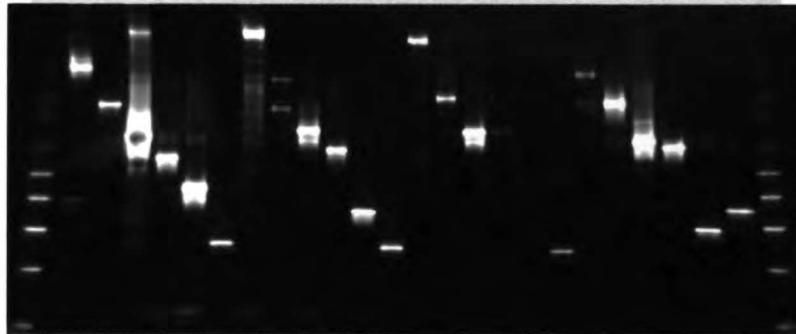
immunoblot

PHB2

phosphorimage
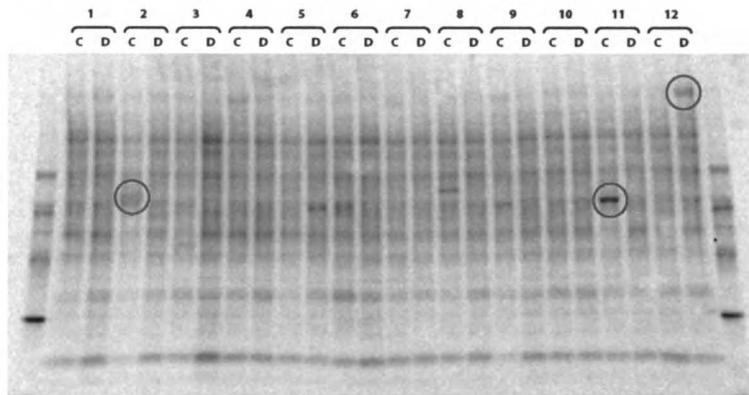
immunoblot
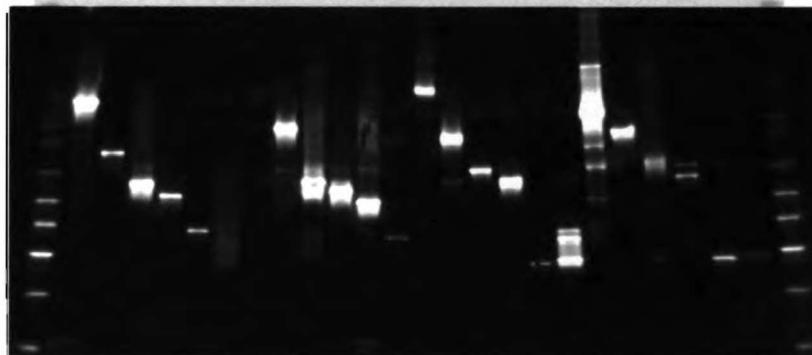
Figure 6: TEV cleavage of Pcl1/Pho85 substrates. (A) To distinguish whether each phosphoprotein was the TAP fusion itself (left) or an associated protein (right), we performed the in extract assay, then cleaved with TEV protease. If the phosphoprotein is the TAP fusion itself, there will be a mobility shift resulting from loss of the ZZ tag. If the phosphoprotein is an associated protein, there will be such mobility shift. (B) Results of TEV cleavage. Shown is a phosphorimage of candidate substrates with and without addition of TEV.

A

If the TAP fusion itself
is phosphorylated:

If an associated protein
is phosphorylated:

-TEV   +TEV

-TEV   +TEV

SDS-PAGE/
phosphorimage

B

RPB2 RPB7 RPB11 SAP185 SSD1 ROM2 IPP1 KCC4 VIP1 RPG1 UGP1 NIP1   PFK1 CDC19 NPL3 PRT1 CDC10 CKA2

− + − + − + − + − + − + − + − + − + − + − + − +   − + − + − + − + − + − +
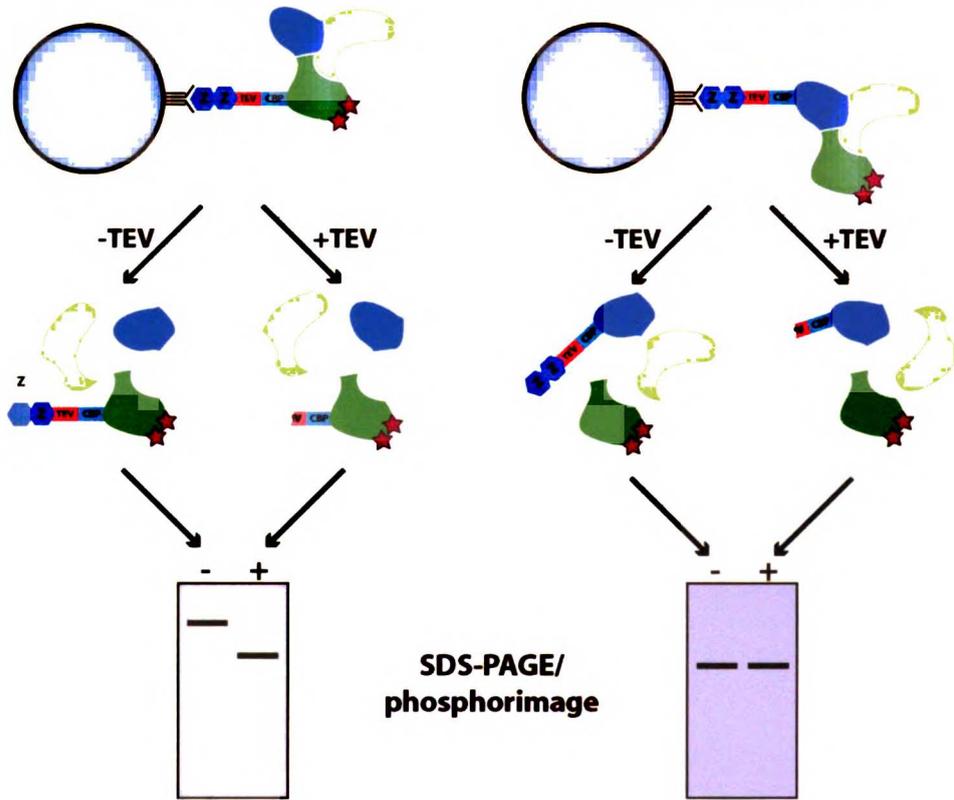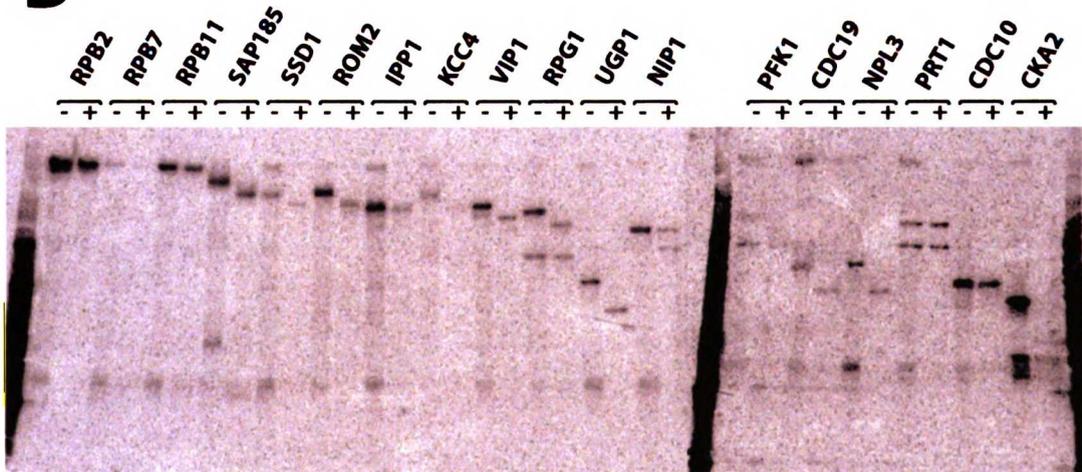
Table 2: Summary of putative Pcl1/Pho85 substrates identified in the screen. The ORF and gene name of each putative substrate are noted, as well as the results of TEV cleavage, and a description of the function of each gene from the *Saccharomyces* genome database.

