Molecular Structure of *Saccharomyces cerevisiae* Gal1p, a Bifunctional Galactokinase and Transcriptional Inducer*

James B. Thoden[‡], **Christopher A. Sellick**[§], **David J. Timson**[¶], **Richard J. Reece**[§1], and **Hazel M. Holden**[‡2]

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, §Faculty of Life Sciences, University of Manchester, The Michael Smith Building, Oxford Road, Manchester M13 9PT, United Kingdom, and ¶School of Biology and Biochemistry, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, United Kingdom

Gal1p of *Saccharomyces cerevisiae* is capable of performing two independent cellular functions. First, it is a key enzyme in the Leloir pathway for galactose metabolism where it catalyzes the conversion of α-D-galactose to galactose 1-phosphate. Second, it has the capacity to induce the transcription of the yeast *GAL* genes in response to the organism being challenged with galactose as the sole source of carbon. This latter function is normally performed by a highly related protein, Gal3p, but in its absence Gal1p can induce transcription, albeit inefficiently, both *in vivo* and *in vitro*. Here we report the x-ray structure of Gal1p in complex with α-D-galactose and Mg-adenosine 5′-(β,γ-imido)triphosphate (AMPPNP) determined to 2.4 Å resolution. Overall, the enzyme displays a marked bilobal appearance with the active site being wedged between distinct N- and C-terminal domains. Despite being considerably larger than other galactokinases, Gal1p shares a similar molecular architecture with these enzymes as well as with other members of the GHMP superfamily. The extraordinary levels of similarity between Gal1p and Gal3p (~70% amino acid identity and ~90% similarity) have allowed a model for Gal3p to be constructed. By identifying the locations of mutations of Gal3p that result in altered transcriptional properties, we suggest potential models for Gal3p function and mechanisms for its interaction with the transcriptional inhibitor Gal80p. The *GAL* genetic switch has long been regarded as a paradigm for the control of gene expression in eukaryotes. Understanding the manner in which two of the proteins that function in transcriptional regulation interact with one another is an important step in determining the overall molecular mechanism of this switch.

There are four enzymes in the Leloir pathway that are responsible for the conversion of β-D-galactose to the more metabolically useful glucose 1-phosphate (1). In the first step of this pathway, β-D-galactose is epimerized to α-D-galactose by galactose mutarotase. The following step involves the ATP-dependent addition of a phosphate group to the galactose C-1 hydroxyl by galactokinase to yield galactose 1-phosphate. Subsequently, a UMP moiety, derived from UDP-glucose, is transferred to galactose 1-phosphate to generate glucose 1-phosphate and UDP-galactose. This reaction is catalyzed by galactose-1-phosphate uridylyl-transferase. To complete the pathway, UDP-galactose is converted to UDP-glucose by UDP-galactose 4-epimerase.

The enzymes of the Leloir pathway have been identified in all biological kingdoms. In the yeast, *Saccharomyces cerevisiae*, the genes encoding these four enzymatic functions (collectively termed the *GAL* genes) are coordinately regulated at the level of transcription (2). The *GAL* genes are essentially inert in yeast cells grown in the absence of galactose, but the genes are rapidly expressed, and to a high level, when the cells are switched to a medium containing galactose as the sole carbon source (3). Transcriptional activation is brought about through the concerted action of three proteins: a transcriptional activator, Gal4p; a transcriptional repressor, Gal80p; and an inducer, Gal3p (4). In the presence of galactose and ATP, Gal3p associates with Gal80p to alleviate its repressing effects, thereby allowing Gal4p to recruit the RNA polymerase II transcriptional machinery to each of the *GAL* genes (5–7).

Gal3p is highly related to the galactokinases of the Leloir pathway, but the protein does not possess a galactokinase activity (8, 9). It can, however, be converted into a galactokinase through the addition of two amino acids into its sequence (a serine and an alanine after Ser 164), (9) underscoring the similarity between it and the Leloir pathway enzyme. The yeast galactokinase Gal1p shows 70% amino acid identity (>90% similarity) with Gal3p over its entire length.

