

# A Genome-wide Screen for Site-specific DNA-binding Proteins\*

Tony R. Hazbun‡ and Stanley Fields§

We used a biochemical genomics method of assaying *Saccharomyces cerevisiae* proteins, derived from a nearly complete set of glutathione S-transferase fusions, to develop an approach that is able to identify proteins that bind to a DNA element. Using the upstream activation sequence (UAS) of the promoter for the invertase gene, *SUC2*, we identified both specific and nonspecific binding activities, which could be classified based on whether they bound with equivalent affinity to a nonspecific DNA competitor. Three transcription factors, Mig1, Yer028c, and Rgt1, were found to be binding activities specific to the *SUC2* UAS. Mig1 and Yer028c had been reported previously to bind to elements within the *SUC2* UAS, validating the ability of the method to identify sequence-specific factors. The third activity, Rgt1, had not been reported previously to bind to *SUC2*. Additional gel shift assays narrowed the Rgt1 binding site to the *SUC2*-B element within the *SUC2* UAS, which is similar to previously identified Rgt1 binding sites present in other genes. *In vivo* levels of invertase activity in an *rgt1Δ* strain were reduced relative to an isogenic *RGT+* strain when these strains were grown under inducing (low glucose) conditions, suggesting that Rgt1 may have a role in the activated transcription of *SUC2*. This report demonstrates the feasibility of identifying DNA binding activities by rapidly assaying a large fraction of the predicted open reading frames of an organism for binding to a regulatory DNA motif. *Molecular & Cellular Proteomics* 1:538–543, 2002.

The determination of transcription factor regulatory networks is a crucial component of the effort to understand biological pathways. For example, the *Saccharomyces cerevisiae* cell cycle is controlled by a series of transcription factors that form an underlying circuit (1). The development of chromatin immunoprecipitation in conjunction with the use of intergenic DNA arrays has allowed the genome-wide localization of transcription factor binding sites (1–3). This method allows the pairing of a large set of promoter regions to a specific transcription factor. Conversely, computational analyses are identifying many putative regulatory sequences within the intergenic regions of the yeast genome (4–6); how-

ever, in many cases, the corresponding DNA-binding proteins are unknown. Thus, the development of new methods to pair transcription factors with their target genes is needed.

We sought to develop an approach to determine the set of proteins that bind to a specific DNA sequence by combining a classical biochemical technique, the gel shift assay, with purified protein pools representing the yeast proteome (7). The upstream activation sequence (UAS)<sup>1</sup> for the *SUC2* gene, which encodes invertase, a periplasmic enzyme that hydrolyzes external sucrose into glucose and fructose, was chosen as a test case for our genome-wide screen. This sequence has been subjected to numerous genetic experiments to identify proteins involved in its regulation (8–10). In addition, the UAS has been delineated precisely by mutation and deletion (11–14). Analysis of the UAS has implicated Mig1, Mig2 (15), and Yer028c (16), transcription factors that bind to a pair of GC-rich sequence motifs in the *SUC2* UAS termed the *SUC2*-A and *SUC2*-B elements (see Fig. 1). Mig1 and Mig2 are repressors of *SUC2* expression at high levels of glucose (repressed growth conditions). Furthermore, there is evidence for an unknown activator(s) that functions at very low levels of glucose (derepressed growth conditions) (12, 14, 16).

We report here the use of a genome-wide gel shift assay to identify proteins capable of binding to the *SUC2* UAS. Additional gel shift assays, under high stringency conditions, allowed us to sort the activities into specific and nonspecific DNA-binding factors. In addition to identifying two transcription factors known to bind to the *SUC2* UAS, we identified a novel binding activity, Rgt1. Further evidence for the role of Rgt1 as a transcriptional activator comes from the observation that an *rgt1* strain had reduced *SUC2* expression.

## EXPERIMENTAL PROCEDURES

**Strain Growth and Protein Purification**—An array of 6144 individual yeast strains, each containing a different yeast ORF fused to GST, was obtained from Eric Phizicky. Yeast strains were grown in yeast extract, peptone, 2% glucose medium (YEPD) (17) to saturation in a 96-well microtiter plate format before pooling and storage in glycerol at –70 °C. Pools of 96 strains were grown overnight in 2 ml of SD Ura<sup>–</sup> (17) liquid medium and washed with an equal volume of SD Leu<sup>–</sup> Ura<sup>–</sup>, and the resuspended cells were used to inoculate 400-ml cultures of SD Leu<sup>–</sup> Ura<sup>–</sup>. Cultures were grown for 18–22 h to an *A*<sub>600</sub> of 0.8–1 and induced with 0.5 mM copper sulfate for 2 h. Cells were harvested, and protein was extracted using Y-PER (Pierce) supple-

From the Departments of Genome Sciences and Medicine, Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195

Received, April 4, 2002, and in revised form, July 22, 2002  
Published, MCP Papers in Press, August 5, 2002, DOI 10.1074/mcp.T200002-MCP200

<sup>1</sup> The abbreviations used are: UAS, upstream activation sequence; GST, glutathione S-transferase; ORF, open reading frame; SD, synthetic dextrose medium; Ura, uracil.

mented with 5 mM dithiothreitol and protease inhibitors (Roche Complete), as per manufacturers' instructions.