ORF	Gene	Cleaved by TEV?	Description
YOR151C	RPB2	-	RNA Pol. II subunit
YDR404C	RPB7	-	RNA Pol. II subunit
YOL005C	RPB11	-	RNA Pol. II subunit
YJL098W	SAP185	+	Associated with Sit4 protein phosphatase; G1/S transition
YDR293C	SSD1	+	RNA binding; cell wall organization and biogenesis
YLR371W	ROM2	+	GEF for Rho1; involved in actin organization, bud growth, cell wall organization and establishment of polarity
YBR011C	IPP1	-	Inorganic pyrophosphatase; associates with Rom2
YCL024W	KCC4	?	Kinase involved in septin organization
YLR410W	VIP1	+	Actin cytoskeleton organization and biogenesis
YBR079C	RPG1	+	Translation initiation factor eIF3 subunit
YOR361C	PRT1	-	Translation initiation factor eIF3 subunit
YMR309C	NIP1	-	Translation initiation factor eIF3 subunit
YGR240C	PFK1	-	metabolism
YAL038W	CDC19	+	Pyruvate kinase; glycolysis; required for START
YDR432W	NPL3	+	mRNA binding and export
YKL035W	UGP1	+	UDP glucose metabolism; transcription induced by inhibition of PHO85
YCR002C	CDC10	-	Septin
YOR061W	CKA2	?	Protein kinase; involve in many processes including G1/S transition, G2/M transition, cell polarity, flocculation, DNA damage
YPL204W	HRR25	ND	Casein kinase I isoform; may be involved in DNA-damage repair

137

Chapter 5:

Conclusion

Difficulties in proteomics and the importance of protein level quality control

In this dissertation, I have described two different approaches, both "proteomic" in nature, towards identifying novel substrates of Pho85. From the failure of the GST collection approach in chapter 2, the seeming success of the in extract assay on the TAP collection in chapter 4, and even from the construction process of the TAP and GFP collections in chapter 3, an important theme emerges. In taking on systematic or global proteomics-level approaches, it is imperative to understand the quality of the reagents used, and how they are represented in whatever assay is chosen. Further, the most meaningful metric of this quality is at the level of the proteins themselves. Without some understanding at this level, these approaches are not really "proteomics" at all, but rather large scale hit or miss biology. Clearly no approach will ever be fully representative of the proteome, but having a quantitative understanding of just how representative one's approach is makes the results from such experiments far more interpretable.

As the quality of proteome level data is directly dependent on the quality of the reagents used, we undertook the construction of two new collections of yeast for this purpose. We verified accurate construction at multiple levels, most importantly verifying expression of the fusion proteins by immunoblot and microscopy for the TAP and GFP collections respectively.

Uses of the TAP and GFP collections

These collections represent exciting possibilities for performing large scale analyses *in vivo* or in conditions closely resembling the physiological context of the cell. We chose to adapt the TAP collection to a high throughput kinase assay, but any extract-based assay should be amenable to such modification. Many different post-translational modifications could be examined, or other protein characteristics, such as interaction with other proteins, could be monitored, as well as protein level responses to various environmental stimuli. Further refinement of these collections and the methodology for utilizing them will undoubtedly continue to yield interesting and useful data in the future.

Development of a novel systematic approach to substrate identification

Difficulty in interpretation of the results of our substrate identification approach with the GST collection led us to develop new methodology and reagents to search for novel protein kinase substrates in a systematic manner. For the assay itself, our primary concern was in maintaining the physiological specificity of kinases for their substrates. We have found that in many cases, and notably with Pho85 kinases, that proper physiological substrate specificity is not maintained *in vitro*. Though the assay we developed is still an *in vitro* assay, the conditions used resemble the cellular context far closer than a purified components approach. The reaction takes place in a native extract, so all cellular proteins are represented, and close to physiological concentrations are used for kinase itself. By attempting to replicate the physiological enzymatic context of the

kinase, we hope that proper substrate specificity is maintained. In the case of Pho85 and its substrate Pho4 we see that, unlike a purified components approach, proper specificity is maintained in this assay, even for closely related Pho85 kinases.

We adapted this assay to a high throughput systematic analysis of the proteome in order to identify novel substrates of Pho85 kinases, and performed a partial screen of the proteome with the Pho80/Pho85 and Pcl1/Pho85 kinases. These two kinases yielded completely unique profiles of hits in these screens, which speaks to the kinase specificity maintained in the approach. Further, the particular substrates identified, especially for Pcl1/Pho85, are consistent with processes in which this kinase is known to be involved.

Preliminary data seems promising in validating this approach towards protein kinase substrate identification. True validation, however, must await the verification of the physiological relevance of the identified substrates.

Substrates of Pcl1/Pho85

The putative substrates identified for Pcl1/Pho85 in the screen provide a compelling list of possible functions of this kinase. Most notably, *ROM2* displays a synthetic lethal interaction with *PHO85(1)*. Though it is not immediately obvious how a kinase could be synthetic lethal with its own substrate, many members of the PKC pathway are genetically linked to Pho85 by a synthetic lethal interaction. The genetics of this pathway indicate that it does not function in a simple linear manner and interacts with other pathways as well, such as the cell cycle, mating,

nutrient signaling, and the cytoskeleton*(2)*. Rom2 functions as the guanine

nucleotide exchange factor (GEF) for Rho1*(3)*, with Bem2 acting as the

corresponding GTPase activating protein (GAP). *BEM2* also displays a synthetic

lethal interaction with *PHO85*, and a *pcl1Δbem2Δ* strain phenocopies this

inviability*(4)*.

The interaction of Pcl1/Pho85 with the PKC pathway is also supported by

its associated phenotypes. Inviability of many mutants in the PKC pathway can

be suppressed by osmotic stabilization with 1M sorbitol, as is also true with the

pcl1Δpcl2Δcln1Δcln2Δ strain*(4)*. Moreover, a *pho85Δ* strain is sensitive to

Calcofluor white and this is phenocopied by deletion of the *PCL 1,2* subfamily*(1)*.

Mutants in the PKC pathway are frequently sensitive to this agent*(2)*. The

interaction between Pho85 and the PKC pathway could be quite complex, as

indicated by the fact that Swi4, a downstream effector of the PKC pathway,

regulates the transcription of *PCL1* and *PCL2(5)*.

The other substrates identified include three others with defined roles in

polarized growth and morphogenesis (Sap185, Ssd1, and Vip1). A second group

of substrates consists of Npl3 and Rpg1, proteins which perhaps function in

promoting localized translation at the emerging bud. A third class of substrates

function in carbon source utilization, in which Pho85 is involved.

Other proteins which were either not cleaved or not resolved in the TEV

experiment await further characterization, but can be connected in many cases

to functions of Pho85. They seem to be involved in a variety of processes,

including transcription (Rpb2, Rpb7, Rpb11), septin organization (Kcc4, Cdc10),

response to DNA damage (Hrr25), G1 progression (Cka2), and carbon source utilization (Pfk1). The true identity of the substrates, and their connection to Pho85, remains to be seen.

Speculation on Pho85

An emerging picture of the role of Pcl1/Pho85 in cell wall integrity and polarized growth and morphogenesis is supported by many of the putative substrates identified in the screen. Though these substrates are related functionally, how they are related at a molecular level is not clear. Do they act together in performing one task, or do they all perform different aspects of the same general biological process?

It is becoming clear that a linear paradigm of signaling does not apply in this case. The role of Pcl1/Pho85 in G1 progression and in cell wall integrity and morphogenesis may be to coordinate the many processes involved in these functions. This coordination could involve the spatial or temporal aspects of these processes, or the crosstalk of different pathways. In this way, the impact of phosphorylation by Pho85 may not be readily assayable, especially if there exists variation within the cell or through the cell cycle. Further, Pho85 may regulate these processes in a graded manner, tweaking the processes involved in response to small environmental variations to achieve the proper outcome.

This type of regulatory role for Pho85 stands in contrast to its role in responding to phosphate starvation. In this case, Pho80/Pho85 phosphorylates a single well characterized substrate in response to a known environmental

condition. The effective output of this phosphorylation is transcriptional, which is readily assayable by a number of different methods.

For Pcl1/Pho85, we do not understand what environmental variable regulates this kinase. Further, it seems to have many substrates, with varying degrees of functional knowledge. How these substrates interact functionally is not known, and in many cases it is not clear how to assay their activity.

In order to form a clearer picture of the biology of the Pcl1/Pho85 kinase, it will clearly be necessary to further investigate whatever substrates are identified. It will also be important, however, to understand how these substrates interact and function together in promoting whatever process is regulated by Pcl1. This will require drawing on knowledge from a variety of different lines and types of investigation, and will require a holistic perspective on the downstream effectors of Pcl1/Pho85 function.

Literature cited

1. Huang, D., Moffat, J. & Andrews, B. 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 (2002).

2. Heinisch, J.J., Lorberg, A., Schmitz, H.P. & Jacoby, J.J. The protein kinase C-
   mediated MAP kinase pathway involved in the maintenance of cellular
   integrity in Saccharomyces cerevisiae. *Mol Microbiol* **32**, 671-680 (1999).

3. Ozaki, K. et al. Rom1p and Rom2p are GDP/GTP exchange proteins (GEPs)
   for the Rho1p small GTP binding protein in Saccharomyces cerevisiae.
   Embo J **15**, 2196-2207 (1996).

4. Lenburg, M.E. & O'Shea, E.K. Genetic evidence for a morphogenetic function
   of the Saccharomyces cerevisiae Pho85 cyclin-dependent kinase.
   Genetics **157**, 39-51 (2001).

5. Baetz, K., Moffat, J., Haynes, J., Chang, M. & Andrews, B. Transcriptional
   coregulation by the cell integrity mitogen-activated protein kinase Slt2 and
   the cell cycle regulator Swi4. *Mol Cell Biol* **21**, 6515-6528 (2001).

