Novel Interactions of *Saccharomyces cerevisiae* Type 1 Protein Phosphatase Identified by Single-Step Affinity Purification and Mass Spectrometry†

Edmund P. Walsh, Douglas J. Lamont, Kenneth A. Beattie, and Michael J. R. Stark*

School of Life Sciences Biocentre, University of Dundee, Dundee DD1 5EH, U.K.

Received September 27, 2001; Revised Manuscript Received December 13, 2001

ABSTRACT: The catalytic subunit of *Saccharomyces cerevisiae* type 1 protein phosphatase ($PP1_C$) is encoded by the essential gene *GLC7* and is involved in regulating diverse cellular processes. To identify potential regulatory or targeting subunits of yeast $PP1_C$, we tagged Glc7p at its amino terminus with protein A and affinity-purified Glc7p protein complexes from yeast. The purified proteins were separated by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) and identified by peptide mass fingerprint analysis using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. To confirm the accuracy of our identifications, peptides from some of the proteins were also sequenced using high-performance liquid chromatography (HPLC) coupled to tandem mass spectrometry. Only four of the Glc7p-associated proteins that we identified (Mhp1p, Bni4p, Ref2p, and Sds22p) have previously been shown to interact with Glc7p, and multiple components of the CPF (cleavage and polyadenylation factor) complex involved in messenger RNA 3′-end processing were present as major components in the Glc7p-associated protein fraction. To confirm the interaction of Glc7p with this complex, we used the same approach to purify and characterize the components of the yeast CPF complex using protein A-tagged Pta1p. Six known components of the yeast (CPF) complex, together with Glc7p, were identified among the Pta1p-associated polypeptides using peptide mass fingerprint analysis. Thus Glc7p is a novel component of the CPF complex and may therefore be involved regulating mRNA 3′-end processing.

Protein phosphatase type 1 (PP1)[1] is a serine/threonine-specific phosphatase that is encoded by the essential gene *GLC7* in *Saccharomyces cerevisiae* (*1*, *2*). The *GLC7* gene encodes the catalytic subunit of PP1 ($PP1_C$), and Glc7p is involved in the regulation of multiple cellular processes including carbohydrate metabolism, mitosis, meiosis, and protein synthesis (see, e.g., ref *3*). The activity of $PP1_C$ is controlled by its association with a range of regulatory or targeting subunits that influence its intracellular localization and substrate specificity, and almost all $PP1_C$ is thought to be present in complexes with such proteins (*4*). Understanding the many cellular roles of PP1 will therefore require identification and characterization of the full range of $PP1_C$ regulators, and model organisms such as yeasts provide well-characterized systems in which this can be achieved. In *S. cerevisiae*, previous studies have identified a number of proteins that interact with Glc7p, many of which represent known or potential Glc7p regulatory subunits (*3*, *4*). These include proteins such as Sds22p (*5*, *6*), Reg1p (*7*, *8*), Reg2p (*9*), Gac1p (*10*), Red1p (*11*, *12*), and Glc8p, the yeast homologue of $PP1_C$ inhibitor-2 (*1*, *13*), each of which has a well-established role as a regulator of Glc7p function.

Recent genome-wide approaches in yeast using the two-hybrid system have identified many more potential Glc7p-interacting proteins (*14*−*16*), but in most cases their interaction with Glc7p has remained unverified by independent approaches. We therefore adopted an alternative strategy to identify novel Glc7p regulatory subunits. Proteins that interact with Glc7p were isolated by affinity purification following in vivo expression of a protein A-tagged version of Glc7p (PrA-TEV-Glc7p). By including a cleavage site for TEV protease between the PrA tag and Glc7p, Glc7p protein complexes could be specifically recovered under mild conditions by TEV protease cleavage following affinity isolation on immobilized immunoglobulin (IgG). Similar approaches have proved highly successful for characterizing protein−protein interactions in both the nuclear pore complex and the spindle pole body (*17*−*19*). We affinity-purified PrA-TEV-Glc7p from yeast cells in either exponential or early stationary phase using IgG-conjugated magnetic beads, separated the Glc7p protein complexes using sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE), and identified individual proteins by peptide mass fingerprint analysis using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. Using this procedure we identified 36 yeast polypeptides that interact

† This work was supported by Medical Research Council Grant G9901184.

* Address correspondence to this author at the Division of Gene Regulation and Expression, School of Life Sciences, MSI/WTB Complex, University of Dundee, Dundee DD1 5EH, U.K. Phone: +44 (0)1382 344250. Fax: +44 (0)1382 344782. E-mail: m.j.r.stark@dundee.ac.uk.

[1] Abbreviations: 5-FOA, 5-fluoroorotic acid; BSA, bovine serum albumin; CFP, cyan fluorescent protein; CPF, cleavage and polyadenylation factor; FRET, fluorescence resonance energy transfer; HPLC, high-performance liquid chromatography; IgG, immunoglobulin; LRR, leucine-rich repeat; MALDI, matrix-assisted laser desorption/ionization; MAP, microtubule-associated protein; ORF, open reading frame; PBS, phosphate-buffered saline; PP1, protein phosphatase 1; $PP1_C$, catalytic subunit of PP1; PrA, protein A tag consisting of two IgG-binding domains; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TOF, time of flight; YFP, yellow fluorescent protein.

10.1021/bi015815e CCC: $22.00
Published on Web 01/19/2002

Table 1: Yeast Strains Used in This Study

strain[a]	genotype	source/ref
AYS927	*MAT***a**/*MAT*α *ade2-1/ade2-1 his3-11, 15/his3-11, 15 leu2-3, 112/leu2-3, 112 trp1-1/trp1-1 ura3-1/ura3-1 can1-100/can1-100 ssd1-d2/ssd1-d2* Gal^+	*60*
PWY1D	*MAT***a**/*MAT*α *ade2-1/ade2-1 his3-11, 15/his3-11, 15 leu2-3, 112/leu2-3, 112 trp1-1/trp1-1 ura3-1/ura3-1 can1-100/can1-100 PTA1/PTA1::PrA::his5*$^+$*ssd1-d2/ssd1-d2* Gal^+	this study
PWY2	*MAT***a** *ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100 PTA1::PrA::his5*$^+$ *ssd1-d2* Gal^+	from PWY1D; this study
LKY150	*MAT***a** *ade2-1 his3-11, 15 leu2-3, 112 ura3-1 can1-100 ssd1-d2 glc7::LEU2 trp1-1::*YIplac204-*PrA-TEV-GLC7* (Trp^+)	*27*
PWY3	*MAT***a** *ade2-1 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 can1-100 ssd1-d2* Gal^+ [pNOPPATA-1L] (Leu^+)	this study
SBY-SS**a**	*MAT***a** *ade2-1 his3-11 leu2-3, 112 trp1-1 ura3 can1-100 ssd1-d2 glc7::LEU2* Gal^+ [YCp-*GLC7*(*URA3*)]	*24*

[a] All strains are in the W303 background.

with Glc7p in protein complexes, including multiple components of the yeast cleavage and polyadenylation factor (CPF) complex (*20*). In the converse experiment, Glc7p was identified among the CPF complex proteins isolated using protein A-tagged Pta1p (Pta1-TEV-PrAp), confirming that Glc7p is a component of the CPF complex. The nature and diversity of Glc7p interacting proteins identified in this study suggest the previously unknown involvement of Glc7p in several key cellular processes, and the approach described is likely to be useful in other large-scale analyses aimed at identifying yeast protein−protein interactions.

MATERIALS AND METHODS

Yeast Strains and General Methods. The *S. cerevisiae* strains used in this study are all derived from the W303 wild-type strain and are listed in Table 1. Yeast strains were grown using standard methods and culture media (*21*) and were grown in at 29 °C in yeast peptone dextrose (YPD) medium unless otherwise indicated. Transformation was performed according to ref *22*, while other standard microbiology and molecular biology techniques were carried out as described by Ausubel et al. (*23*).

Plasmid and Strain Construction. A yeast strain (LKY150) expressing an N-terminal protein A-tagged Glc7 protein (PrA-TEV-Glc7p) was constructed as follows. To generate an integrative plasmid encoding protein A-tagged Glc7p, the HA-tagged *GLC7* construct from YCpHA-*GLC7* (from Kim Arndt: see ref *24*) was excised as a *Hin*dIII−*Msc*I fragment and cloned between the *Hin*dIII and *Sma*I sites of YIplac204 (*25*). A small *Spe*I−*Cla*I fragment at the 5′-end of *GLC7* (carrying the HA epitope tag) was removed and replaced with a larger *Spe*I−*Cla*I fragment carrying a protein A tag and TEV protease cleavage site, generating YIplac204-PrA-*GLC7*. This fragment was made by polymerase chain reaction (PCR) using primers PrA-*GLC7*-5′ and PrA-*GLC7*-3′ with pZZ-His5 (*26*) as a template for the protein A segment, as described in ref *27*. The fragment was cleaved by *Spe*I and *Cla*I prior to subcloning, and the sequence of the amplified region was subsequently verified by DNA sequencing. To generate a strain solely dependent on PrA-TEV-Glc7p for Glc7p function, YIplac204-PrA-*GLC7* was integrated into SBY-SS**a** at the *trp1-1* locus followed by selection on 5-FOA to remove YCp-*GLC7*(*URA3*) as described elsewhere (*27*). A control yeast strain expressing PrA-TEV (PWY3) was constructed by transforming AY925 with plasmid pNOP-PATA-1L (*28*), a YCp carrying PrA-TEV and a *LEU2* selectable marker [*LEU2 CEN NOP1* promoter::protein A (2× IgG binding domains)::TEV site::*ADH1* terminator].