Whole cell extracts were passed over a 2-ml glutathione-agarose (Sigma) column, followed by three washes with 10 ml of wash buffer (20 mM Tris-HCl (pH 8), 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 0.5% *N*-octyl-glucoside, and 10% glycerol). Proteins were eluted by addition of 2 ml of elution buffer (20 mM Tris-HCl (pH 8), 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 0.5% *N*-octyl-glucoside, 10% glycerol, and 25 mM glutathione). Eluted proteins samples were adjusted to 40% glycerol and stored at  $-20^{\circ}\text{C}$ . Typically, the total protein concentrations in the eluted samples were  $\sim 0.2$  mg/ml as determined by Bio-Rad protein assay reagent.

For the deconvolution of the pools, the growth conditions were as above except that the strains were grown in 50-ml cultures, and the purification was scaled accordingly. Finally, individual yeast strains corresponding to DNA binding activities were grown in 50-ml cultures and purified accordingly. To confirm the identities of the GST-ORFs encoded by yeast strains that yielded gel shift activities, plasmids from yeast cells within individual wells were extracted and transformed into *Escherichia coli*, followed by sequencing.

**DNA Probes and Labeling**—The 145-bp *SUC2* UAS DNA probe and 272-bp non-homologous sequence probe were generated by PCR from the plasmid template, pRB58 (18), using the following sets of primer pairs: *SUC2* UAS forward, 5' GGCGTGCCTTTGTGAAC 3'; *SUC2* UAS reverse, 5' TGGGGTCGATTAACGCTA 3'; non-homologous forward, 5' AGTGCCTTCCAAAGCTA 3'; non-homologous reverse, 5' ACTATTCTAAATGACGTACTT 3'. The non-homologous probe was equivalent to nucleotides  $-608$  to  $-880$  relative to the *SUC2* transcriptional start site.

The 24-bp *SUC2*-A and 21-bp *SUC2*-B element probes were synthesized as oligonucleotides and annealed to make double-stranded probes. The sequences of the primers were as follows: *SUC2*-A forward, 5' AATAAAAATGCGGGGAATAATGGA 3'; *SUC2*-A reverse, 5' TCCATTATCCCCGCATTTTTATT 3'; *SUC2*-B forward, 5' TTCGCCCCCGGATAATTTCT 3'; *SUC2*-B reverse, 5' AGGAAATTATCGGGGGCGAA 3'.

PCR-generated products were agarose gel-purified, and the DNA was extracted with a Qiagen kit before labeling. DNA was end-labeled in 10  $\mu\text{l}$  containing 50 nM DNA fragment (60 ng of 145-bp DNA fragment), 5 units of T4 polynucleotide kinase (Invitrogen), 1 $\times$  polynucleotide kinase exchange buffer, and 70  $\mu\text{Ci}$  of ATP (3000 Ci/mmol) and incubated at  $37^{\circ}\text{C}$  for 1 h. Unincorporated nucleotide was removed by a gel filtration column (Centrisep 20) equilibrated with TES buffer (Tris-HCl (pH 8.0), 1 mM EDTA, and 100 mM NaCl). The *SUC2*-A and *SUC2*-B probes were generated by labeling oligonucleotides as above followed by annealing a 2-fold excess of cold complementary oligonucleotide. Unincorporated nucleotide was removed by gel filtration.

**Gel Shift Assay**—Protein-DNA complexes were formed by mixing equal volumes of purified protein eluates with 0.5 nM labeled DNA in 7.5  $\mu\text{l}$  of assay buffer (20 mM HEPES (pH 7.5), 20 mM  $\text{MgCl}_2$ , 20 mM  $\text{CaCl}_2$ , 50  $\mu\text{M}$   $\text{ZnCl}_2$ , 0.02% Nonidet P-40, 50 mM NaCl, 100 mg/ml acetylated bovine serum albumin, and 3 ng/ $\mu\text{l}$  poly[d(I-C)]) followed by incubation on ice for 15 min. Variations from the above amount of poly[d(I-C)] are noted in the figure legends. The experiment in Fig. 1 used poly[d(A-T)] instead of poly[d(I-C)]. Protein-DNA complexes were fractionated on non-denaturing 5% (49:1) polyacrylamide gels containing 0.5 $\times$  Tris borate EDTA buffer (5% glycerol, 1 mM  $\text{MgCl}_2$ , 1 mM  $\text{CaCl}_2$ , and 50  $\mu\text{M}$   $\text{ZnCl}_2$ ) at 5 volts/cm at  $4^{\circ}\text{C}$ . The gel was dried onto filter paper and exposed overnight to a phosphorimaging screen followed by scanning of the screen with a Molecular Dynamics Storm phosphorimager. Other gel shift assays were performed as above except with a 10% (49:1) polyacrylamide gel.