145

Appendix A:

Pho86p, an endoplasmic reticulum (ER) resident protein in Saccharomyces cerevisiae, is required for ER exit of the high-affinity phosphate transporter Pho84p

W.-T. Walter Lau*, Russell W. Howson†, Per Malkus‡, Randy Schekman‡, and Erin K. O'Shea*†§

*Program in Biophysics, †Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0448; and ‡Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720

In the budding yeast *Saccharomyces cerevisiae*, *PHO84* and *PHO86* are among the genes that are most highly induced in response to phosphate starvation. They are essential for growth when phosphate is limiting, and they function in the high-affinity phosphate uptake system. *PHO84* encodes a high-affinity phosphate transporter, and mutations in *PHO86* cause many of the same phenotypes as mutations in *PHO84*, including a phosphate uptake defect and constitutive expression of the secreted acid phosphatase, Pho5p. Here, we show that the subcellular localization of Pho84p is regulated in response to extracellular phosphate levels; it is localized to the plasma membrane in low-phosphate medium but quickly endocytosed and transported to the vacuole upon addition of phosphate to the medium. Moreover, Pho84p is localized to the endoplasmic reticulum (ER) and fails to be targeted to the plasma membrane in the absence of Pho86p. Utilizing an *in vitro* vesicle budding assay, we demonstrate that Pho86p is required for packaging of Pho84p into COPII vesicles. Pho86p is an ER resident protein, which itself is not transported out of the ER. Interestingly, the requirement of Pho86p for ER exit is specific to Pho84p, because other members of the hexose transporter family to which Pho84 belongs are not mislocalized in the absence of Pho86p.

Yeast cells respond to phosphate starvation by up-regulating the activity of a high-affinity phosphate uptake system (1). Pho84p and Pho86p function in this phosphate transport system. Both *PHO84* and *PHO86* are regulated

147

transcriptionally in response to extracellular phosphate levels by a phosphateresponsive signal transduction pathway (the PHO pathway) (2). Expression of these genes is greatly induced in low-phosphate medium (3, 4). *PHO84* and *PHO86* are essential for growth under low-phosphate conditions and are required for the transcriptional repression of *PHO5*, which encodes a secreted acid phosphatase, in high-phosphate medium.

Strains containing loss-of-function mutations in *PHO84* and *PHO86* exhibit similar phenotypes, e.g., phosphate uptake defects and constitutive expression of phosphate responsive genes. *PHO84* was identified in a screen for mutants that express *PHO5* constitutively (Pho^c) (5). *PHO84* encodes a member of the hexose transporter family that contains 12 transmembrane domains. Biochemical experiments demonstrate that Pho84p is a high-affinity phosphate transporter (6) that is conserved in plants and fungi (7–11). *PHO86* was also identified in a genetic selection for mutants that exhibit the Pho^c phenotype (12) and in a screen for mutants that confer arsenate resistance (13). *PHO86* encodes a protein that associates with membranes, presumably through its two predicted transmembrane domains (4).

PHO86 is not required for transcriptional activation of *PHO84* (13), raising the possibility that Pho86p is directly involved in the high-affinity phosphate uptake system. Pho86p could be a phosphate transporter that associates with Pho84p at the plasma membrane for phosphate uptake. Alternatively, Pho86p might be required for proper localization of Pho84p to the plasma membrane. Proteins destined for the plasma membrane are synthesized, processed, and

148

folded in the endoplasmic reticulum (ER), and, once folded, they are packaged into ER-derived COPII vesicles for transport to the Golgi apparatus and then to the plasma membrane (14, 15). Accessory proteins have been identified that assist in the transport of secretory proteins through the secretory pathway. For example, Vps10p is required for the sorting of the soluble vacuolar protein carboxypeptidase Y from the Golgi to the vacuole (16), whereas Ast1p ensures efficient transport of the plasma membrane ATPase (Pma1p) from the Golgi to the plasma membrane (17). Some accessory proteins function in an early stage of the secretory pathway. One such example is Shr3p, which is required for the ER exit of the general amino acid permease (Gap1p) (18, 19). Because of the physiological importance of Pho84p in low-phosphate conditions, Pho86p may function exclusively to ensure rapid and faithful transport of the permease to the cell surface.

In this paper, we report that Pho86p is required for specific packaging of Pho84p into COPII vesicles derived from ER membranes, but itself is not packaged into COPII vesicles, indicating that Pho86p belongs to a class of "outfitters" (20), resident ER proteins that facilitate the loading of cargo into transport vesicles.

Materials and Methods

Media, Genetic Methods, and Strains

Standard yeast media are as described (21) and media contained 2% glucose unless otherwise specified. No-phosphate medium is as described (12).

Crosses, sporulation, and tetrad analysis were performed by standard genetic methods (22). *S. cerevisiae* strains used in this study were EY0664 (*MATa ura3-52 leu2-3*, isogenic to RSY255) and its derivatives, EY0665 (*pho86Δ::LEU2*), EY0666 (*pho84::PHO84-GFP*), EY0667 (*pho84::PHO84-HA*), EY0668 (*pho86Δ::LEU2 pho84::PHO84-HA*), and EY0669 (*pho86Δ::LEU2 pho84::PHO84-GFP*). EY0665 was constructed by transforming EY0664 with EB1123 partially digested with *Kpn*I and *Sac*I, and selecting on synthetic medium lacking leucine. EY0666 and EY0667 were constructed by transforming EY0664, respectively, with EB1124 and EB1125 that were partially digested with *Xho*I, and selecting on minimal medium lacking uracil; the resulting strains were streaked for single colonies on 5-fluoroorotic acid plates. EY0666 was obtained by screening for Pho84p-GFP by direct fluorescence, whereas EY0667 was obtained by screening for Pho84p-HA by immunoblotting. EY0668 and EY0669 were constructed by disrupting *PHO86* in EY0666 and EY0667, respectively, with EB1123 partially digested with *Kpn*I and *Sac*I. Other yeast strains were EY0687 (*MATa end4ts his4 leu2 ura3*, isogenic to RSY1580), EY0688 (*MATa pma1::HA-PMA1::LEU2 ade2–101c his3-Δ200 leu2-Δ1 lys2–801am trp1-Δ63 ura3–52*, isogenic to RSY1578), and its derivative EY0689 (*pho86Δ::TRP1*). To construct EY0689, EB0477 was digested with *Kpn*I and *Sac*I, transformed into EY0688, and transformants were selected on synthetic medium lacking tryptophan.

Plasmids

Plasmids p*PHO84-GFP* (EB0666) and p*PHO86-GFP* (EB0667) were

constructed by fusing PCR-generated *XhoI/ Eco*RI fragments of *PHO84* and

PHO86, respectively, with their native promoters in-frame to the N terminus of

the green fluorescent protein (GFP) in pRS316. A disruption vector that replaces

the entire ORF of *PHO86* with *TRP1* (EB0477) was constructed as follows. A

450-bp *KpnI/Eco*RI fragment derived from the 5' noncoding region of *PHO86* and

a 580-bp *Bam*HI/ *Xba*I fragment from the 3' noncoding region of *PHO86* were

generated by PCR and inserted into the Bluescript plasmid to create pTS-

PHO86. An *Eco*RI/*Bgl*II fragment containing the *TRP1* gene from plasmid pJJ248

(23) was inserted into the *Eco*RI/*Bam*HI cut pTS-*PHO86* to create EB0477.

Another disruption vector (EB1123) for *PHO86* was constructed by ligating a 4-kb

*Pst*I/*Bam*HI fragment from EB0477 to a *Pst*I/ *Bam*HI fragment containing the

LEU2 gene from pJJ252 (23). To replace the chromosomal copy of *PHO84* with

PHO84-HA, we constructed an integration vector (EB1124) in three steps. First,

an *XhoI/Sac*I PCR product containing *PHO84* was cloned into pRS306 digested

with *Xho*I and *Sac*I to create pRS306- *PHO84*. In this step, a polylinker

containing a *Not*I site was added in-frame to the C-terminal coding region of

PHO84 immediately preceding the stop codon. Second, a *Not*I fragment

containing three tandem copies of the hemagglutinin epitope (HA) was inserted

into the *Not*I site of pRS306-*PHO84* to create pRS306-*PHO84-HA*. Third, a 0.5-

kb *Eco*RI/*Sac*I PCR product derived from the 3' noncoding region of *PHO84* was

cloned into the *Eco*RI/*Sac*I sites of pRS306-*PHO84-HA* to create EB1124. To

integrate *PHO84-GFP* into the yeast genome, we constructed the integration

vector EB1125 by first cloning an *XhoI/Sac*I fragment of p*PHO84-GFP* into

pRS306 to create pRS306-*PHO84-GFP*. A *Bam*HI/*Sac*I PCR product derived

from the 3' noncoding region of *PHO84* was then cloned into pRS306- *PHO84-*

GFP at the *Bam*HI/*Sac*I sites to make EB1125. Epitopetagged *PHO86* (p*PHO86-*

HA, EB1126) was constructed as follows. A PCR-generated *Xho*I/*Eco*RI

fragment of *PHO86* was cloned into the *Xho*I/*Eco*RI sites of pRS316, and a

polylinker containing a *Not*I site was added in-frame to the C-terminal coding

region of *PHO86* immediately preceding the stop codon. A *Not*I fragment

containing three copies of HA was inserted into the *Not*I site of this construct to

create EB1126. A low copy plasmid containing *GAL2-GFP* (*URA3* marked) was a

generous gift from A. Kruckeberg (University of Amsterdam, The Netherlands).