A haploid yeast strain (PWY2) expressing Pta1p tagged with protein A at its C-terminus (Pta1-TEV-PrAp) was generated as follows. A TEV-PrA tag linked to the *Schizosaccharomyces pombe his5*$^+$ marker sequence was fused precisely to the 3′-end of the *PTA1* open reading frame (ORF) by homologous recombination with a suitable PCR product. A 5′-TEV-PrA-His5$^+$-3′ gene tag bracketed by sequences flanking the *PTA1* stop codon was amplified by PCR from pZZ-His5 (*26*) using *Taq* polymerase (Bioline) and primers Pta1-PAtag-For(5′-CTAAAGCAAGAAGATGAAGGCTTACA-CAAGCAGTGCGATTCACTGCTTGACAGGCTAAAAGGAG-CAGGGGCGGGTGC-3′) and Pta1-PAtag-Rev (5′-TGCAA-GTTGAAGGAAGACCCTACACATGCGTATATATGAT-GTATGTAATGGTTGTGATCAGAGGTCGACGGTATC-GATAAG-3′). Each primer has a pZZ-His5 specific sequence (underlined) and a 60-base *PTA1*-specific sequence. A diploid strain heterozygous for *PTA1*-TEV-PrA (PWY1D) was generated by transformation of AYS927 with the 5′-TEV-PrA-*his5*$^+$-3′ PCR product followed by selection for the *his5*$^+$ marker. Transformants heterozygous for *PTA1*-TEV-PrA-*his5*$^+$ were identified by screening of genomic DNA by PCR with suitable primers and by Western blot analysis to detect the fusion polypeptide (data not shown). A haploid PTA1-TEV-PrA strain dependent solely on the modified *PTA1* gene (PWY2) was obtained from PWY1D following sporulation, tetrad dissection, and identification of His^+ segregants. PWY2 was verified by PCR and Western blot analysis as described above.

Affinity Purification of PrA-TEV-Glc7p. LKY150 and PWY3 were grown in 3 L of YPD at 29 °C to either exponential phase (1.2×10^7 cells/mL) or early stationary phase (1.1×10^8 cells/mL). Cultures were centrifuged at 5000*g* for 30 min to pellet the cells, which were resuspended in 30 mL of lysis buffer (100 mM KCl, 0.1% Triton X-100, 0.1 mM EDTA, pH 8.0, 1 mM $MgCl_2$, 50 mM HEPES−KOH, pH 7.5). Complete protease inhibitors (Roche) were added to the lysis buffer according to the manufacturer's instructions. Glass beads (20 mL of 0.4 mm diameter; Sigma) were added to the cell suspension, and the cells were vortexed for 30 min at 4 °C. The suspension was centrifuged at 5000*g* for 5 min at 4 °C to pellet the beads, and then the suspension was cleared by centrifugation for 30 min at 20000*g* at 4 °C. The cleared lysate was frozen in liquid nitrogen and stored at −80 °C.

Proteins were affinity-isolated from cell lysates using immunomagnetic beads. Rabbit γ-globulin antibodies (Jackson Immunoresearch, Stratech, U.K.) were first conjugated to Dynabeads M-280 tosyl-activated magnetic beads (Dynal,

U.K.) according to the manufacturer's instructions (see http://www.dynal.net) using 3 μg of antibody per 10^7 Dynabeads. Antibody-conjugated beads were stored at a final concentration of 2×10^9 beads/mL in Dynal buffer D (PBS, pH 7.4, 0.1% w/v BSA). Before use, the antibody-conjugated beads were washed three times with PBS, once with PBS−Triton (PBS, 1% Triton X-100), and three times with PBS. Cell lysates (from 3 L of culture) were thawed, and 1 mL (2×10^9) of Dynabeads and complete protease inhibitors (two tablets) were added to each sample, followed by a 2 h incubation with rotation at 4 °C. In most experiments, PrA-TEV-Glc7p complexes bound to the beads were then washed three times with 10 mL of lysis buffer, but in some experiments lysis buffer containing either 500 mM KCl or 1.0 M KCl was used for the second and third washes as indicated. After three washes with 10 mL of IP buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% Nonidet P40) and one with 10 mL of TEV cleavage buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.1% NP40, 0.5 mM EDTA, pH 8.0, 1 mM dithiothreitol), the beads were finally resuspended in 1 mL of TEV cleavage buffer containing 10 μL (100 units) of TEV protease (Roche) and then incubated with rotation for 2 h at room temperature. After the beads were removed, samples were concentrated to 40 μL using a Vivaspin 500 ultrafiltration concentrator (Vivascience) by centrifugation at 12000*g* for 30 min at 4 °C and then stored at −80 °C until required.

Affinity Purification of Pta1-TEV-PrAp. Cultures (3 L) of PWY2 and PWY3 haploid yeast strains were grown in YPD at 29 °C to late exponential phase (7×10^7 cells/mL). Protein extracts were prepared and Pta1p-associated proteins affinity purified, concentrated, and stored as described in the previous section, except that 3 mL of IgG-conjugated magnetic beads was used for each protein sample.

MALDI Mass Spectrometry and Protein Identification Affinity-purified protein complexes were separated on SDS−PAGE gels and the proteins analyzed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry as follows. Samples (40 μL) were dried down to a volume of 10 μL in a Speedvac vacuum centrifuge (Savant) and supplemented with 10 μL of 4× Novex sample buffer and 5 μL of 50 mM dithiothreitol in 100 mM NH_4HCO_3. The sample mixture was incubated for 10 min at 70 °C and then cooled prior to addition of 5 μL of 300 mM iodoacetamide in 100 mM NH_4HCO_3. After incubation at room temperature in the dark for 30 min, 10 μL of water was added before gel analysis. Each 40 μL sample was loaded onto a 10% bis-Tris gel (Novex/Invitrogen) with MES running buffer (Novex/Invitrogen). Gels were run at 200 V for 35 min and then stained with Sypro Orange (Molecular Probes) according to manufacturer's instructions, using 10 μL of 5000× concentrate in 50 mL of 10% acetic acid solution. Stained protein bands were visualized either by using a Fuji-imaging system or a dark reader box (Clare Chemical Research). The appropriate bands were excised with a clean scalpel blade and stored at −20 °C.

In-gel digestion of proteins was carried out as follows (modified from ref *29*). Water (300 μL) was added to each gel slice, and the samples were placed onto a shaker for 15 min. Acetonitrile (300 μL) was then added, and samples were incubated for a further 15 min. The solution was then discarded and 300 μL of 100 mM NH_4HCO_3 added to each gel slice, again discarding the solution after shaking for 15 min. After this step was repeated with 300 μL of 50 mM NH_4HCO_3, the gel slice was crushed with a Teflon stick and 100 μL of acetonitrile added. The crushed slice was incubated on a shaker for 10 min, the solution removed, and the gel slice dried in a Speedvac vacuum centrifuge (Savant) for 10 min. Gel slices were rehydrated for 30 min with trypsin digestion buffer [12.5 μg/mL modified bovine trypsin (Roche), 0.1% *n*-octyl glucoside (Calbiochem), 20 mM NH_4HCO_3], and then additional buffer (lacking trypsin) was added if required to cover the gel pieces. After overnight incubation on a shaker at 30 °C, an equal volume of acetonitrile/0.3% trifluoroacetic acid was added, and the slices were incubated at 30 °C on a shaker for 10 min to extract the peptides. One-tenth of the peptide solution from each digest (1−5 μL) was added to a MALDI sample plate and then supplemented with 1 μL of a 10 mg/mL solution of α-cyano-4-hydroxy-*trans*-cinnamic acid matrix (Sigma) in 50% acetonitrile in water containing 0.3% trifluoroacetic, mixed, and allowed to air-dry prior to analysis. MALDI mass spectra were generated using a Voyager DE-STR MALDI-TOF MS system (PerSeptive Biosystems) with delayed extraction in the reflectron mode. A three-point internal calibration was carried out on each spectrum to improve mass accuracy for database interrogation. Protein identification was from a comparison of peak list data generated from the Data Explorer application (PerSeptive Biosystems) against NCBInr (nonredundant) and Swissprot databases using the ProteinProspector V3.4.1 software MS-Fit (http://www.prospector.ucsf.edu).

Tandem Mass Spectrometry and Protein Identification. Peptide sequencing was carried out using capillary high-performance liquid chromatography coupled to a hybrid quadrupole-TOF mass spectrometer (cLC-Q-TOF-MS/MS). The peptide sequence data were obtained from digests separated using a water/acetonitrile/formic acid gradient on an Ultimate capillary HPLC system (LC Packings) with a 75 μm × 15 cm pepMap column (C18, LC Packings) coupled to a hybrid quadrupole time-of-flight system (Q-Tof2, Micromass) with a split flow at 200 nL/min. The HPLC conditions were as follows. A linear gradient of 5−50% over 20 min was formed using buffer A (2% acetonitrile, 0.1% formic acid) and buffer B (80% acetonitrile, 0.1% formic acid). The gradient profile was monitored by measuring the absorbance at 215 nm with a Z-flow cell (3 nL volume) prior to analysis by mass spectrometry. The automated data acquisition facility within Masslynx version 3.5 (Micromass) was used to generate CID-MS spectra. The PepSeq application (Masslynx) was then used to interpret the data and for assignment of sequence tags. Protein assignment was by comparison of sequence tag data against NCBInr and Swissprot databases using the ProteinProspector V3.4.1 software MS-Tag (http://prospector.ucsf.edu).