**Invertase Activity**—The yeast strains used to measure invertase

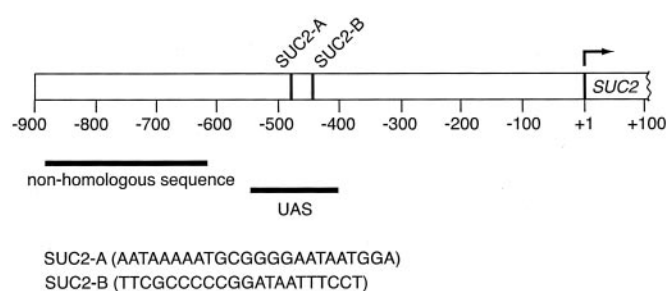


FIG. 1. An illustration of the *SUC2* promoter region indicating the relative size and position of the probes used in gel shift assays (solid lines). The sequences of the probes used to represent the GC-rich elements (*SUC2*-A and *SUC2*-B) are shown.

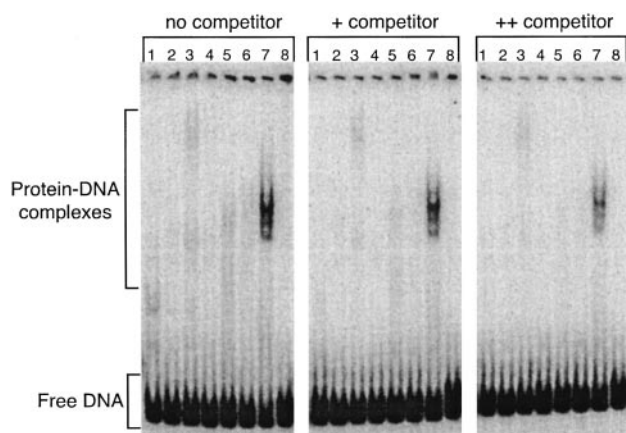
activity were derived from the *Saccharomyces* genome deletion project (19). The *rgt1 $\Delta$*  strain (ATCC number 4004887) and its isogenic parent strain (BY4741 *MATa his3 $\Delta$ 1 leu2 $\Delta$ 0 met15 $\Delta$ 0 ura3 $\Delta$ 0*) were obtained from the American Type Culture Collection. Periplasmic invertase activity was measured from whole cells in log phase based on a method described previously (20). At least three individual colonies from each strain were grown in YEPD overnight, and the cells were derepressed by inoculation into yeast extract, peptone, glycerol medium (YEPG) containing 0.1% glucose followed by growth for 3 h at  $30^{\circ}\text{C}$ . One unit of invertase-specific activity is 1  $\mu\text{mol}$  of glucose released/min/1 absorbance unit of cells at 600 nm.

## RESULTS

**Establishment of Conditions for the Gel Shift Assay**—To screen the yeast proteome, we generated pools of purified proteins with defined constituent proteins in each pool (7). An array of 6144 yeast strains, each overexpressing a single yeast ORF fused to GST, was split into 64 pools of 96 strains each. Protein pools were then generated by purification of the GST fusion proteins from whole cell extracts of each pool of strains. The protein pools were used in a radioactive gel shift assay, a sensitive method to detect the binding of proteins to a radiolabeled DNA fragment in which protein-DNA complexes migrate with reduced electrophoretic mobility in a non-denaturing gel compared with the unbound DNA. Various parameters of the gel shift assay were investigated, including the size of the probe DNA, assay buffer components, and gel conditions, resulting in the use of a 145-bp fragment corresponding to nucleotides  $-401$  to  $-546$  relative to the *SUC2* transcriptional start (Fig. 1). This fragment includes two GC-rich elements, *SUC2*-A and *SUC2*-B, which are bound by Mig1, Mig2, and Yer028c (15, 16).

Seven protein pools, including one known to contain transcriptional regulators of *SUC2*, were selected to assess the effect of unlabeled nonspecific competitor on protein binding to the *SUC2* UAS (Fig. 2). Our aim was to establish a level of nonspecific competitor that eliminated the majority of observed binding activities, thereby enriching for specific interactions. In the absence of unlabeled nonspecific competitor, binding activities could be observed in nearly every pool tested, indicating the many DNA binding activities found in yeast. Upon the addition of increasing amounts of poly[d-

(A-T)], most of these disappeared. At the highest level of poly[d(A-T)], only two samples had DNA binding activities (Fig. 2), which are likely *SUC2* UAS-specific binding activities. Each activity had a distinct electrophoretic signature of mobility, signal intensity, and banding pattern, dependent on factors such as the efficiency of protein expression and purification, possible proteolytic degradation during purification, DNA binding affinity, and the mass and charge of each protein. The electrophoretic signature of each activity was consistent regardless of whether the protein was present in a pool