The plasmid that carries *ERO1-HA* was a gift from J. Weissman (University of

California, San Francisco).

In Vitro Vesicle Budding Assay

This assay was a modification of a previously described procedure (19),

and was performed at 30°C unless otherwise specified. Yeast cultures grown in

yeast extract/peptone/dextrose (YPD) or synthetic solid media were used to

inoculate overnight stock cultures. Cells were inoculated into YPD or synthetic

high-phosphate media until the cultures reached an optical density at 600 nm

(OD_{600}) of approximately 0.5. Yeast cells were washed three times in sterile

water, inoculated into no-phosphate medium at an OD_{600} of 0.5, and grown for 2

h. Cells were harvested, washed twice in no-phosphate medium, and

resuspended in no-phosphate medium to 2 OD_{600}/ ml, and shaken for 15 min.

The cultures were labeled with 1 mCi/20 OD$_{600}$/ml 35S-Promix (1,200 Ci/mmol; Amersham Pharmacia) for 3 min. Metabolic activity was stopped by the addition of NaN$_3$ and NaF (20 mM final). Spheroplast preparation and lysis were performed as described (24). Gently lysed spheroplasts were washed once with "low-salt B88" [20 mM Hepes (pH 6.8)/50 mM KOAc/250 mM sorbitol/5 mM MgOAc], and twice with B88 [20 mM Hepes (pH 6.8)/150 mM KOAc/250 mM sorbitol/5 mM MgOAc]. The pellet was resuspended in 0.5 ml B88, and 0.5 ml "high-salt B88" [20 mM Hepes (pH 6.8)/1 M KOAc/250 mM sorbitol/5 mM MgOAc] was added. The sample was allowed to chill on ice for 10 min, and washed twice with B88. Each budding reaction contained membranes (2.5 OD600/ml lysed pheroplasts), COPII components (1–2 mg Sec13/31p, 1–2 mg Sec23/24p, and 1 mg Sar1p), or 100 mg crude cytosol (25) supplemented with 1 mg Sar1p in a total volume of 125 ml with 1X ATP regeneration system (19) and 0.1 mM GTP. Reactions were incubated at 28°C for 45 min, and 10% of the total reaction was removed for analysis (Total). The remaining reaction mixture was centrifuged at medium speed (12,000 X g) for 4 min, and 75 ml of the supernatant fraction was collected for analysis (Vesicle fraction, 60% of Total). Proteins of interest were analyzed by immunoprecipitation from Total or Vesicle fractions.

Immunoprecipitation

Immunoprecipitations (IPs) were performed as described in ref. 19 with minor modifications. After each sample was incubated for 10 min at 55°C in 20

ml of 5% SDS, the volume was adjusted to 1ml with IP buffer without SDS [150 mM NaCl/1% Triton X-100/50 mM TrisCl (pH 7.5)] with protease inhibitors (1 mM PMSF/0.5 mg/ml leupeptin/0.7 mg/ml pepstatin), 50 ml of 30% (vol/vol) protein G-Sepharose in IP buffer without detergents, and 2 ml (1.2 mg/ml) monoclonal anti-HA antibody (12CA5) or 2 ml polyclonal anti-Vph1p serum (a gift from T. Stevens, University of Oregon, Eugene, OR). IP reactions were washed, resolved by SDS/PAGE, and analyzed with a PhosphorImager (Molecular Dynamics).

Fluorescence Microscopy and Immunofluorescence

Phosphate starvation and GFP direct fluorescence in live yeast cells were performed as described (26). To study Gal2p-GFP localization, we grew cells expressing Gal2p-GFP in raffinose medium to log phase, and 1% galactose was added to the culture to induce GAL2 expression for 1 h. Immunofluorescence was conducted as described (27) by using formaldehyde fixation. For Pma1p and Kar2p localization, the mouse monoclonal anti-HA antibody 16B12 (Babco, Richmond, CA) and a rabbit anti-Kar2p polyclonal antibody were used, respectively. Secondary antibodies were goat anti-mouse IgG-conjugated to BODIPY TMR-X, and goat anti-rabbit IgG-conjugated to BODIPY FL (Molecular Probes). Images of coimmunofluoresence of Pho84p-GFP or Pho86p-GFP and Kar2p were documented by using a confocal microscope (LEICA TCSNT, Wetzlar).

154

Results

Localization of Pho84p-GFP Is Regulated in Response to Extracellular

Phosphate Levels.

To investigate mechanisms for regulating the activity of Pho84p, we sought to study regulation of its localization in live cells by fluorescence microscopy. We constructed a fusion protein between Pho84p and GFP driven by the *PHO84* promoter on a low copy plasmid (EB0666). This plasmid allowed a *pho84* mutant strain to grow on low-phosphate plates and complemented the Phoc phenotype. As expected, Pho84p-GFP is localized to the plasma membrane in no-phosphate medium in wild-type cells (EY0664) (Fig. 1).

To investigate changes in the localization of Pho84p in response to changes in phosphate levels, we added phosphate to a phosphate-starved culture. Localization of Pho84p-GFP was monitored in the absence of new protein synthesis by the addition of cycloheximide. Within 30 min, most Pho84p-GFP was internalized and localized to the vacuole (Fig. 1). However, if KCl instead of KH_2PO_4 was added under the same conditions, most Pho84p-GFP remained at the plasma membrane (data not shown), suggesting that the changes in Pho84p-GFP localization were because of changes in the rate of internalization in response to different phosphate levels in the medium. If arsenate, a phosphate analog, was added to the phosphate-starved culture in the presence of cycloheximide, Pho84p-GFP was not internalized, indicating that the changes in Pho84p localization are specific to phosphate levels in the culture (data not shown). Because the transcription of *PHO84* is regulated by the PHO

155

pathway, we wished to determine if the change in localization of Pho84p was also regulated by this signaling pathway. To test whether this was the case, we studied Pho84p-GFP localization in other PHO mutants. In a *pho85Δ* strain and in a strain containing *PHO81[c]*, a hyperactive allele of *PHO81*, changes in localization of Pho84p-GFP are similar to wild type (data not shown), demonstrating that regulation of Pho84p localization is independent of the phosphate signaling pathway.

To test whether internalization of Pho84p-GFP was mediated by endocytosis, we examined whether internalization of Pho84p was blocked in an *end4[ts]* mutant that is defective in endocytosis at a restrictive temperature. In the *end4[ts]* mutant, grown at the permissive temperature (25°C), Pho84p-GFP was endocytosed upon addition of phosphate (data not shown). However, at the restrictive temperature (33°C), Pho84p-GFP remained at the plasma membrane even after addition of phosphate (Fig. 1), indicating that Pho84p-GFP is internalized through the endocytic pathway. We conclude that Pho84p is regulated by a posttranslational mechanism; its localization is regulated in response to extracellular phosphate levels.

PHO86 Encodes an ER Resident Protein.

Because mutations in *PHO84* and *PHO86* result in essentially the same phenotype, it is possible that Pho84p and Pho86p participate in the same function, uptake of inorganic phosphate. It has been proposed that Pho86p might interact with Pho84p to form a protein complex that is required for high-

156

affinity phosphate uptake (28). If so, Pho86p would be localized to the plasma membrane in low-phosphate medium. Alternatively, Pho86p may reside in an intracellular compartment involved in the biogenesis of Pho84p. To distinguish between these two models, we studied the localization of Pho86p in low- and high-phosphate medium. We tagged the *PHO86* gene at the C terminus with GFP and found that a low copy construct (EB0667) complemented the Phoc and low-phosphate lethal phenotypes of a *pho86Δ* strain (EY0665). In low-phosphate medium, Pho86p-GFP is localized to a perinuclear compartment coincident with Kar2p/Bip, a known ER resident protein (Fig. 2). Pho86p-GFP was also found in the ER in high-phosphate medium (data not shown), indicating that its localization is not regulated in response to extracellular phosphate levels. Thus, Pho86p is not likely to be a phosphate transporter. Instead, it may be required for the synthesis or transport of Pho84p from the ER.

Pho84p-GFP Is Localized to the ER in the pho86Δ *Mutant.*

    If Pho86p is involved in the transport of Pho84p, the permease may be retained in the ER in the absence of Pho86p. To test this hypothesis, we integrated *PHO84-GFP* (EB1125) into the yeast genome and investigated the localization of Pho84p-GFP in a strain lacking *PHO86* (EY0669) by immunofluorescence in low-phosphate medium. As shown in Fig. 3, Pho84p-GFP is predominantly localized to the ER in a *pho86* mutant, as indicated by its colocalization with Kar2p. By coimmunofluorescence, we also showed that Pho84p-GFP colocalizes with Ero1p-HA, another known ER resident protein

(data not shown). Together with the observation that Pho84p stably accumulates in a *pho86* mutant (data not shown), these results support the view that Pho86p has a role in Pho84p transport.