RESULTS

Purification of PrA-TEV-Glc7p-Associated Proteins. Haploid yeast strains were constructed that express either a protein A-tagged version of Glc7p (PrA-TEV-Glc7p) as their sole source of $PP1_C$ or, as a control, unfused protein A (PrA-TEV) expressed in a strain containing a normal, untagged *GLC7* gene. Both strains showed essentially identical growth kinetics compared to a W303 wild-type strain, and both PrA-

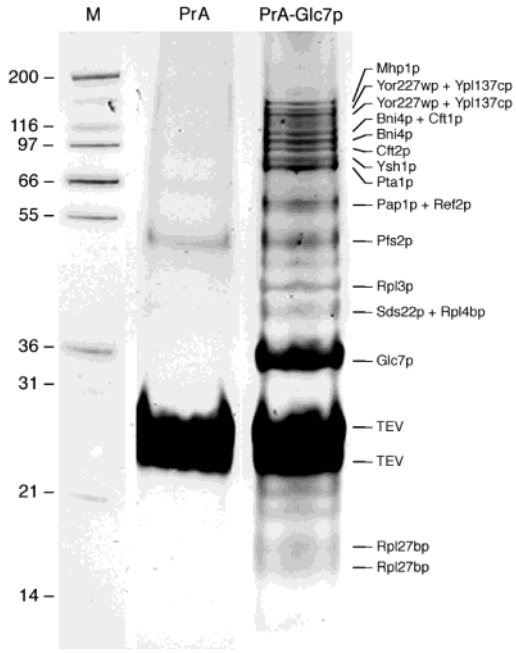

FIGURE 1: Affinity isolation of Glc7p protein complexes from exponentially growing cells. Protein complexes were isolated using IgG-conjugated magnetic beads from whole cell extracts prepared from yeast cultures expressing either PrA-TEV-Glc7p (LKY150) or PrA-TEV (PWY3) as a control. Proteins released from the beads following addition of TEV protease were analyzed by SDS−PAGE electrophoresis, staining the gel with Sypro Orange. Protein bands were then excised and analyzed by MALDI-TOF mass spectrometry as described in Materials and Methods. Proteins identified in each band by mass fingerprint analysis are indicated on the right. Lanes: M, molecular mass standards (sizes in kilodaltons); PrA-Glc7p, proteins from LKY150 eluted from IgG-conjugated beads by treatment with TEV protease (TEV); PrA, unfused protein A control (PWY3).

TEV-Glc7p and PrA-TEV were expressed at similar levels as determined by Western blot analysis (data not shown). These results showed that expression of neither recombinant protein had any obvious detrimental effect on yeast cell growth and indicated that the PrA-TEV-Glc7p is functional in vivo.

The PrA-TEV-Glc7p fusion protein was affinity-purified from yeast cell extracts using IgG-conjugated magnetic beads, to which protein A or protein A-tagged polypeptides bind with high affinity. The unfused PrA-TEV protein was purified from yeast cell extracts to provide a control for proteins that bound nonspecifically either to the beads or to protein A under the conditions used. Separate experiments were carried out using either exponential phase or early stationary phase cells, since Glc7p might interact with different subsets of proteins under distinct physiological conditions. We therefore hoped that this strategy would enable us to identify a wider range of Glc7p-interacting proteins. The purified proteins from the exponential and stationary phase cultures were separated by SDS−PAGE as shown in Figures 1 and 2, respectively. Protein bands detected in the PrA-TEV-Glc7p purification that were not present in the PrA-TEV control were therefore likely candidates for Glc7p-associated polypeptides or members of Glc7p-containing protein complexes. In some experiments such as that shown in Figure 2, additional bands were seen that corresponded to BSA carried over from the IgG-conjugated bead storage buffer. In addition, in all experiments the TEV protease used to elute the protein complexes appeared as two major bands of approximate molecular mass 27 and 29 kDa that might obscure proteins at these locations. We therefore attempted removal of these nonspecific proteins from the purified protein samples using the TAP tag strategy, which involves the use of tandem protein A and calmodulin binding domain (CBP) tags and a second purification step (*30*). Tandem protein A-TEV and CBP tags were added to the N-terminus of Glc7p (PrA-TEV-CBP-Glc7p; data not shown). We found that the double-tagged protein also exhibited normal expression and function in yeast cells and enabled effective removal of the contaminating TEV protease. Although no additional bands with similar migration to TEV protease were apparent as a result, the additional optimization and washing steps involved with the tandem tag resulted in a considerable reduction in the recovery of the Glc7p-associated proteins without apparently improving the quality of the sample (data not shown). The use of a single PrA tag therefore proved to be the most effective way to achieve a good yield while maintaining a low level of nonspecific protein contamination.

Identification of Glc7p-Interacting Proteins Using Mass Spectrometry. The purified proteins separated using SDS−PAGE were identified by peptide mass fingerprint analysis using MALDI-TOF mass spectrometry. A total of 37 polypeptides (including Glc7p) were identified from the experiments in this study: 16 were identified in the exponential phase experiment and 33 were identified in the early stationary phase experiment, as indicated in Figures 1 and 2A, respectively. Representative peptide mass fingerprints for three of the polypeptides (Mhp1p, Pta1p, and Glc7p) are shown in Figure 3. The presence of BSA contamination in the early stationary phase protein sample may have obscured some of the protein bands in the SDS−PAGE gel, and some proteins may consequently have escaped identification. Some of the proteins we identified were found in both exponential and early stationary phase samples while others were only found under one of the two conditions. The proteins identified are summarized in Table 2 (exponential phase) and Table 3 (early stationary phase). To check the accuracy of the mass fingerprint analysis using MALDI-TOF mass spectrometry, peptide mass fingerprint identities were confirmed by subsequent analysis of representative protein samples using capillary HPLC coupled with tandem mass spectrometry to produce sequence tags. Figure 4 shows data for Mhp1p (Figure 4A), Pta1p (Figure 4B), and Glc7p (Figure 4C), all of which yielded amino acid sequence tags that confirmed the protein identifications derived by fingerprint analysis.

Almost all of the 36 polypeptides that we identified in this study are the products of characterized genes whose likely function is already known, and over half of the Glc7p-associated proteins were ribosomal proteins (Tables 2 and 3). However, all but one of these ribosomal proteins (Rpp0p) are highly basic, raising the possibility that their association with Glc7p was not specific despite the inclusion of detergents to discourage such nonspecific interactions. We therefore examined whether any of the Glc7p-associated proteins were removed by washing in 0.5 and 1.0 M salt as an additional test of specificity. Figure 2B shows that essentially all of the bands corresponding to ribosomal

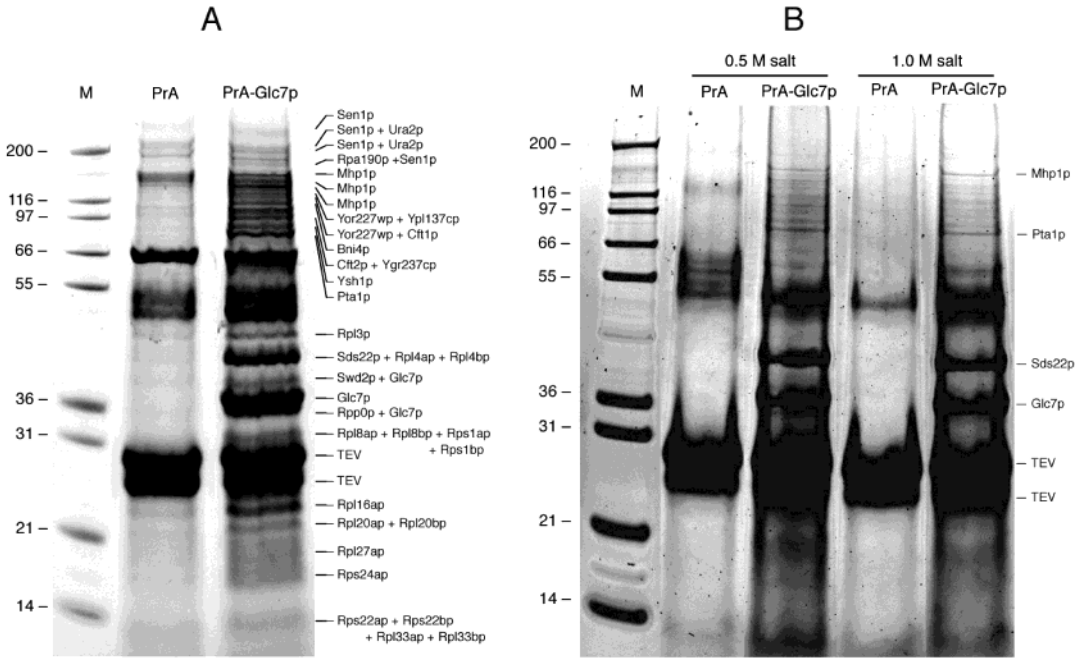

FIGURE 2: Affinity isolation of Glc7p protein complexes from early stationary phase cells. Protein complexes were isolated using IgG-conjugated magnetic beads from whole cell extracts prepared from yeast cultures expressing either PrA-TEV-Glc7p (LKY150) or PrA-TEV (PWY3) as a control. Proteins released from the beads following addition of TEV protease were analyzed by SDS−PAGE electrophoresis, staining the gel with Sypro Orange. Protein bands were then excised and analyzed by MALDI-TOF mass spectrometry as described in Materials and Methods. Proteins identified in each band by mass fingerprint analysis are indicated on the right. The major protein bands at approximately 50, 55, 66, 120, 180, and 240 kDa are contaminating BSA that was carried over from the buffer used to store the IgG-conjugated beads. (A) Protein complexes washed in lysis buffer containing 100 mM KCl before release by TEV protease. (B), Protein complexes washed in lysis buffer containing 0.5 or 1.0 M KCl as indicated before release by TEV protease. Lanes: M, molecular mass standards (sizes in kilodaltons); PrA-Glc7p, proteins from LKY150 eluted from IgG-conjugated beads by treatment with TEV protease (TEV); PrA, unfused protein A control (PWY3).