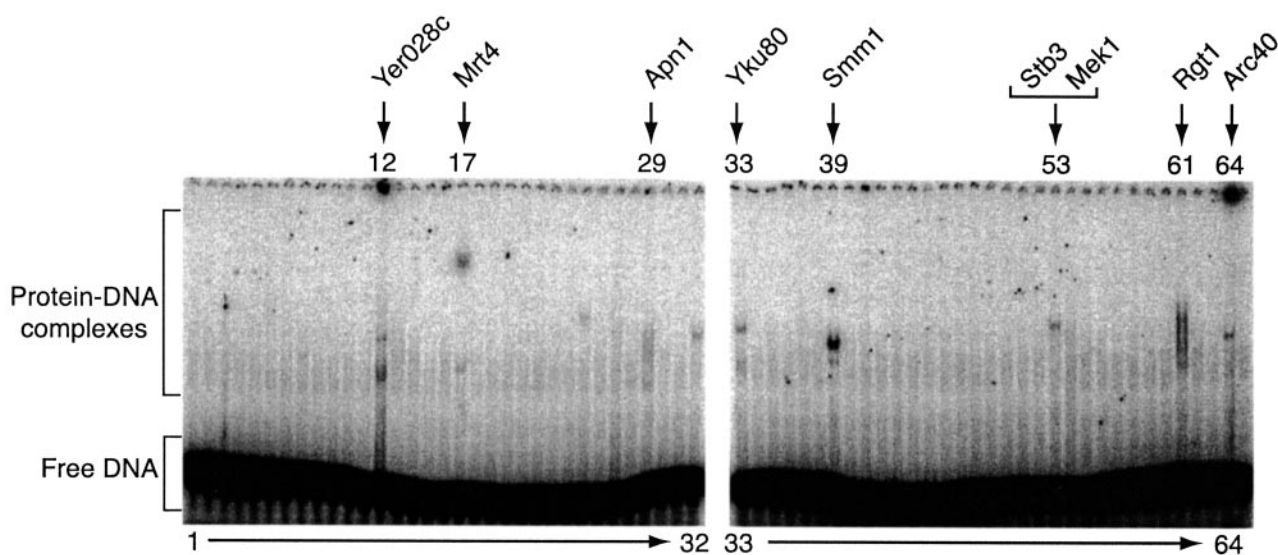


**FIG. 2. A pilot gel shift assay demonstrating the effect of unlabeled nonspecific competitor.** Protein-DNA complexes were formed by mixing eight pools of proteins with the *SUC2* UAS probe. Unlabeled nonspecific competitor, in the form of poly[d(A-T)], was added in the following amounts: *no competitor*, *+ competitor* (10 ng/ $\mu$ l), and *++ competitor* (50 ng/ $\mu$ l). Lanes 1–7 corresponded to pools 8, 9, 14, 22, 54, 57, and 61, respectively. Lane 8 had no protein added.

of 96 or was purified individually (see below).

**Identification of *SUC2* UAS Binding Activities**—Of the 64 pools representing the yeast proteome, eight pools with protein-DNA complexes of varying electrophoretic mobilities and radioactive intensities were investigated to determine the identities of the GST-ORFs responsible for the observed complex. Additional pools with protein-DNA complexes that are faintly visible in Fig. 3, but do not have their lane labeled, were not deconvoluted, because it became apparent that faint bands are more likely to be nonspecific complexes (see experiments below). The identities of the GST-ORFs in these eight pools were determined by a deconvolution strategy in which yeast strains from each 96-well microtiter plate were combined systematically as pools of those in each of the eight rows and each of the 12 columns. For example, purification of the GST fusion proteins from the row and column pools of plate 61 showed that protein-DNA complexes of an identical pattern and electrophoretic mobility as that present in the plate 61 pool were also present in the pools of *column 7* and of *row D* (Fig. 4). Based on these coordinates, the relevant protein in pool 61 was identified as Rgt1. Similarly, other protein-DNA complexes were deconvoluted and yielded the identities observed in Fig. 3. Deconvolution of pool 53 identified two closely migrating protein-DNA complexes yielding a total of nine complexes from the eight pools.

At least three transcription factors have been reported previously to bind the *SUC2* UAS, Yer028c, Mig1, and Mig2. Pool 12 contained the Yer028c complex. Pool 14 contained the Mig1 complex as observed in previous assays (Fig. 2, *lane 3*), but this complex was not clearly evident in Fig. 3 because of variability in the assay. However, the Mig1-DNA complex was



**FIG. 3. A genome-wide gel shift assay using 64 pools of purified proteins representing the yeast proteome.** The pools with protein-DNA complexes whose identities were deconvoluted are labeled with the pool number and the identity of the GST-ORF responsible for the electrophoretic mobility shift (*above the lane*). GST-ORF identities were determined by a deconvolution procedure as in Fig. 4. Two GST-ORFs were responsible for the gel shift signal observed in pool 53. Mig1 binding activity was observed previously in pool 14 but is not readily apparent in this image.

clearly visible upon deconvolution of pool 14 (data not shown). Mig2 binding activity was not detected in our assays, although a correct size insert with proper junction sequence was present in the appropriate transformant of pool 9.