Pho86p Is Not Required for Targeting of Gal2p-GFP or Pma1p-HA to the Plasma Membrane.

To investigate the specificity of Pho86p, we examined the subcellular localization of other plasma membrane proteins in the *pho86Δ* strain by indirect immunofluorescence. Pma1p-HA was localized to the plasma membrane in a *pho86Δ* strain (Fig. 4). However, Pho86p might still be required for the ER exit of a specific family of proteins. To determine whether deletion of *PHO86* affects other transporters related in sequence to Pho84p, we investigated the localization of Gal2p-GFP in the *pho86Δ* strain. Gal2p is a galactose transporter in yeast, and shares 22% identity and 37% similarity with Pho84p (29). In both wild-type and *pho86Δ* strains, Gal2p-GFP is localized to the plasma membrane (Fig. 4). Moreover, *pho86Δ* cells are able to grow in a variety of conditions, including different carbon sources, and they can mate normally. From these data, we conclude that the function of Pho86p is likely to be specific to Pho84p.

Pho84p Is Packaged into COPII Transport Vesicles in Vitro.

Folded and mature secretory proteins are packaged into COPII vesicles, which bud from the ER membrane and then fuse with the Golgi apparatus. Pho86p may function in folding, cargo selection into COPII vesicles, or cargo

retention in the Golgi membrane. To characterize the details of Pho84p exit from the ER, we studied its packaging into COPII vesicles derived from wild-type membranes utilizing an *in vitro* vesicle budding assay. This assay has been used to study packaging of Gap1p and Pma1p (19), and a subunit of the V_0 complex of the vacuolar ATPase, Vph1p (P.M. and R.S., unpublished data). COPII vesicles can be generated from perforated spheroplasts by addition of ATP and GTP together with whole cytosol or purified COPII proteins that consist of a small GTPase, Sar1p, and coat proteins Sec23p, Sec24p, Sec13p, and Sec31p.

We tagged Pho84p at its C terminus with three tandem copies of the hemagglutinin antigen (HA) epitope and integrated this construct (EB1124) into the yeast genome to replace the wildtype copy of Pho84p. The resulting strain (EY0667) behaves like wild type with respect to growth in low-phosphate medium and *PHO5* repression in high-phosphate medium. EY0667 cells were starved for phosphate and pulse-radiolabeled with [^{35}S]methionine for 3 min, a short time period intended to label all newly synthesized proteins, including precursors freshly assembled in the ER. Perforated spheroplasts were prepared and incubated with purified COPII proteins to generate COPII vesicles *in vitro*. Vesicles were separated from membranes, and Pho84p-HA was immunoprecipitated under denaturing conditions from either the total budding reaction or the vesicle fraction and quantitatively analyzed. Pho84p-HA was packaged with approximately 20% efficiency into COPII vesicles when COPII proteins were included; the efficiency was reduced by more than 3-fold if Sar1p

or any one of the purified coat proteins was omitted (data not shown), indicating that the packaging of Pho84p-HA into transport vesicles requires each of the COPII components (Fig. 5). As a control, we performed another immunoprecipitation to detect the packaging of Vph1p, from both the total and vesicle fraction of the budding reaction after Pho84p had been immunoprecipitated (Fig. 5). When perforated spheroplasts pulse-labeled for a longer time (18 min) were used in the budding reaction, the efficiency of Pho84p packaging was reduced to about 3% (data not shown), consistent with a greater proportion of Pho84p located in the compartments beyond the ER. These results demonstrate that standard COPII proteins and nucleotides are sufficient to package Pho84p-derived vesicles from wild-type membranes.

Pho86p Is Required for Packaging Pho84p into Transport Vesicles in Vitro.

We next addressed the fate of Pho84p in membranes isolated from a *pho86Δ* mutant. Perforated spheroplasts were prepared from radiolabeled cells derived from a *pho86Δ* mutant expressing Pho84p-HA (EY0668), and the budding reactions were performed. Pho84p-HA was not detected above background levels in the vesicle fraction when purified COPII proteins were included in the budding reaction (Fig. 6). To test whether deletion of *PHO86* has a general effect on cargo packaging into COPII vesicles, we investigated the packaging of Vph1p into COPII vesicles after Pho84p-HA had been immunoprecipitated. Approximately 12% of newly synthesized Vph1p was packaged by using *pho86Δ* mutant membranes, and this level was comparable

160

to that achieved with wild-type membranes (see Fig. 5). These results suggest that Pho86p is not required for the general formation of COPII vesicles from the ER, or any later step in the secretory pathway. Rather, it is required at an early step for packaging of Pho84p into COPII vesicles. Thus, Pho86p belongs to a class of substrate-specific accessory proteins that have selective roles in the packaging of membrane proteins.

Pho86p Is Not Packaged into COPII Vesicles in Vitro.

Accessory proteins that facilitate cargo transport from ER to Golgi can be divided into three classes: "outfitters," "escorts," or "guides"(20). A fundamental feature of an "outfitter" is that it is not expected to accompany its cargo to the Golgi, and thus would not be packaged into COPII vesicles. Even though Pho86p is localized to the ER when yeast cells are grown in high- or lowphosphate medium, we could not rule out the possibility that Pho86p shuttles between the ER and the Golgi, as is expected of "escort" and "guide" proteins.

To follow the fate of Pho86p, we constructed a low copy plasmid (ARS/CEN) in which three copies of the HA epitope were introduced at the C terminus of Pho86p under the control of its native promoter (EB1126). The epitope-tagged Pho86p complemented the Phoc phenotype of a *pho86Δ* strain (EY0665) and allowed the strain to grow in low-phosphate medium. Indirect immunofluorescence indicated that Pho86p-HA was localized to the ER both in high- and low-phosphate medium (data not shown), confirming the previous

161

observations made with Pho86p-GFP. To study the packaging of Pho86p into COPII vesicles, we transformed Pho86p-HA into a *pho86Δ* strain, pulse-labeled the cells for 3 min, and prepared perforated spheroplasts. We performed the budding reactions with either whole cytosol or purified COPII proteins supplemented with nucleotides. As shown in Fig. 7, Pho86p-HA did not appear above background levels in the vesicle fractions. As a control, Vph1p was packaged into the transport vesicles under the same conditions (data not shown). These data indicate that Pho86p is an ER resident protein that is likely to function as an "outfitter," and thus, its role must precede the incorporation of Pho84p into COPII vesicles.

Discussion

We have shown that Pho86p functions in the secretory pathway —it is required for packaging of Pho84p, a high-affinity phosphate transporter, into COPII vesicles. Pho86p is likely to belong to a class of "outfitters," resident ER proteins that facilitate the export of cargo molecules from the ER (20). Pho86p might be required for folding, maturation, or oligomerization of Pho84p in the ER so that it can assume a conformation competent for incorporation into COPII vesicles. Another possibility is that Pho86p specifically recruits COPII proteins to the ER and then is displaced by have been reported Pho84p during the budding reaction.

Recently, many examples have been reported of accessory proteins that facilitate the ER exit of cargoes. Three such proteins in *S. cerevisiae*, Vma21p,

Chs7p, and Gsf2p, are required for the ER exit of Vph1p (30), chitin synthase,

Chs3p (31), and a glucose transporter, Hxt1p (32), respectively. Though these

three accessory proteins are localized to the ER in steady state, it is not known if

they function prior to the incorporation of their respective cargoes into COPII

vesicles or if these proteins themselves are transported out of the ER. However,

both Vma21p and Gsf2p possess C-terminal, cytoplasmic retrieval (-KKXX)

sequences, which suggests that these molecules cycle between the ER and the

Golgi apparatus. In contrast, Pho86p appears to function in a manner similar to

that of Shr3p. Neither Pho86p nor Shr3p has a recognizable retrieval signal.

Shr3p is required for packaging of a family of amino acid permeases into COPII

vesicles, but itself is not included in these vesicles(19). Interestingly, other than

the C-terminal retrieval signals on two of these proteins, these accessory

proteins do not share homology, and they facilitate exit of different classes of

cargoes from the ER. It is unclear whether these proteins function as part of the

quality control procedure, or part of the cargo selection process involving COPII

coat proteins and other sorting receptors.

Our data suggest that the function of Pho86p is specific to Pho84p.

Pho86p is not required for the ER exit of other plasma membrane proteins that

were investigated. Interestingly, it is not required for the plasma membrane

targeting of Gal2p, which belongs to the same 12-transmembrane-domain

protein family. In this respect, Pho86p apparently functions in a manner different

from Shr3p and Gsf2p, which are required for the ER exit of a family of

transporters; Shr3p is required for Gap1p and Hip1p (19), and Gsf2p for Hxt1p

163

and Gal2p(32). It is not known whether Pho86p directly interacts with Pho84p. A direct interaction between these two proteins could explain the specificity of Pho86p in substrate selection. The nature of this interaction could be explored by using Pho84p homologues from other species expressed in wild-type and *pho86* mutants, or by the use of Pho84p-Gal2p chimeric proteins.