proteins were removed by this treatment, whereas in comparison, an almost identical pattern of nonribosomal protein-containing bands remained, albeit with reduced yield. Among these, Mhp1p, Pta1p, and Sds22p were reconfirmed by MALDI-MS as major components (Figure 2B). Although previous studies indicated that rat liver $PP1_C$ interacts with ribosomal protein L5 and associates with ribosomes in an L5-dependent manner (*31*), we did not find the yeast homologue of this protein (Rpl5p/YPL131w product) in our Glc7p-associated fraction. Given that almost all of the ribosomal protein bands were removed by the high-salt wash, they most likely associated with Glc7p in a nonspecific manner. One possible exception to this is Rpp0p, the band corresponding to which remained after the 0.5 M salt wash (Figure 2B and data not shown).

Of the well-established yeast $PP1_C$ regulatory subunits (see ref *3*), only Sds22p was present among the Glc7p-associated polypeptides. Sds22p forms a stable, 1:1 complex with Glc7p that can be demonstrated in a variety of ways (*5*, *6*, *27*, *32*) and is required for the normal nuclear localization of Glc7p (*27*). Interestingly, Sds22p appeared to be a much more abundant partner of Glc7p in early stationary phase extracts, and the Sds22p band was more difficult to detect following SDS−PAGE analysis of the exponential phase Glc7p complexes (compare Figures 1 and 2), a finding that was reproducible over several experiments (data not shown).

We identified a total of nine Glc7p-associated polypeptides involved in RNA transcription and processing. Seven of these proteins (Cft1p, Cft2p, Ysh1p, Pta1p, Pap1p, Ref2p, and Pfs2p) are involved in mRNA 3′-end processing (*20*, *33*). With the exception of Ref2p, all of these proteins are known components of the yeast cleavage and polyadenylation factor (CPF) complex (*20*, *34*). Since it is very unlikely that Glc7p interacts directly with each individual member of the CPF complex, our data suggest that it is an integral component of CPF. Although Pap1p and Pfs2p were only identified from exponential phase cells, this was most likely because BSA obscured them in the early stationary phase extracts (see above). Two other proteins involved in RNA transcription and processing that were identified are Rpa190p and Sen1p. Rpa190p is the major subunit of RNA polymerase I which is involved in ribosomal RNA gene expression (*35*), while Sen1p is a putative RNA helicase involved in nuclear RNA processing of pre-transfer RNA, pre-ribosomal RNA, small nuclear RNA, and small nucleolar RNA (*36*−*38*). Neither of these proteins has previously been found to associate with $PP1_C$. However, it is interesting to note that, like Rpa190p (*35*), Glc7p is also enriched in the nucleolus (*39*). Yeast RNA polymerase I is also known to be phosphorylated (*40*), raising the possibility that the PP1 with which it is associated could regulate its phosphorylation state.

We also identified a family of three related polypeptides in Glc7p-associated protein fractions, namely, Mhp1p, Yor227p, and Ypl137p. Mhp1p is 25% identical to Yor227p and 27% identical to Ypl137p, while Ypl137p and Yor227p are more closely related to each other (43% identity). Ypl137p and Yor227p are currently of unknown function, but Mhp1p is a microtubule-binding protein functionally related to the *Drosophila* 205K MAP and is apparently involved in microtubule function (*41*, *42*). The sequence similarity between these three proteins suggests that they may be involved in similar processes or share overlapping

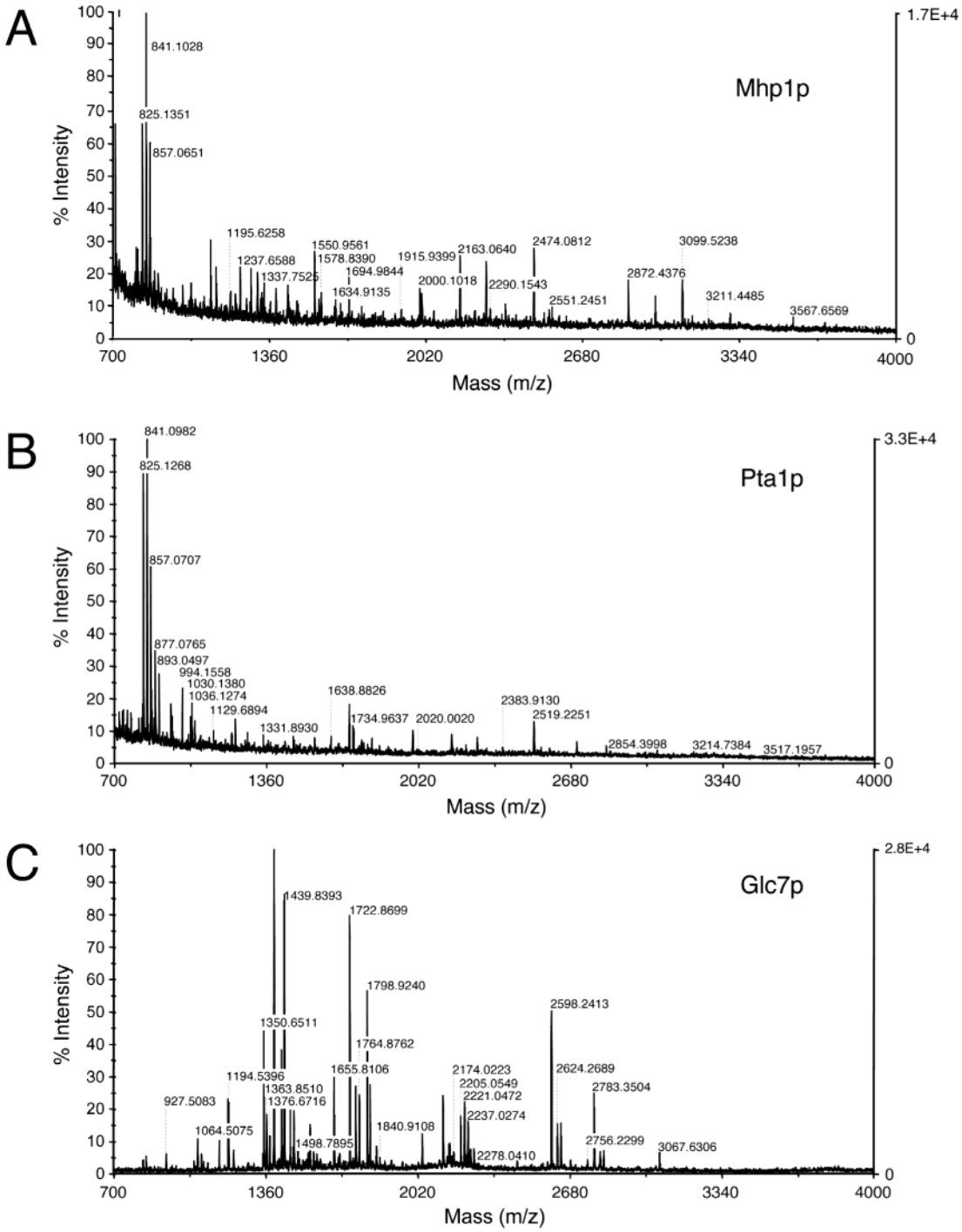

FIGURE 3: Identification of affinity-isolated proteins using MALDI-TOF mass spectrometry. Tryptic peptide mass fingerprints of (A) Mhp1p, (B) Pta1p, and (C) Glc7p.

functions. Mhp1p has already been identified in a genome-wide yeast two-hybrid screen as a candidate Glc7p-interacting protein (*16*) and in our hands was one of the most abundant and salt wash resistant Glc7p-associated proteins (Figure 2B). Our data therefore provide strong evidence for a genuine interaction between Mhp1p and Glc7p, while the fact that all three members of the Mhp1p protein family copurified with Glc7p underscores the significance of this finding. Another cytoskeletal protein identified in the Glc7p-associated polypeptides is Bni4p, a protein that has been shown to be involved in cell wall maintenance and cytokinesis (*43*). Bni4p interacts with the septin ring at the bud neck, with Chs4p (an activator of chitin synthase III; *43*) and with Glc7p in the yeast two-hybrid system (*16*). Furthermore, Glc7p and Bni4p can be co-immunoprecipitated from yeast cell extracts, and their close proximity in the cell can be demonstrated by FRET using CFP- and YFP-tagged proteins (L. Kozubowski and K. Tatchell, personal communication), clearly establishing Bni4p as a Glc7p-interacting protein.