The functional annotations for the other activities indicated that some nonspecific DNA binding activities were detected even in the presence of unlabeled competitor DNA. For example, the protein-DNA complex from pool 33 corresponded to Yku80, a protein that binds to double-stranded ends of DNA with little or no sequence specificity (21). The protein-

DNA complex observed in pool 29 was due to Apn1, an apurinic/aprymidinic endonuclease (22) that binds DNA under our experimental conditions. The functional annotations for Arc40, Mek1, Mrt4, Smm1, and Stb3 make them also unlikely to be sequence-specific activities for the *SUC2* UAS. The assumption that these activities are not sequence-specific for the *SUC2* UAS is corroborated by experimental results below and in Fig. 5.

Additional gel shift assays were performed to determine which binding activities identified in the initial genome-wide gel shift screen were specific for the *SUC2* UAS. To this end, purified proteins were compared in their binding to the *SUC2* UAS probe and a 272-bp non-homologous sequence. The non-homologous sequence is derived from an upstream region of the *SUC2* promoter (Fig. 1) that lacks the GC elements present in the UAS. Additionally, high levels of nonspecific competitor were used to sort activities into specific and non-specific categories.

Fig. 5B shows that Rgt1, Mig1, and Yer028c exhibited specific binding to the *SUC2* UAS, with minimal binding to the non-homologous sequence even at low levels of nonspecific competitor and significant binding to the *SUC2* UAS even at high levels of nonspecific competitor. In contrast, Stb3, Mrt4, and Mek1 bound similarly to both the *SUC2* UAS and to the non-homologous sequence probe at low levels of nonspecific competitor, and the addition of a 5-fold higher level of nonspecific competitor abolished binding to both probes (Fig. 5A). These results indicate that these proteins bound DNA nonspecifically. Other activities identified in the genome-wide screen also exhibited similar nonspecific activity (data not shown).

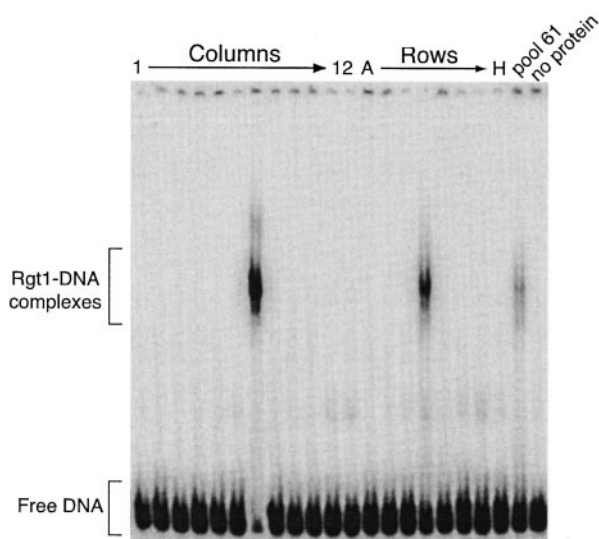


FIG. 4. **Deconvolution of pool 61 by gel shift assay.** The activity in pool 61 is identified as Rgt1 by determining the row and column with the same activity.