Among more than 6,000 gene products in yeast, it is estimated that at least 10% of them go through the secretory pathway (14). When phosphate is limiting, the major components of the secretory pathway are very different from those that are present in phosphate-rich conditions. It is likely that there exist mechanisms to ensure efficient ER exit of different cargoes under certain physiological conditions and that the sorting itself may be regulated by nutritional conditions. Nitrogen regulation of sorting and stability of the general amino acid permease, Gap1p, in the Golgi apparatus is an example of nutritionally regulated transport in yeast (33). Pho86 may play such a role because its transcription is highly induced under low-phosphate conditions.

Literature Cited

1. Nieuwenhuis, B. J. & Borst-Pauwels, G. W. (1984) *Biochim. Biophys. Acta.* **770,** 40–46.

2. Oshima, Y. (1982) in *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression,* eds. Strathern, J. N., Jones, E. W. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 159–180.

3. Bun-Ya, M., Nishimura, M., Harashima, S.&Oshima, Y. (1991) *Mol. Cell. Biol.* **11,** 3229–3238.

4. Yompakdee, C., Bun-ya, M., Shikata, K., Ogawa, N., Harashima, S. & Oshima, Y. (1996) *Gene* **171,** 41–47.

5. Ueda, Y. & Oshima, Y. (1975) *Mol. Gen. Genet.* **136,** 255–259.

6. Berhe, A., Friestedt, U. & Persson, B. L. (1995) *Eur. J. Biochem.* **227,** 566–572.

7. Harrison, M. J. & van Buuren, M. L. (1995) *Nature (London)* **378,** 626–629.

8. Muchhal, U. S., Pardo, J. M. & Raghothama, K. G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10519–10523.

9. Leggewie, G., Willmitzer, L. & Riesmeier, J. W. (1997) *Plant Cell* **9,** 381–392.

10. Mitsukawa, N., Okumura, S., Shirano, Y., Sato, S., Kato, T., Harashima, S. & Shibata, D. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7098–7102.

11. Liu, H., Trieu, A. T., Blaylock, L. A. & Harrison, M. J. (1998) *Mol. Plant– Microbe Interact.* **11,** 14–22.

12. Lau, W.-T. W., Schneider, K. R. & O'Shea, E. K. (1998) *Genetics* **150,** 1349–1359.

13. Bun-ya, M., Shikata, K., Nakade, S., Yompakdee, C., Harashima, S. & Oshima, Y. (1996) *Curr. Genet.* **29,** 344–351.

14. Kaiser, C. A., Gimeno, R. E. & Shaywitz, D. A. (1997) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Cell Cycle and Cell Biology,* eds. Pringle, J. R., Broach, J. R. & Jones, E. W. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 91–228.

15. Barlowe, C. (1998) *Biochim. Biophys. Acta.* **1404,** 67–76.

16. Marcusson, E. G., Horazdovsky, B. F., Cereghino, J. L., Gharakhanian, E. & Emr, S. D. (1994) *Cell* **77,** 579–586.

17. Chang, A. & Fink, G. R. (1995) *J. Cell Biol.* **128,** 39–49.

18. Ljungdahl, P. O., Gimeno, C. J., Styles, C. A. & Fink, G. R. (1992) *Cell* **71,** 463–478.

19. Kuehn, M. J., Schekman, R. & Ljungdahl, P. O. (1996) *J. Cell Biol.* **135,** 585–595.

20. Herrmann, J. M., Malkus, P. & Schekman, R. (1999) *Trends Cell. Biol.* **9,** 5–7.

21. Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J.&Struhl, K. (1993) *Current Protocols in Molecular Biology* (Wiley, New York).

22. Sherman, F., Fink, G. & Lawrence, C. (1978) *Methods in Yeast Genetics* (Cold Spring Harbor Lab. Press, Plainview, NY).

23. Jones, J. S. & Prakash, L. (1990) *Yeast* **6,** 363–366.

24. Rexach, M. F., Latterich, M. & Schekman, R. W. (1994) *J. Cell Biol.* **126,** 1133–1148.

25. Spang, A. & Schekman, R. (1998) *J. Cell Biol.* **143,** 589–599.

26. Kaffman, A., Rank, N. M. & O'Shea, E. K. (1998) *Genes Dev.* **12,** 2673–2683.

27. Schwappach, B., Stobrawa, S., Hechenberger, M., Steinmeyer, K. & Jentsch, T. J. (1998) *J. Biol. Chem.* **273,** 15110–15118.

28. Yompakdee, C., Ogawa, N., Harashima, S. & Oshima, Y. (1996a) *Mol. Gen. Genet.* **251,** 580–590.

29. Hodges, P. E., McKee, A. H. Z., Davis, B. P., Payne, W. E. & Garrels, J. I. (1999) *Nucleic Acids Res.* **27,** 69–73.

30. Graham, L. A., Hill, K. J. & Stevens, T. H. (1998) *J. Cell Biol.* **142,** 39–49.

31. Trilla, J. A., Duran, A. & Roncero, C. (1999) *J. Cell Biol.* **145,** 1153–1163.

32. Sherwood, P. W. & Carlson, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 7415–7420.

33. Roberg, K. J., Rowley, N. & Kaiser, C. A. (1997) *J. Cell Biol.* **137,** 1469–1482.

Fig. 1. Localization of Pho84p-GFP is regulated in response to extracellular phosphate levels. Wild type (EY0664, *Upper*) and an *end4^{ts}* mutant (EY0687, *Lower*) harboring p*PHO84-GFP* (EB0666) were studied. Direct fluorescence microscopy of Pho84p-GFP was performed in no-phosphate medium at 25°C (*Left*) and then in medium containing phosphate (final concentration 10 mM) and cycloheximide (final concentration 0.1 mg/ml) at 33°C.

No Phosphate **+ Phosphate**

WT

end4ts

25°C **33°C**
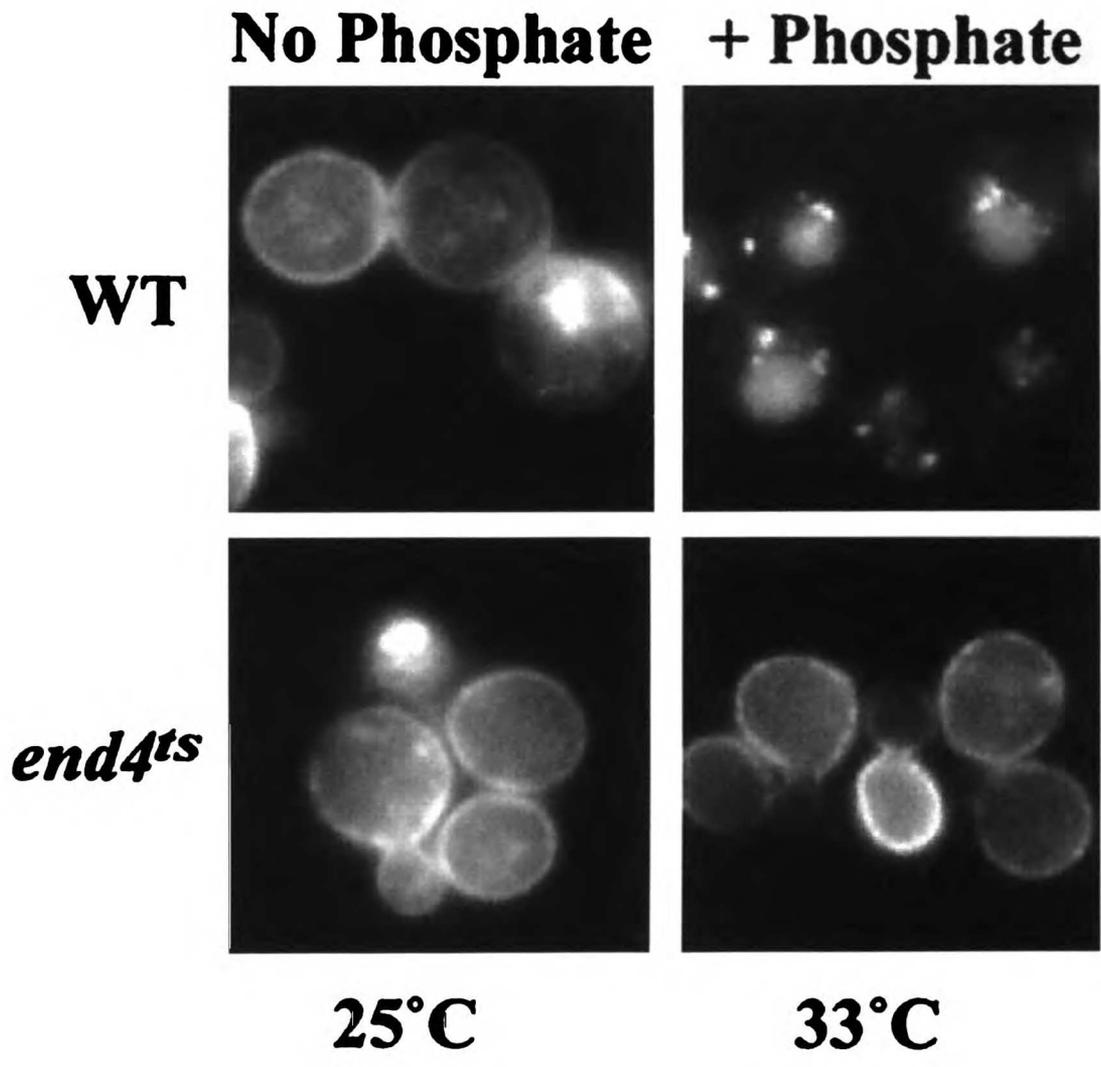
Fig. 2. Pho86p-GFP is localized to the ER. Cells lacking *PHO86* (EY0665) carrying p*PHO86-GFP* (EB0667) were starved for phosphate. Direct fluorescence of Pho86p-GFP was observed. Indirect immunofluorescence was performed by using an antibody to the ER-resident protein Kar2p/Bip. From left to right are fluorescence images corresponding to Pho86p-GFP (green), Kar2p (red), or an overlay of the two.
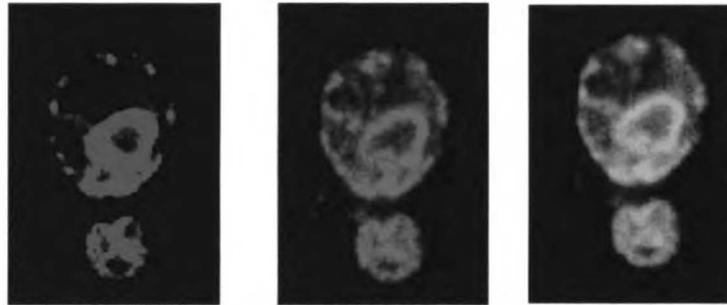
Pho86p-GFP   Kar2p   Overlay