In the absence of Gal3p, the *GAL* genes are still induced but at a very slow rate (10, 11). The absence of both Gal3p and Gal1p, however, renders yeast cells unable to grow using galactose as their sole carbon source (12). In addition, the effect of the loss of Gal3p can be largely overcome if Gal1p is produced at higher than wild-type levels (8). Other galactokinases, *e.g.* from *Escherichia coli*, can substitute functionally for the galactokinase activity of Gal1p but are unable to induce *GAL* gene expression (12). Combined, these data suggest that although Gal1p is primarily a galactokinase, it can also function as a weak transcriptional inducer. This conclusion has also been supported through a number of *in vitro* experiments (9, 13).

Here, we report the crystal structure of *S. cerevisiae* Gal1p complexed with α-D-galactose and the nonhydrolyzable ATP analog AMPPNP (14). This analysis provides a new molecular template upon which to base a more complete understanding of the molecular architecture of the ligand sensor, Gal3p, and the location of mutations in Gal3p that give rise to a constitutive phenotype.

EXPERIMENTAL PROCEDURES

Cloning of the Gene Encoding Gal1p—The gene encoding Gal1p was PCR-amplified from a yeast expression plasmid pAP60 (9) such that the forward and reverse primers added NheI and XhoI cloning sites, respectively. PCR-amplification was accomplished with Platinum *Pfx* DNA polymerase (Invitrogen) according to the manufacturer's instructions and standard cycling conditions. The PCR product was purified with the QIAquick PCR purification kit (Qiagen) followed by A-tailing and sub-

sequent ligation into the pGEM-T vector (Promega) for sequencing purposes. *E. coli* DH5α cells were transformed with the resulting vector. The gene was sequenced with the ABI Prism™ Big Dye primer cycle sequencing kit (Applied Biosciences, Inc). The pGEM-T vector construct was then digested with NheI and XhoI, and the gene was separated from digestion by-products on a 1.0% agarose gel. The gene was excised from the gel, purified with the QIAquick gel extraction kit (Qiagen), and ligated into the expression vector derived from pET-28b(+) (Novagen) that was previously digested with NheI and XhoI. This modified pET-28b encodes a TEV protease recognition site thereby resulting in a construct that contains an additional 20 amino acid residues at the N terminus with the following sequence: MGSSHHHHHHSSENLYFQGH.

E. coli DH5α cells were transformed with the ligation mixture and plated onto LB/agar plates for selection with kanamycin. Individual colonies were selected, cultured overnight and plasmid DNA extracted with a QIAprep Spin Miniprep kit (Qiagen). Plasmids were tested for insertion of the gene by digestion with NheI and XhoI.

Protein Expression and Purification—For protein expression, the pET28-Gal1p plasmid was used to transform *E. coli* HMS174(DE3) cells (Novagen). A starter culture from a single colony was grown overnight at 37 °C in LB medium supplemented with kanamycin. Subsequently, 10 ml of the mixture was transferred to 1000 ml of supplemented TB medium (50 mg/liter kanamycin and 1 g/liter galactose) in a 2-liter shaker flask and grown at 37 °C until an optical density of ~0.8 was achieved at 600 nm. The culture was then transferred to a shaker at room temperature (~20 °C) and allowed to grow until an optical density of greater than 1.8 was obtained, at which point isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 0.05 mM. Cell growth was allowed to continue at room temperature for an additional 18 h.

The cells were harvested by centrifugation at 4000 × *g* for 15 min and frozen in liquid nitrogen. Frozen cells (250 g) were thawed in 750 ml of lysis buffer consisting of 50 mM NaH$_2$PO$_4$, 100 mM galactose, 10 mM imidazole, and 300 mM NaCl (pH 8.0). The thawed cells were placed in an ice bath and disrupted by seven rounds of sonication (1-min duration each) separated by 5 min of cooling. Cellular debris was removed by centrifugation at 20,000 × *g* for 25 min. The clarified supernatant was loaded onto a 25 ml column of nickel-nitrilotriacetic acid-agarose (Qiagen) that had been equilibrated previously with lysis buffer. The column was then washed with lysis buffer until the absorbance reading at 280 nm reached background level. The protein was eluted with a linear gradient of 10–250 mM imidazole in lysis buffer. Protein-containing fractions were pooled based on purity as judged by SDS-PAGE and dialyzed against 10 mM Tris-HCl (pH 8.0), 100 mM galactose, and 100 mM NaCl. The dialyzed protein was further purified by anion exchange HPLC chromatography using a 6-ml Resource-Q column. The protein was eluted at pH 8.0 (25 mM Tris-HCl plus 100 mM galactose) with a linear gradient from 50 to 400 mM NaCl. Protein-containing fractions were again pooled based on purity as judged by SDS-PAGE and dialyzed against 10 mM Tris-HCl (pH 8.0), 100 mM galactose, and 200 mM NaCl and concentrated to 15.0 mg/ml based on an extinction coefficient of 1.31 cm/(mg·ml) as calculated with the program Protean (DNASTAR, Inc., Madison, WI). The N-terminal tag was removed by treatment with TEV protease. Uncleaved enzyme and TEV protease were removed by passage over nickel-nitrilotriacetic acid-agarose. A typical yield was ~40 mg of protein/250 grams of cells.