We also identified Ura2p, a protein that catalyzes the initial steps of pyrimidine biosynthesis in yeast (*44*), as a Glc7p-associated protein. Ura2p is a 242 kDa multifunctional protein that contains carbamyl phosphate synthetase and aspartate transcarbamylase domains but which has not previously been found associated with $PP1_C$. Finally, we

Table 2: Summary of Glc7p-Associated Proteins Identified in Exponential Phase Cells[a]

ORF name	gene product	molecular mass (kDa)	peptides matched (%) [b]	sequence coverage (%) [b]	protein function	RVXF motif
YJL042W[c]	Mhp1p	155.2	29	22	cytoskeleton	yes
YDR301W	Cft1p	153.4	10	17	RNA 3′-processing	yes
YPL137C	Ypl137cp	140.8	17/13	19/14	cytoskeleton	yes
YOR227W	Yor227wp	139.3	21/17	20/24	cytoskeleton	yes
YNL233W[c]	Bni4p	100.4	11/9	29/18	cytoskeleton	no[d]
YLR115W	Cft2p	96.1	15	21	RNA 3′-processing	no
YLR277C	Ysh1p	87.6	22	28	RNA 3′-processing	no
YAL043C[c]	Pta1p	88.3	9	20	RNA 3′-processing	yes
YKR002W	Pap1p	64.4	11	24	RNA 3′-processing	no
YDR195W[c]	Ref2p	59.8	11	25	RNA 3′-processing	yes
YNL317W	Pfs2p	53.2	6	17	RNA 3′-processing	yes[e]
YOR063W	Rpl3p	43.7	16	33	80S ribosome	no
YDR012W	Rpl4bp	38.9	9	38	80S ribosome	no
YKL193C[c]	Sds22p	38.9	7	66	PP1 regulator	no
YER133W	Glc7p	35.9	33	60	PP1	no
YDR471W	Rpl27bp	15.4	12/11	33/33	80S ribosome	no

[a] References for these proteins can be found at the Yeast Proteome Database website (http://www.proteome.com). [b] Multiple numbers in these columns represent values for the same protein identified in different gel protein bands (values are shown for the highest molecular mass bands first; see Figure 1). [c] Proteins for which there exist other data indicative of a potential interaction with Glc7p. [d] Contains a candidate motif lacking an upstream basic residue. [e] Contains a candidate motif lacking a downstream acidic residue.

Table 3: Summary of Glc7p-Associated Proteins Identified in Early Stationary Phase Cells[a]

ORF name	gene product	molecular mass (kDa)	peptides matched (%) [b]	sequence coverage (%)[b]	protein function	RVXF motif
YLR430W	Sen1p	252.5	30/15/12/26	16/10/10/15	RNA processing	no
YJL130C	Ura2p	244.9	16/22	7/14	biosynthesis	no
YOR341W	Rpa190p	186.4	25	13	RNA polymerase I	yes
YJL042W[c]	Mhp1p	155.2	32/13/12	23/10/10	cytoskeleton	yes
YDR301W	Cft1p	153.4	13	15	RNA 3′-processing	yes
YPL137C	Ypl137cp	140.8	20	21	unknown	yes
YOR227W	Yor227wp	139.3	20/15	21/18	unknown	yes
YNL233W[c]	Bni4p	100.4	14	31	cytoskeleton	no[d]
YLR115W	Cft2p	96.1	9	16	RNA 3′-processing	no
YGR237C[c]	Ygr237cp	89.1	9	22	unknown	yes
YLR277C	Ysh1p	87.6	31	46	RNA 3′-processing	no
YAL043C[c]	Pta1p	88.3	31	47	RNA 3′-processing	yes
YOR063W	Rpl3p	43.7	21	36	80S ribosome	no
YBR031W	Rpl4ap	39	11	30	80S ribosome	no
YDR012W	Rpl4bp	38.9	11	30	80S ribosome	no
YKL193C[c]	Sds22p	38.9	20	36	PP1 regulator	no
YKL018W	Swd2p	36.9	10	29	unknown	no
YER133W	Glc7p	35.9	25/30/11	49/67/30	PP1	no
YLR340W	Rpp0p	33.6	12	26	80S ribosome	no
YLR441C	Rps1ap	28.7	9	24	80S ribosome	no
YML063W	Rps1bp	28.7	9	24	80S ribosome	no
YHL033C	Rpl8ap	28	15	39	80S ribosome	no
YLL045C	Rpl8bp	28	15	39	80S ribosome	no
YIL133C	Rpl16ap	22.1	11	38	80S ribosome	no
YMR242C	Rpl20ap	21.3	8	43	80S ribosome	no
YOR312C	Rpl20bp	20.7	8	44	80S ribosome	no
YDR418W	Rpl12bp	15.8	6	43	80S ribosome	no
YHR010W	Rpl27ap	15.4	7	58	80S ribosome	no
YER074W	Rps24ap	15.3	12	29	80S ribosome	no
YJL190C	Rps22ap	14.5	9	59	80S ribosome	no
YLR367W	Rps22bp	14.5	9	59	80S ribosome	no
YPL143W	Rpl33ap	12.1	14	54	80S ribosome	no
YOR234C	Rpl33bp	12.1	12	32	80S ribosome	no

[a] References for these proteins can be found at the Yeast Proteome Database website (http://www.proteome.com). [b] Multiple numbers in these columns represent values for the same protein identified in different gel protein bands (values are shown for the highest molecular mass bands first; see Figure 2). [c] Proteins for which there exist other data indicative of a potential interaction with Glc7p. [d] Contains a candidate motif lacking an upstream basic residue.

identified two other proteins of as yet unknown function in the purified Glc7p complexes, Ygr237p and Swd2p. Swd2p contains four WD40 domains (*45*) and again has not previously been found associated with $PP1_C$. Since we were unable to identify distinct bands corresponding to Ura2p and Swd2p after washing in either 0.5 or 1 M salt (Figure 2B), it is not possible to say whether these interactions with Glc7p are likely to be specific. However, Ygr237p has previously been detected in HA-tagged Glc7p immunoprecipitates (K. Arndt, personal communication), and this, together with our current data, supports the notion that it is specifically associated with $PP1_C$.

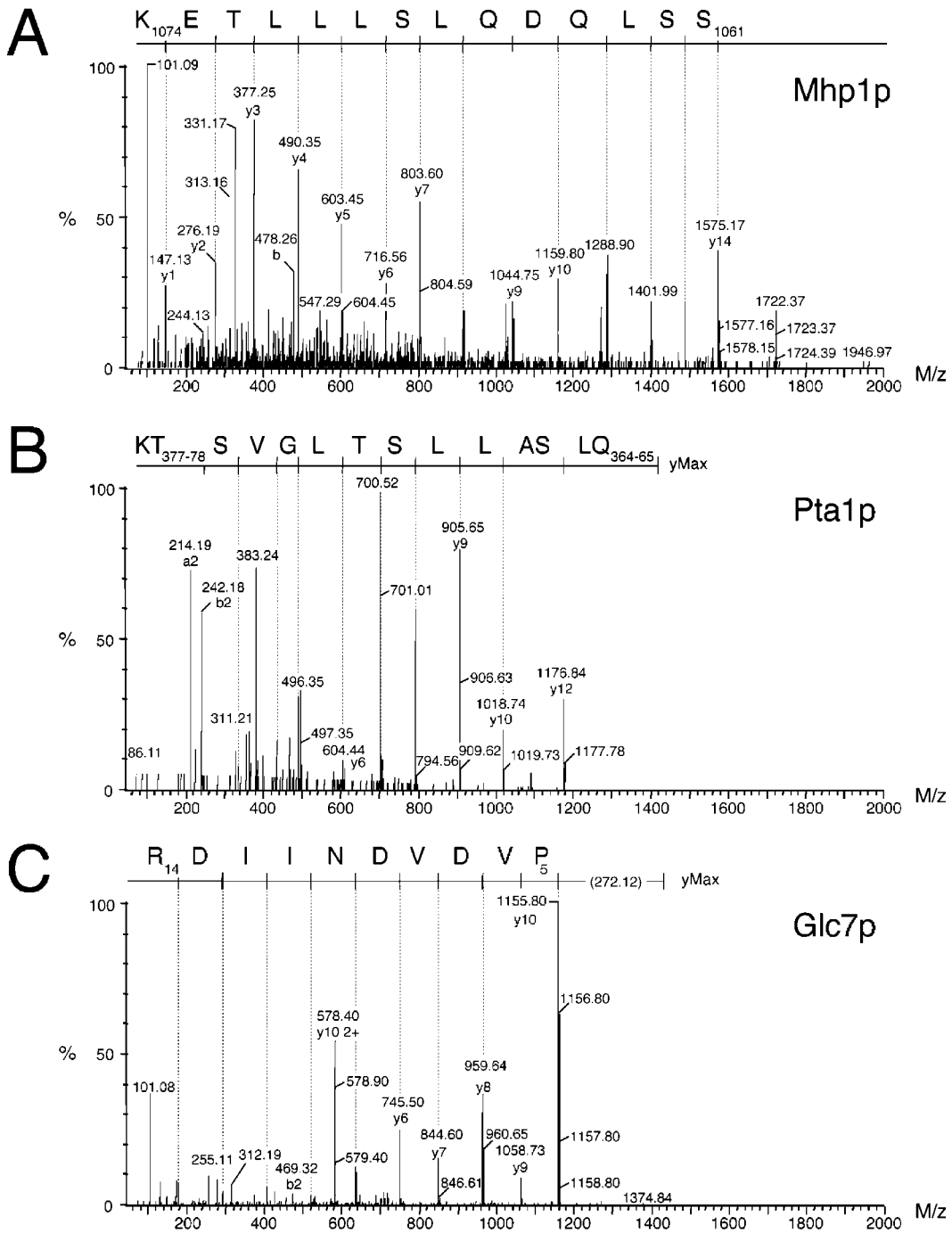

FIGURE 4: Identification of affinity- purified proteins using capillary HPLC coupled to tandem mass spectrometry. Amino acid sequence tags were identified for (A) Mhp1p, (B) Pta1p, and (C) Glc7p. Note that tandem mass spectrometry cannot discriminate between the isobaric amino acids Q/K or I/L. The peptide parent *m*/*z* values (for $[M + 2H]^{2+}$) were 1026.64, 709.51, and 714.46, respectively.