Fig. 3. Pho84p-GFP is localized to the ER in the absence of *PHO86*. A *pho86Δ* strain with integrated *PHO84-GFP* (EY0669) was starved for phosphate and coimmunofluorescence was performed as in Fig. 2. From left to right are fluorescence images corresponding to Pho84p-GFP (green), Kar2p (red), or an overlay of the two.
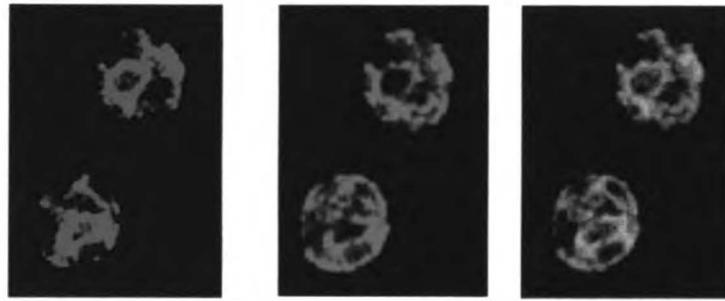
Pho84p-GFP Kar2p   Overlay

Fig. 4. Pho86p is not required for proper localization of Pma1p-HA and Gal2p-GFP. (*Left*) Fluorescence images from wild-type cells.(*Right*) Images from strains lacking *PHO86*. Indirect immunofluorescence was performed on wild-type (EY0688) and *pho86Δ* (EY0689) strains expressing Pma1p-HA by using anti-HA antibodies (*Upper*). Direct fluorescence microscopy was performed on wild-type (EY0664) and *pho86Δ* (EY0665) strains expressing Gal2p- GFP grown in synthetic medium containing 2% raffinose, and 1% galactose (*Lower*).

	WT	pho86Δ
Pma1p-HA		
Gal2p-GFP		
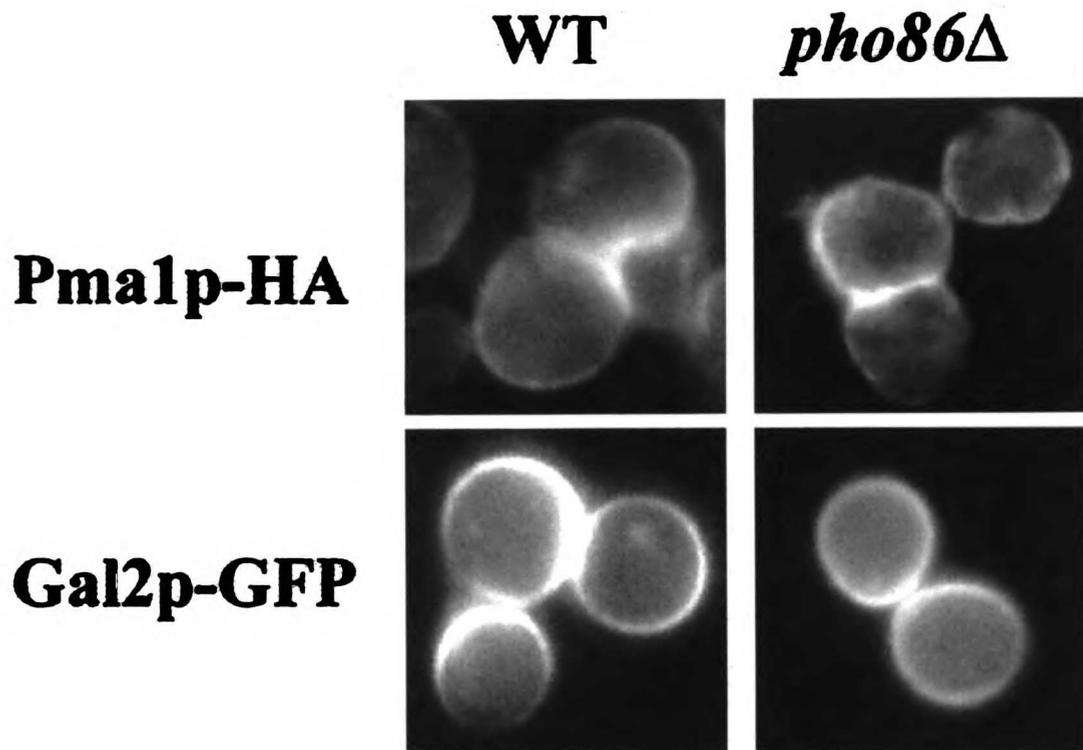
Fig. 5. COPII proteins and nucleotides are sufficient to package Pho84p into

COPII vesicles from wild-type membranes. Permeabilized spheroplasts prepared

from wild-type cells expressing Pho84p-HA (EY0667) were used in vesicle

budding reactions *in vitro*. Reactions contained: nucleotides only (ATP and GTP),

COPII 1Sar1p (purified COPII proteins supplemented with nucleotides and

Sar1p), and COPII2Sar1p (COPII proteins with nucleotides but without Sar1p).

Pho84p-HA was immunoprecipitated from 10% of the total reactions, and the

vesicle fractions were derived from 60% of the total reactions. A second

immunoprecipitation was then performed to detect the amount of Vph1p in the

total reactions and vesicle fractions. The immunoprecipitated proteins were

resolved separately on 8%SDS/PAGEand quantified with a PhosphorImager.