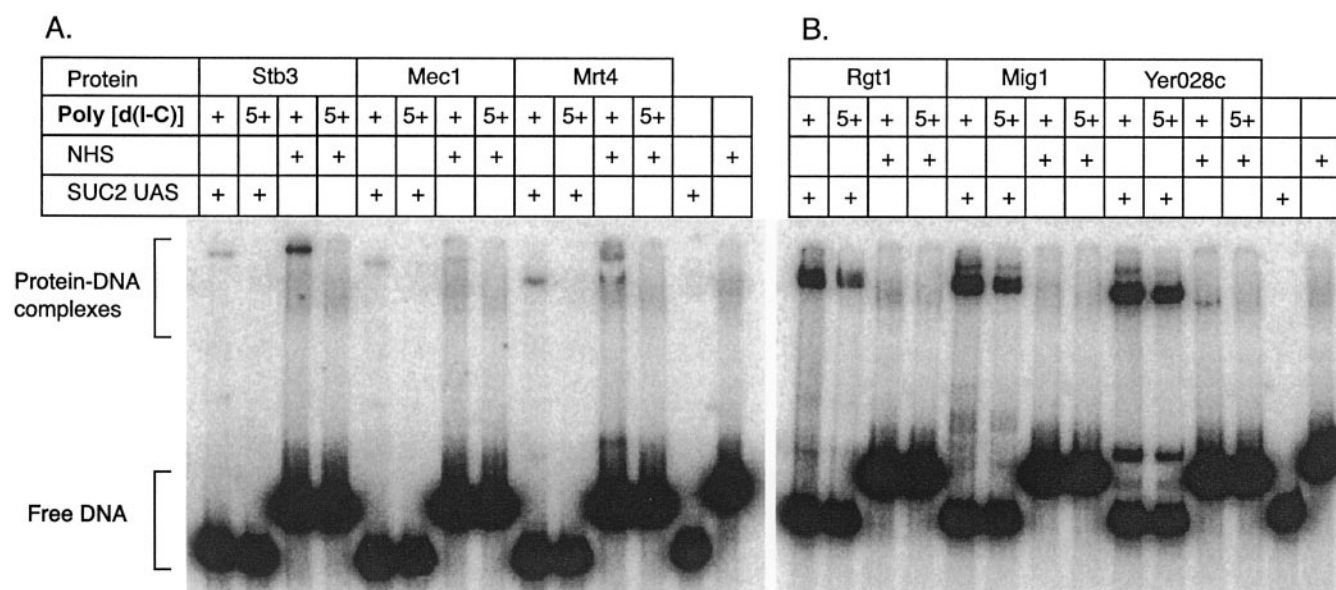
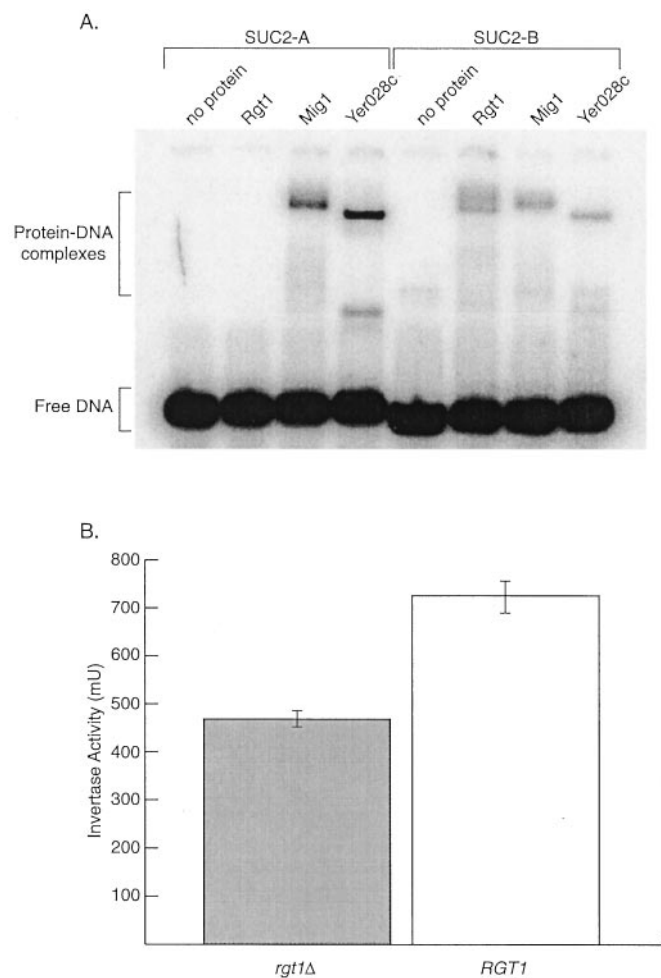


FIG. 5. **Determination of binding specificity by comparing DNA binding to unrelated sequences and at varying levels of poly[d(I-C)].** Panel A shows GST-ORFs that bind to both the *SUC2* UAS and the non-homologous sequence (NHS) probes at low levels of poly[d(I-C)]. DNA binding by these GST-ORFs was abolished by a 5-fold increase in the level of poly[d(I-C)]. Panel B shows sequence-specific binding to the *SUC2* UAS by GST-ORFs.



**FIG. 6. Evidence for the role of Rgt1 in the regulation of SUC2.** A, Rgt1 binds specifically to the SUC2-B element and not to the SUC2-A. Mig1 and Yer028c bind to both SUC2-B and SUC2-A elements. B, invertase activity in an *rgt1Δ* strain compared with the parent isogenic strain (*RGT1*).

**Regulation of the SUC2 Promoter by Rgt1**—Although previous studies identified Rgt1 binding sites in promoters of the *HXT* (hexose transporter) genes (23, 24), the binding of this factor to *SUC2* had not been observed. We performed gel shift assays with annealed oligonucleotide probes corresponding to the SUC2-B and SUC2-A elements and found that Rgt1 binds to only the SUC2-B probe (Fig. 6A). The sequence of the SUC2-B element (Fig. 1) includes a site with similar sequence and spacing from the transcriptional start site as the Rgt1 binding sites observed in *HXT* promoters. Mig1 and Yer028c bound specifically to both GC-rich elements (Fig. 6A), consistent with previous results indicating the presence of binding sites within these fragments.