Identification of Pta1p-Interacting Proteins. The finding that the CPF complex is associated with Glc7p was unexpected and suggests the interesting possibility that PP1 may regulate some aspect of mRNA 3′-end processing. To determine whether Glc7p is a genuine component of the mRNA 3′-end processing CPF complex, we therefore adopted a similar strategy to identify the proteins found within this complex. Isolation of the CPF complex was carried out by affinity purification of Pta1p, a well-established CPF component (*20*, *46*). A haploid yeast strain (PWY3) was constructed that expressed a protein A-tagged version of Pta1p as its sole source of Pta1p (PTA1-TEV-PrA). PWY3 showed growth kinetics similar to that of a W303 wild-type strain, indicating that expression of the recombinant protein did not adversely affect growth and that PrA-tagged Pta1p is functional in vivo (data not shown). The Pta1p-TEV-PrA fusion protein was purified from yeast cell extracts using IgG-conjugated magnetic beads, and after cleavage with TEV protease, eluted proteins were compared using SDS−PAGE to those obtained from the control strain (expressing unfused PrA-TEV) under the same conditions. Proteins that were found in the Pta1p-TEV-PrA purification but not in the PrA-TEV control were then identified by peptide mass fingerprint analysis using MALDI-TOF mass spectrometry (Figure 5). We identified a total of 16 polypeptides as summarized in Table 4. These included all of the

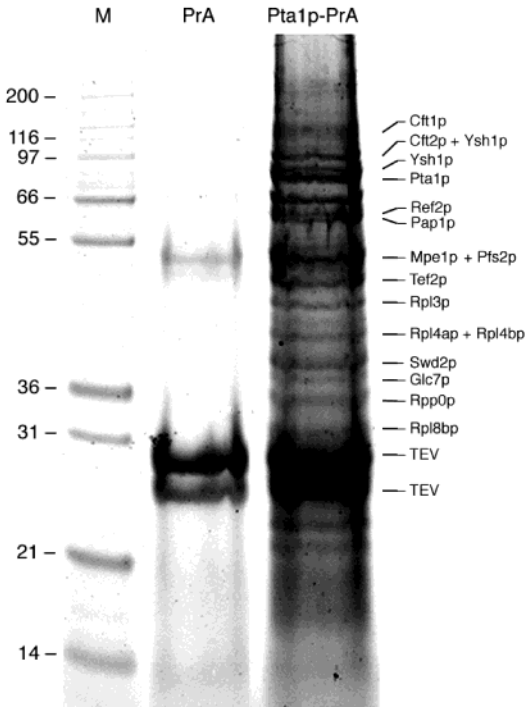

FIGURE 5: Affinity isolation of Pta1p protein complexes from yeast cells. Protein complexes were isolated using IgG-conjugated magnetic beads from whole cell extracts prepared from yeast cultures expressing either Pta1p-TEV-PrA (PWY2) or PrA-TEV (PWY3) as a control. Proteins released from the beads following addition of TEV protease were analyzed by SDS−PAGE electrophoresis, staining the gel with Sypro Orange. Protein bands were then excised and analyzed by MALDI-TOF mass spectrometry as described in Materials and Methods. Proteins identified in each band by mass fingerprint analysis are indicated on the right. Lanes: M, molecular mass standards (sizes in kilodaltons); Pta1p-PrA, proteins from PWY2 eluted from IgG-conjugated beads by treatment with TEV protease (TEV); PrA, unfused protein A control (PWY3).

CPF complex components that we had previously identified as Glc7p-associated proteins, namely, Cft1p, Cft2p, Ysh1p, Pta1p, Pap1p, and Pfs2p, with the additional presence of Mpe1p. Strikingly, Glc7p was also identified within the Pta1p-associated polypeptides (Figure 5), confirming its presence as a component of the CPF complex. We also identified seven other proteins, namely, Ref2p, Swd2p, EF1α (Tef2p), and four ribosomal proteins (Rpl3p, Rpl4ap, Rpp0p, and Rpl8bp) that were also present in the Glc7p-associated protein samples. The significance of finding ribosomal proteins or EF1α associated with the CPF complex is unclear, and at least in the former case, they may represent nonspecific contaminants that were isolated because of their abundance and highly charged nature as discussed above. However, Swd2p is not highly charged and may therefore represent a new CPF-associated polypeptide, consistent with its identification in the Glc7p-associated fraction. Ref2p is also a very basic protein (p*I* 10.43) but, unlike the ribosomal proteins, has already been found to be required for efficient mRNA 3′-end processing (*33*, *47*), albeit not as an overt component of CPF. We have therefore identified one novel component of CPF (Glc7p) and two strong candidates for novel CPF-associated polypeptides (Ref2p and Swd2p).

DISCUSSION

Glc7p-Associated Proteins. We report in this paper the analysis of proteins that interact with *S. cerevisiae* $PP1_C$ as a means to identify candidate PP1 regulatory and targeting subunits. Previous studies using genetic analysis, yeast two-hybrid screens, and immunoprecipitation have identified a number of potential Glc7p regulatory subunits (see refs *3* and *14*−*16*). We chose to carry out a systematic search for proteins that interact with Glc7p using affinity purification and a proteomic-based strategy as a complementary approach to these studies. A version of Glc7p carrying a protein A tag was affinity-purified from yeast, and the Glc7p-associated proteins were identified by peptide mass fingerprint analysis using MALDI-TOF mass spectrometry. The accuracy of these protein identifications was confirmed by amino acid sequence tag analysis of representative protein samples using capillary HPLC coupled to tandem mass spectrometry. Of the 36 polypeptides that we identified about half were highly basic ribosomal proteins (Tables 2 and 3). Since most of these potential interactions were lost when Glc7p complexes were washed in high salt (Figure 2B), they probably reflect nonspecific interactions with Glc7p. Among the remaining polypeptides we identified one known regulatory subunit of PP1, two known cytoskeletal proteins, eight proteins involved in RNA processing, one protein involved in ribosomal RNA gene expression, one protein involved in nucleotide metabolism, and several proteins whose functions have not yet been determined.

An initially surprising result of this study was that while many novel Glc7p-associated polypeptides were identified, only one of the well-established regulatory subunits of yeast $PP1_C$ was found, namely, Sds22p. Reg1p (*7*, *8*), Reg2p (*9*), Gac1p (*10*), Red1p (*11*, *12*), and Glc8p (*1*, *13*) are all known both to interact directly with Glc7p and to regulate its function in a range of cellular processes, and yet none were found using our approach. Since $PP1_C$ regulatory subunits could be present at varying levels in the cell and might bind to Glc7p with a range of affinities, the simplest explanation is that we have identified those proteins present in the most abundant and stable complexes with Glc7p under our chosen growth conditions. In support of this, many of the above proteins are known to be of low abundance or not to be expressed in haploid cells undergoing mitotic growth. For example, Gac1p is expressed mainly after the diauxic shift, and it is undetectable even by Western blot analysis when present in single copy (*48*), while Red1p is the product of a meiosis-specific gene that is expressed in diploid cells during sporulation (*49*). In addition, we also did not detect Bud14p/Fun2p, a protein that we have independently shown to interact with Glc7p and which is highly unstable, being readily detectable by Western blot analysis only in extracts from *pep4* strains that lack vacuolar proteases (A. Engles and M. J. R. Stark, unpublished data). In addition to identifying Sds22p, we have also verified the association of yeast $PP1_C$ with several other potential Glc7p partners identified previously through genome-wide two-hybrid screening, namely, Mhp1p, Bni4p, and Ref2p (*14*−*16*). With the exception of the ribosomal proteins whose association with Glc7p is probably largely nonspecific (see above), there were relatively few differences between the two types of protein sample analyzed in this work. Pap1p, Pfs2p, and Ref2p were only found associated with Glc7p in exponential phase cells, while Sen1p, Ura2p, Rpa190p, Ygr237p, and Swd2p were only identified from early stationary phase extracts. It is not yet clear whether such differences might be significant.