The percent of Pho84p-HA or Vph1p in the vesicle fraction compared with the

corresponding total reaction (histogram) was calculated from the amount of

radiolabeled Pho84p-HA or Vph1p (*Lower*), respectively. T, Total reaction. V,

Vesicle fraction. Data are reported as the mean values of three independent

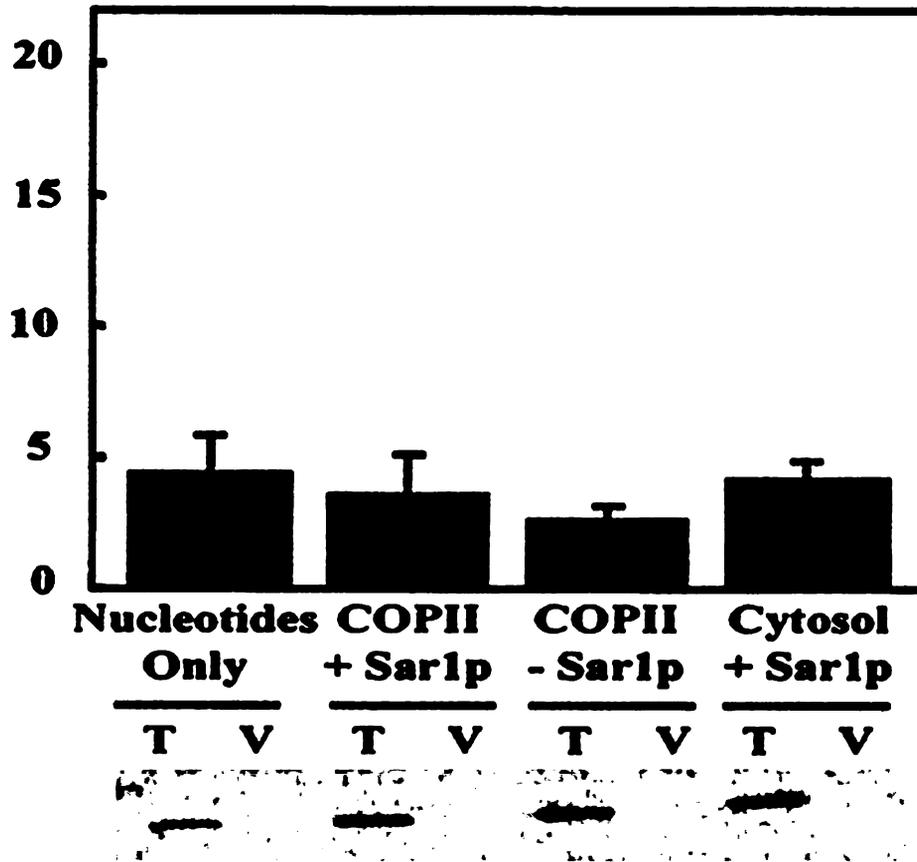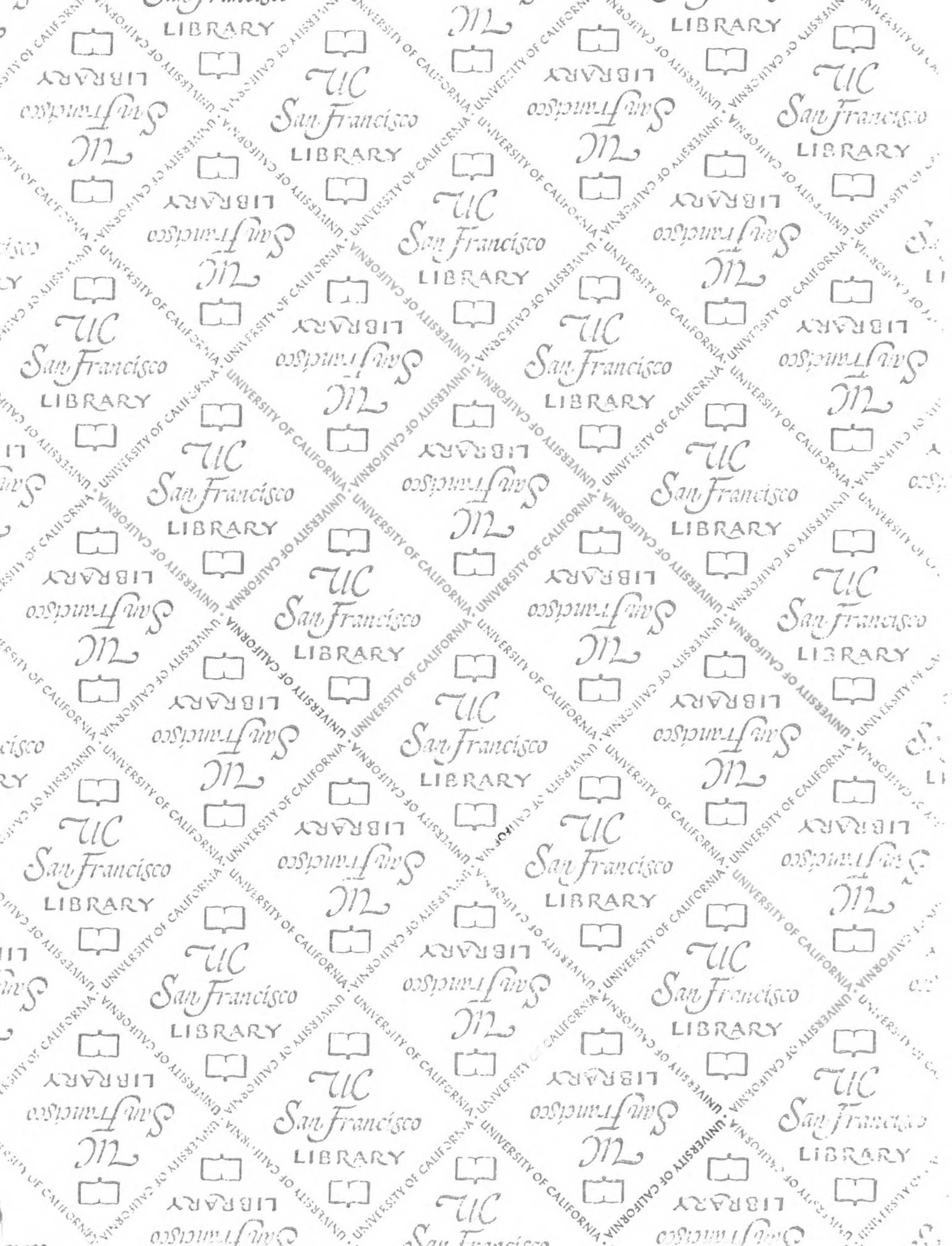experiments, and the error bars indicate the standard deviation.
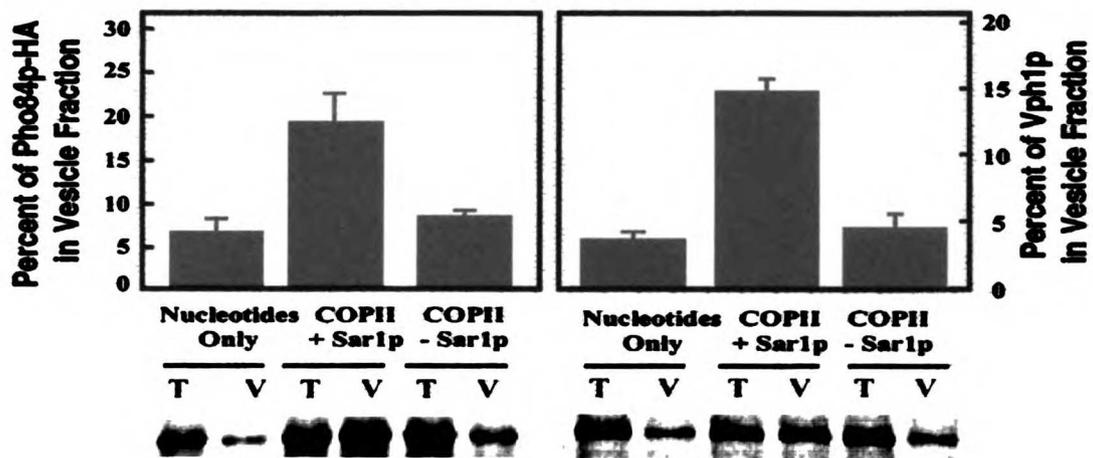
Fig. 6. Pho86p is required for packaging of Pho84p-HA, but not Vph1p, into

COPII vesicles. Vesicle budding experiments and quantitation were performed as

in Fig. 5. Permeabilized spheroplasts prepared from a *pho86Δ* mutant expressing

Pho84p-HA (EY0668) were utilized in this experiment. Pho84p-HA was

immunoprecipitated from both total and vesicle fractions by using anti-HA

antibodies. Vph1p was then immunoprecipitated by using anti-Vph1 antibodies.

Both proteins were resolved separately on 8% SDS/PAGE and analyzed with a

PhosphorImager.
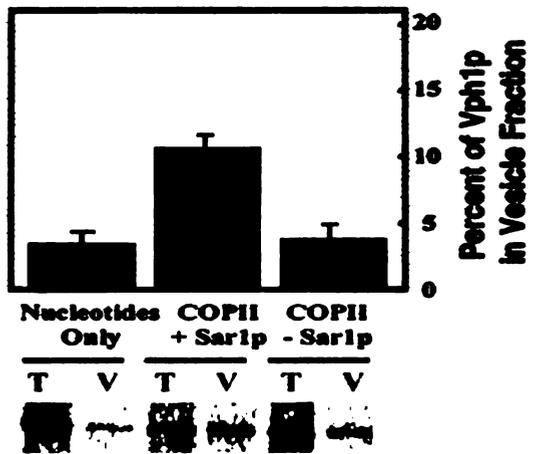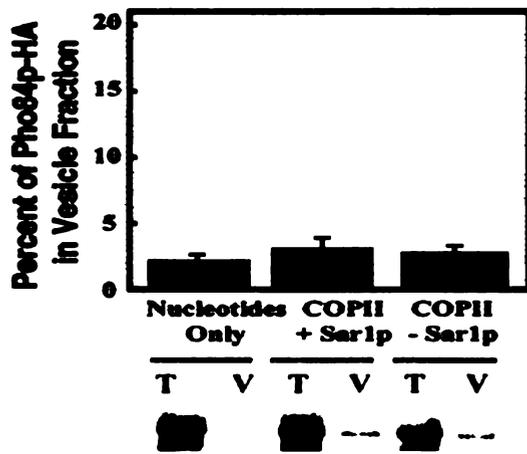
Fig. 7. Pho86 is not packaged into COPII vesicles. Cells lacking *PHO86*

(EY0665) and carrying plasmid p*PHO86-HA* (EB1126) were pulse-labeled, and

perforated spheroplasts were prepared. Vesicle budding experiments were

performed as in Fig. 5, except cytosol 1 Sar1p (cytosol supplemented with Sar1p

and nucleotides) were also included in the reactions. Pho86p-HA was analyzed

by immunoprecipitation and quantitation on a PhosphorImager.