We also investigated the effect of an *rgt1* deletion on the expression of invertase to address whether Rgt1 plays a role in the regulation of the *SUC2* gene. We compared invertase expression under inducing conditions in an *rgt1Δ* yeast strain to the isogenic parent strain. Inducing (low glucose) condi-

tions inactivate the Mig1 and Mig2 repressors and stimulate unknown activators for *SUC2* expression. Under these conditions, invertase expression in the *rgt1Δ* strain was reduced to 62% of the parent strain, indicating that Rgt1 is at least partially responsible for fully induced levels of *SUC2* (Fig. 6B).

#### DISCUSSION

We used the gel shift assay on a genome-wide basis to identify yeast proteins that bind in a sequence-specific manner to a DNA element, in this case a regulatory region of the *SUC2* gene. Two homologous transcription factors, Mig1 and Yer028c, were found that had been demonstrated previously to bind a pair of GC-rich elements in the *SUC2* UAS. Whereas Mig1 is required for glucose repression of *SUC2* expression, Yer028c does not appear to be involved in the glucose-mediated regulation of *SUC2* despite its ability to bind sites similar to those bound by Mig1 (16). The target genes of Yer028c are unknown. Mig2 is a transcription factor homologous to Mig1 that is known to bind the *SUC2* UAS but was not identified in our assay even though the gene fusion was present in the array. The inability to detect Mig2 could be because of poor growth of the strain carrying the GST-Mig2 fusion or reduced expression or inefficient purification of the fusion.

We also identified the Rgt1 transcription factor as a novel *SUC2* UAS binding activity. Rgt1 is a key regulator of the *HXT* genes (24) and is classically viewed as having three transcriptional modes, depending on the level of glucose in the growth media, that are as follows: (i) activation in high glucose, (ii) neutral activity in low glucose, and (iii) repression in the absence of glucose. Invertase expression is induced in low glucose media, and interestingly we observed a reduced level of induced invertase expression in an *rgt1Δ* strain. Although similar results had been observed for invertase expression in an *rgt1Δ* strain (11), the effect had not been considered significant. However, the invertase expression results, our identification of an Rgt1 binding site in the *SUC2* promoter, and the similarity of this site to sites in the *HXT* genes support the idea that Rgt1 maybe an activator of *SUC2* transcription at low glucose levels. In addition, a LexA-Rgt1 hybrid can activate a reporter gene with LexA binding sites under low glucose conditions, suggesting that Rgt1 can act as an activator at this level of glucose (23). More rigorous validation of the role of Rgt1 in the regulation of *SUC2* requires the measurement of mRNA levels or *in vivo* binding by chromatin immunoprecipitations.

The assay conditions used in the genome-wide gel shift assay (Fig. 3) were not stringent enough to completely prevent nonspecific proteins from binding to the *SUC2* UAS probe. These activities were classified easily as nonspecific for the *SUC2* UAS by comparing binding to a non-homologous probe. The advantage of the less stringent conditions is that DNA binding activities may be detected from proteins that were unknown to bind DNA. An example of this is the Stb3 protein, which interacts with the transcriptional repressor Sin3

(25), but whose molecular function is unknown.

Previous use of the biochemical genomics strategy with pooled yeast GST fusions has identified enzyme activities (7), whereas the activities identified in this report are observed under equilibrium binding conditions. The ability to detect such binding events is in large part because of the sensitivity of the gel shift assay. Another advantageous aspect of the gel shift assay is that even two binding activities in a pool can be deconvoluted readily, because each protein has a unique charge and size, and thus the complexes migrate differently through a native gel. Nevertheless, improvements of the approach are possible. Specifically, concentrating purified protein samples and removing proteins with nonspecific binding activities should improve the robustness of the procedure. More stringent initial assays that used high levels of nonspecific competitor DNA would streamline the procedure.

Computational analyses of yeast intergenic regions have identified many regulatory motifs in promoters from diverse sets of genes. In many cases, the binding factors for these motifs are unknown. Similar experimental and computational approaches are being applied to gene sets from other organisms. The combination of the gel shift assay and pools of GST-ORFs provides a convenient method to characterize these binding activities.

*Acknowledgments*—We thank Eric Phizicky for the GST-ORF yeast array. Eric Phizicky, Michael Devit, and Marris Vignali provided helpful comments on the manuscript.

\* This work was supported by National Center for Research Resources of the National Institutes of Health Grant P41 RR11823. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 206-616-4523; Fax: 206-543-0754; E-mail: thazbun@u.washington.edu.

§ Investigator of the Howard Hughes Medical Institute.