Table 4: Pta1p-Associated Proteins[a]

ORF name	gene product	molecular mass (kDa)	peptides matched (%)[b]	sequence coverage (%)[b]	protein function	RVXF motif
YDR301W[c]	Cft1p	153.4	29	22	RNA 3′-processing	yes
YLR115W[c]	Cft2p	96.1	28	34	RNA 3′-processing	no
YLR277C[c]	Ysh1p	87.6	59	63	RNA 3′-processing	no
YAL043C[c]	Pta1p	88.3	17	25	RNA 3′-processing	yes
YKR002W[c]	Pap1p	64.4	45	25	RNA 3′-processing	no
YDR195W	Ref2p	59.8	15	16	RNA 3′-processing	yes
YNL317W[c]	Pfs2p	53.2	13	37	RNA 3′-processing	yes[d]
YKL059C[c]	Mpe1p[e]	49.5	14	41	RNA 3′-processing	no
YBR118W	Tef2p	49.9	9	40	EF1α	yes
YOR063W	Rpl3p	43.7	15	40	80S ribosome	no
YBR031W	Rpl4ap	39	23	39	80S ribosome	no
YDR012W	Rpl4bp	38.9	20	34	80S ribosome	no
YKL018W	Swd2p	36.9	9	33	unknown	no
YER133W	Glc7p	35.9	16	46	PP1	no
YLR340W	Rpp0p	33.6	16	32	80S ribosome	no
YLL045C	Rpl8bp	28	11	19	80S ribosome	no

[a] References for these proteins can be found at the Yeast Proteome Database website (http://www.proteome.com). [b] Multiple numbers in these columns represent values for the protein identified in different gel protein bands (high to low molecular mass bands). [c] Known components of the CPF complex. [d] Contains a candidate motif lacking a downstream acidic residue. [e] YKL059C encodes Mpe1p (*53*), which is identical to the 58 kDa CPF protein termed Pfs1p in refs *46* and *58* (B. Dichtl and W. Keller, personal communication). The name Pfs1p has been assigned to a different polypeptide (the YHR185C product) in the current yeast databases.

Table 5: Candidate Glc7p Interacting Proteins That Contain a Putative Consensus RVXF Motif[a]

ORF name	gene product	protein function	hits	candidate RVXF motif	position in sequence: start	position in sequence: end	ref
YOR341W[b]	Rpa190p	RNA polymerase I	1	RYAISVSFRHLD	1572	1583	
YDR301W[b]	Cft1p	RNA 3′-processing	2	KIESTTIAFVKE	265	276	
				RRLTTIKFD	647	655	
YJL042W[b,c]	Mhp1p	cytoskeleton	1	RPLKHVSFATN	444	454	*16*
YPL137C[b]	Ypl137cp	unknown	1	KRVSFAVD	283	290	
YOR227W[b]	Yor227wp	unknown	2	RLPSIKFLD	744	752	
				KRIHFD	778	783	
YNL233W[c]	Bni4p	cytoskeleton	1	NNDQGVRFSSQ[d]	826	836	*16*
YGR237C[b,c]	Ygr237cp	unknown	1	KTFTKHISFD	5	14	
YAL043C[b,c]	Pta1p	RNA 3′-processing	1	KLATIKFISE	166	175	
YDR195W[b,c]	Ref2p	RNA 3′-processing	1	KARISSIKFLD	366	376	*40*
YAR014C[c]	Fun2p	unknown	1	KGNKSVSFND	372	381	*16*
YBR045C[c]	Gip1p	PP1 regulator	1	RSTSSVRFD	487	495	*11*
YER054C[c]	Gip2p	PP1 regulator	1	RSKSVHFD	218	225	*11*
YDR028C[c]	Reg1p	PP1 regulator	1	KNRHIHFND	462	470	*7*
YBR050C[c]	Reg2p	PP1 regulator	1	KPRERHIKFND	163	173	*9*
YDR130C[c]	Fin1p	unknown	1	KSVKFKLPD	238	246	*16*
YOR178C[c]	Gac1p	PP1 regulator	1	KNVRFAIE	69	76	*48*
YOR329C[c]	Scd5p	vesicular transport	1	KSKKVRFSE	270	278	*11*

[a] References for these proteins can be found at the Yeast Proteome Database website (http://www.proteome.com). [b] Proteins that were identified as part of Glc7p complexes in our study. [c] Proteins for which there exist other data indicative of a potential interaction with Glc7p. [d] Candidate motif lacking an upstream basic residue.

Many $PP1_C$ regulatory subunits bind to a hydrophobic groove in $PP1_C$ located on the opposite face of the protein to its active site. $PP1_C$ binding is mediated via a short consensus sequence termed the R/K-V/I-X-F motif (*50*, *51*), hereafter referred to as the RVXF motif, and this is also the case for Glc7p (*52*). Using the search consensus [KR] X{0,4} [RHQKAMNST] [VI] [RHSATMK] [FW] X{0,3} [END] derived from ref *50*, putative RVXF motifs were found in a number of the Glc7p-associated proteins identified in this study, as well as in many of the regulatory subunits of Glc7p that have already been identified (Table 5). While the presence of such a motif clearly cannot be taken in isolation as evidence for a direct interaction with Glc7p, it is nonetheless striking that, of the eight proteins identified in our study that contain a putative consensus RVXF motif, for five there is additional evidence to support their association with Glc7p (Tables 2−4), while two more are highly related to another protein (Mhp1p) for which such evidence exists.

A Family of Related, Novel Glc7p-Associated Polypeptides. Three of the Glc7p-interacting polypeptides that we have identified form a family of related proteins, namely, Mhp1p, Ypl137p, and Yor227p, one of which (Mhp1p) has previously been shown to bind to microtubules both in vitro and in yeast cells (*42*). While a potential tubulin-binding domain in the C-terminal region of Mhp1p does not appear to be conserved in the other two polypeptides (*42*), the proteins nonetheless show considerable sequence similarity and may therefore share one or more common functions. Given that all three were found to interact with Glc7p in our study, we examined whether they shared any potential $PP1_C$-interacting domain(s). The RVXF motifs detected in

Mhp1p and Ypl137p are indeed located within a conserved region in the amino-terminal portion of the proteins. Although the RVXF motifs of Yor227p listed in Table 5 are not located in this amino-terminal region, Yor227p does share the amino-terminal motif found in the other two polypeptides but with the phenylalanine residue substituted by a less favorable isoleucine (not shown). This region could therefore constitute a conserved $PP1_C$-interacting region. All three proteins also share a domain that contains several imperfect leucine-rich repeat (LRR) sequences. Sds22p is an LRR protein that lacks any obvious RVXF motifs (*5*, *6*), and the LRRs may mediate its interaction with $PP1_C$ (see ref *27*), so it is also possible that the LRR-related sequences in Mhp1p, Ypl137p, and Yor227p might be involved in their interaction with Glc7p. It will be interesting to examine which domains in these proteins are required for the PP1 interaction and to establish how their interaction with Glc7p relates to their function.

The CPF Complex. We identified six known components of the CPF (cleavage and polyadenylation factor) complex as prominent bands in our Glc7p-associated protein samples, namely, Cft1p, Cft2p, Ysh1p, Pta1p, Pap1p, and Pfs2. This complex, which is involved in mRNA 3′-end processing, has previously been shown to contain nine polypeptides (*20*, *34*), but in the Glc7p purification experiments we did not identify Fip1p, Mpe1p, or Yth1p. It is possible that Yth1p (25 kDa) was present in the purified protein samples but missed because it has the same molecular mass as the TEV protease bands (Figures 1 and 2). However, this potential problem is unlikely to have affected the identification of Fip1p (a 36 kDa protein that migrates with an apparent mass of >50 kDa; *46*) or Mpe1p (~58 kDa; *53*). To confirm that Glc7p is indeed a member of the CPF complex, we affinity-purified Pta1p, one of its major components (*20*, *46*), and identified the Pta1p-associated proteins. We recovered all of the CPF complex components found associated with Glc7p as Pta1p-associated polypeptides and additionally identified Mpe1p, although Fip1p and Yth1p were still not apparent (Table 4). Fip1p was one of the least abundant members of the CPF complex found associated with either Pfs2p or Mpe1p (*53*), suggesting perhaps that it is just less stably associated with the complex and more easily lost during isolation. Importantly, Glc7p was clearly recovered in association with protein A-tagged Pta1p, demonstrating that it is a bona fide member of the CPF complex. Given the large number of components present in the CPF complex, it is unlikely that all of them interact directly with Glc7p. Of the proteins listed in Table 4, Ref2p contains a candidate PP1 binding motif and has also been shown to interact with Glc7p by two-hybrid analysis (*16*), making it a good candidate to mediate the interaction between PP1 and CPF. However, two of the other CPF proteins listed in Table 4 also contain potential PP1 binding domains. One of these (Pta1p) was the most prominent CPF member remaining in association with Glc7p after being washed in 1 M salt (Figure 2B) and was also the only CPF component found to be co-immunoprecipitated with HA-tagged Glc7p (K. Arndt, personal communication). Pta1p is therefore perhaps the best candidate for mediating the association of Glc7p with CPF. Our finding that Glc7p is a component of CPF suggests that PP1 might be involved in the regulation of mRNA 3′-end processing, but the exact role of Glc7p in this process requires further analysis. However, it is interesting to note that the poly(A) addition activity of Pap1p in vitro is downregulated by phosphorylation but can be restored by phosphatase treatment with PP1 (*54*). Further experiments will be necessary to determine whether CPF-associated PP1 carries out this function in the cell and to examine whether any other aspects of CPF function are regulated by PP1 dephosphorylation.

The purification of Pta1p also revealed interactions with several other proteins that have not previously been shown to be associated with the CPF complex (Table 4). Of these, the significance of the potential association with four ribosomal proteins and with EF1α is unclear for the reasons discussed above. Intriguingly, though, EF1α has previously also been found to associate with CF 1A, a factor required along with CPF and CF 1B for the mRNA 3′-cleavage reaction that precedes polyadenylation (*55*). In comparison with EF1α, Ref2p and Swd2p are both much better candidates for novel CPF components. Ref2p is already known to influence the efficiency of cleavage at weak poly(A) addition sites (*33*), and previous two-hybrid studies provide evidence that Ref2p can interact with Pap1p, the CPF-associated poly(A) polymerase, via their mutual interaction with Fir1p (*33*, *56*). Swd2p contains multiple WD40 domains, a property shared with the CPF subunit Pfs2p. Consistent with the existence of novel CPF complex polypeptides, we noticed a number of unidentified proteins that copurify with the CPF complex in previous studies (*46*, *53*, *57*−*59*), some of which appear to migrate at similar positions on SDS−PAGE gels to Glc7p, Ref2p, and Swd2p.

In conclusion, we have identified several new Glc7p-associated proteins in this work and verified the association of Glc7p with several more proteins with which it had been proposed to interact on the basis of genome-wide two-hybrid screens. Our data suggest the involvement of PP1 dephosphorylation in the regulation of a number of novel processes including mRNA 3′-end formation, and we have shown that Glc7p is a novel component of the CPF complex required for this process. Given the probable bias against low-abundance proteins, the use of capillary HPLC-linked tandem mass spectrometry might be more useful than MALDI for protein identification in future studies of this kind. Low-abundance proteins might also be more easily identified by direct analysis of total protein complexes using HPLC-linked tandem mass spectrometry (i.e., without using gels). Finally, the approach reported in this paper could be efficiently applied to global analysis of protein interactions of the yeast proteome if so wished.

ACKNOWLEDGMENT

The mass spectrometry reported in this work was carried out in the Fingerprints Proteomics Facility at Dundee, supported by a Wellcome Trust JIF award (060269).

REFERENCES

1. Cannon, J. F., Pringle, J. R., Fiechter, A., and Khalil, M. (1994) *Genetics 136*, 485−503.
2. Feng, Z. H., Wilson, S. E., Peng, Z. Y., Schlender, K. K., Reimann, E. M., and Trumbly, R. J. (1991) *J. Biol. Chem. 266*, 23796−23801.
3. Stark, M. J. (1996) *Yeast 12*, 1647−1675.
4. Bollen, M. (2001) *Trends Biochem. Sci. 26*, 426−431.

5. Hisamoto, N., Frederick, D. L., Sugimoto, K., Tatchell, K., and Matsumoto, K. (1995) *Mol. Cell. Biol. 15*, 3767−3776.
6. MacKelvie, S. H., Andrews, P. D., and Stark, M. J. (1995) *Mol. Cell. Biol. 15*, 3777−3785.
7. Tu, J., and Carlson, M. (1995) *EMBO J. 14*, 5939−5946.
8. Dombek, K. M., Voronkova, V., Raney, A., and Young, E. T. (1999) *Mol. Cell. Biol. 19*, 6029−6040.
9. Frederick, D. L., and Tatchell, K. (1996) *Mol. Cell. Biol. 16*, 2922−2931.
10. François, J. M., Thompson-Jaeger, S., Skroch, J., Zellenka, U., Spevak, W., and Tatchell, K. (1992) *EMBO J. 11*, 87−96.
11. Tu, J., Song, W., and Carlson, M. (1996) *Mol. Cell. Biol. 16*, 4199−4206.
12. Bailis, J. M., and Roeder, G. S. (2000) *Cell 101*, 211−221.
13. Tung, H. Y., Wang, W., and Chan, C. S. (1995) *Mol. Cell. Biol. 15*, 6064−6074.
14. Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S., and Sakaki, Y. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 1143−1147.
15. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 4569−4574.
16. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) *Nature 403*, 623−627.
17. Rout, M. P., and Blobel, G. (1993) *J. Cell Biol. 123*, 771−783.
18. Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000) *J. Cell Biol. 148*, 635−651.
19. Wigge, P. A., Jensen, O. N., Holmes, S., Soues, S., Mann, M., and Kilmartin, J. V. (1998) *J. Cell Biol. 141*, 967−977.
20. Zhao, J., Hyman, L., and Moore, C. (1999) *Microbiol. Mol. Biol. Rev. 63*, 405−445.
21. Burke, D., Dawson, D., and Stearns, T. (2000) *Methods in yeast genetics. A Cold Spring Harbor Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, New York.
22. Gietz, D., St. Jean, A., Woods, R. A., and Schiestl, R. H. (1992) *Nucleic Acids Res. 20*, 1425.
23. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994), Greene Publishing Associates and Wiley-Interscience, New York.
24. Andrews, P. D., and Stark, M. J. R. (2000) *J. Cell Sci. 113*, 507−520.
25. Gietz, R. D., and Sugino, A. (1988) *Gene 74*, 527−534.
26. Rayner, J. C., and Munro, S. (1998) *J. Biol. Chem. 273*, 26836−26843.
27. Peggie, M. W., MacKelvie, S. H., Bloecher, A., Knatko, E. V., Tatchell, K., and Stark, M. J. R. (2001) *J. Cell Sci. 115*.
28. Lau, D., Kunzler, M., Braunwarth, A., Hellmuth, K., Podtelejnikov, A., Mann, M., and Hurt, E. (2000) *J. Biol. Chem. 275*, 467−471.
29. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem. 68*, 850−858.
30. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) *Nat. Biotechnol. 17*, 1030−1032.
31. Hirano, K., Ito, M., and Hartshorne, D. J. (1995) *J. Biol. Chem. 270*, 19786−19790.
32. Hong, G., Trumbly, R. J., Reimann, E. M., and Schlender, K. K. (2000) *Arch. Biochem. Biophys. 376*, 288−298.
33. Russnak, R., Nehrke, K. W., and Platt, T. (1995) *Mol. Cell. Biol. 15*, 1689−1697.
34. Tollervey, D., and Caceres, J. F. (2000) *Cell 103*, 703−709.
35. Memet, S., Gouy, M., Marck, C., Sentenac, A., and Buhler, J. M. (1988) *J. Biol. Chem. 263*, 2830−2839.
36. Ursic, D., Himmel, K. L., Gurley, K. A., Webb, F., and Culbertson, M. R. (1997) *Nucleic Acids Res. 25*, 4778−4785.
37. Rasmussen, T. P., and Culbertson, M. R. (1998) *Mol. Cell. Biol. 18*, 6885−6896.
38. Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001) *Nature 413*, 327−331.
39. Bloecher, A., and Tatchell, K. (2000) *J. Cell Biol. 149*, 125−140.
40. Bell, G. I., Valenzuela, P., and Rutter, W. J. (1976) *Nature 261*, 429−431.
41. Irminger-Finger, I., and Mathis, N. (1998) *Cell Struct. Funct. 23*, 209−219.
42. Irminger-Finger, I., Hurt, E., Roebuck, A., Collart, M. A., and Edelstein, S. J. (1996) *J. Cell Biol. 135*, 1323−1339.
43. DeMarini, D. J., Adams, A. E., Fares, H., De Virgilio, C., Valle, G., Chuang, J. S., and Pringle, J. R. (1997) *J. Cell Biol. 139*, 75−93.
44. Souciet, J. L., Nagy, M., Le Gouar, M., Lacroute, F., and Potier, S. (1989) *Gene 79*, 59−70.
45. Neer, E. J., Schmidt, C. J., Nambudripad, R., and Smith, T. F. (1994) *Nature 371*, 297−300.
46. Preker, P. J., Ohnacker, M., Minvielle-Sebastia, L., and Keller, W. (1997) *EMBO J. 16*, 4727−4737.
47. Russnak, R., Pereira, S., and Platt, T. (1996) *Gene Expression 6*, 241−258.
48. Stuart, J. S., Frederick, D. L., Varner, C. M., and Tatchell, K. (1994) *Mol. Cell. Biol. 14*, 896−905.
49. Thompson, E. A., and Roeder, G. S. (1989) *Mol. Gen. Genet. 218*, 293−301.
50. Zhao, S., and Lee, E. Y. (1997) *J. Biol. Chem. 272*, 28368−28372.
51. Egloff, M. P., Johnson, D. F., Moorhead, G., Cohen, P. T., Cohen, P., and Barford, D. (1997) *EMBO J. 16*, 1876−1887.
52. Wu, X., and Tatchell, K. (2001) *Biochemistry 40*, 7410−7420.
53. Vo, L. T. A., Minet, M., Schmitter, J. M., Lacroute, F., and Wyers, F. (2001) *Mol. Cell. Biol. 21*, 8346−8356.
54. Mizrahi, N., and Moore, C. (2000) *Mol. Cell. Biol. 20*, 2794−2802.
55. Kessler, M. M., Zhao, J., and Moore, C. L. (1996) *J. Biol. Chem. 271*, 27167−27175.
56. del Olmo, M., Mizrahi, N., Gross, S., and Moore, C. L. (1997) *Mol. Gen. Genet. 255*, 209−218.
57. Barabino, S. M., Ohnacker, M., and Keller, W. (2000) *EMBO J. 19*, 3778−3787.
58. Ohnacker, M., Barabino, S. M., Preker, P. J., and Keller, W. (2000) *EMBO J. 19*, 37−47.
59. Zhao, J., Kessler, M., Helmling, S., O'Connor, J. P., and Moore, C. (1999) *Mol. Cell. Biol. 19*, 7733−7740.
60. Black, S., Andrews, P. D., Sneddon, A. A., and Stark, M. J. (1995) *Yeast 11*, 747−759.

BI015815E