#### REFERENCES

- Simon, I., Barnett, J., Hannett, N., Harbison, C. T., Rinaldi, N. J., Volkert, T. L., Wyrick, J. J., Zeitlinger, J., Gifford, D. K., Jaakkola, T. S., and Young, R. A. (2001) Serial regulation of transcriptional regulators in the yeast cell cycle. *Cell* **106**, 697–708
- Iyer, V. R., Horak, C. E., Scafe, C. S., Botstein, D., Snyder, M., and Brown, P. O. (2001) Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**, 533–538
- Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T. L., Wilson, C. J., Bell, S. P., and Young, R. A. (2000) Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309
- Cliften, P. F., Hillier, L. W., Fulton, L., Graves, T., Miner, T., Gish, W. R., Waterston, R. H., and Johnston, M. (2001) Surveying *Saccharomyces* genomes to identify functional elements by comparative DNA sequence analysis. *Genome Res.* **11**, 1175–1186
- Sinha, S., and Tompa, M. (2000) A statistical method for finding transcription factor binding sites. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **8**, 344–354
- Zhang, M. Q. (1999) Promoter analysis of co-regulated genes in the yeast genome. *Comput. Chem.* **23**, 233–250
- Martzen, M. R., McCraith, S. M., Spinelli, S. L., Torres, F. M., Fields, S., Grayhack, E. J., and Phizicky, E. M. (1999) A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153–1155
- Carlson, M., Osmond, B. C., and Botstein, D. (1981) Mutants of yeast defective in sucrose utilization. *Genetics* **98**, 25–40
- Vallier, L. G., and Carlson, M. (1991) New SNF genes, GAL11 and GRR1 affect SUC2 expression in *Saccharomyces cerevisiae*. *Genetics* **129**, 675–684
- Carlson, M., Osmond, B. C., Neigeborn, L., and Botstein, D. (1984) A suppressor of SNF1 mutations causes constitutive high-level invertase synthesis in yeast. *Genetics* **107**, 19–32
- Ozcan, S., Vallier, L. G., Flick, J. S., Carlson, M., and Johnston, M. (1997) Expression of the SUC2 gene of *Saccharomyces cerevisiae* is induced by low levels of glucose. *Yeast* **13**, 127–137
- Bu, Y., and Schmidt, M. C. (1998) Identification of cis-acting elements in the SUC2 promoter of *Saccharomyces cerevisiae* required for activation of transcription. *Nucleic Acids Res.* **26**, 1002–1009
- Sarokin, L., and Carlson, M. (1984) Upstream region required for regulated expression of the glucose-repressible SUC2 gene of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4**, 2750–2757
- Wu, J., and Trumbly, R. J. (1998) Multiple regulatory proteins mediate repression and activation by interaction with the yeast Mig1 binding site. *Yeast* **14**, 985–1000
- Lutfiyya, L. L., and Johnston, M. (1996) Two zinc-finger-containing repressors are responsible for glucose repression of SUC2 expression. *Mol. Cell. Biol.* **16**, 4790–4797
- Lutfiyya, L. L., Iyer, V. R., DeRisi, J., DeVit, M. J., Brown, P. O., and Johnston, M. (1998) Characterization of three related glucose repressors and genes they regulate in *Saccharomyces cerevisiae*. *Genetics* **150**, 1377–1391
- Sherman, F., Fink, G. R., and Hicks, J. B. (eds) (1986) *The Laboratory Course Manual for Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Carlson, M., and Botstein, D. (1982) Two differentially regulated mRNAs with different 5' ends encode secreted with intracellular forms of yeast invertase. *Cell* **28**, 145–154
- Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., El Bakkoury, M., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Davis, R. W., et al. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906
- Boehm, J., Ulrich, H. D., Ossig, R., and Schmitt, H. D. (1994) Kex2-dependent invertase secretion as a tool to study the targeting of transmembrane proteins which are involved in ER → Golgi transport in yeast. *EMBO J.* **13**, 3696–3710
- Feldmann, H., Driller, L., Meier, B., Mages, G., Kellermann, J., and Winnacker, E. L. (1996) HDF2, the second subunit of the Ku homologue from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **271**, 27765–27769
- Popoff, S. C., Spira, A. I., Johnson, A. W., and Demple, B. (1990) Yeast structural gene (APN1) for the major apurinic endonuclease: homology to *Escherichia coli* endonuclease IV. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4193–4197
- Ozcan, S., Leong, T., and Johnston, M. (1996) Rgt1p of *Saccharomyces cerevisiae*, a key regulator of glucose-induced genes, is both an activator and a repressor of transcription. *Mol. Cell. Biol.* **16**, 6419–6426
- Ozcan, S., and Johnston, M. (1999) Function and regulation of yeast hexose transporters. *Microbiol. Mol. Biol. Rev.* **63**, 554–569
- Kasten, M. M., and Stillman, D. J. (1997) Identification of the *Saccharomyces cerevisiae* genes STB1–STB5 encoding Sin3p binding proteins. *Mol. Gen. Genet.* **256**, 376–386