Crystallization of Gal1p—A search for crystallization conditions was conducted at both room temperature and at 4 °C via the hanging drop method of vapor diffusion utilizing an "in-house" designed sparse matrix screen composed of 144 conditions. The best crystals were

TABLE ONE

X-ray data collection statistics

Space group	$P2_1$
Unit cell dimensions (Å, deg)	$a = 63.3$, $b = 85.0$ $c = 112.3$, $\beta = 96.7$
Resolution limits (Å)	30.0–2.4 (2.5–2.4)[a]
No. independent reflections	44350 (4591)
Completeness (%)	95.5 (86.0)
Redundancy	3.4 (1.7)
Avg I/avg σ(I)	11.2 (2.9)
R_{sym} (%)[b]	7.5 (27.3)

[a] Statistics for the highest resolution bin.
[b] $R_{sym} = (\Sigma|I - \bar{I}|/\Sigma I) \times 100$.

observed growing at room temperature from hanging drop experiments with precipitant solutions of 18–25% poly(ethylene glycol) 8000 containing 200 mM LiCl buffered with 100 mM MES[3] (pH 6.0) and enzyme concentrations of 15.0 mg/ml plus 5 mM MgAMPPNP. Single crystals grew to maximum dimensions of ~0.8 × 0.5 × 0.1 mm in 2–4 weeks. They belonged to the space group $P2_1$ with unit cell dimensions of $a = 63.3$ Å, $b = 85.0$ Å, $c = 112.3$ Å, and $\beta = 96.7°$. The asymmetric unit contained two polypeptide chains.

Structural Analysis of Gal1p—An x-ray data set was collected to 2.4 Å resolution at 293 K with a Bruker AXS Platinum 135 CCD detector controlled with the PROTEUM software suite (Bruker AXS Inc., Madison, WI). The x-ray source was CuKα radiation from a Rigaku RU200 x-ray generator equipped with Montel optics, operated at 50 kV and 90 mA. The x-ray data were processed with SAINT version 7.06A (Bruker AXS Inc.) and internally scaled with SADABS version 2005/1 (Bruker AXS Inc.). X-ray data collection statistics are presented in TABLE ONE.

The structure was solved by molecular replacement with the program EPMR utilizing as a search model the structure of human *N*-acetylgalactosamine kinase (Protein Data Bank accession no. 2A2C). Three iterations of least-squares refinement with the software package REFMAC (15) and cyclical averaging with solvent flattening using DM (16) followed by manual model building with TURBO (17) allowed for ~80% of the Gal1p model to be built. Alternate cycles of manual model building and least-squares refinement with TNT (18) allowed for the remaining ordered parts of the structure to be traced. The final *R*-factor is 17.4% for all measured x-ray data to 2.4 Å resolution. Refinement statistics are presented in TABLE TWO. The Ramachandran plot reveals no significant outliers. Specifically, 90.2% of the residues fall into the core regions, and 9.8% are located in the allowed regions.

The model for Gal3p was constructed by using the Gal1p coordinates as a template and substituting the appropriate amino acids on the basis of amino acid sequence alignments with the program TURBO (17). No attempts at energy minimization were done.

RESULTS AND DISCUSSION

Overall Structure of Gal1p—Gal1p is a monomer (19). The enzyme, in complex with α-D-galactose and MgAMPPNP, crystallized in the space group $P2_1$ with two molecules in the asymmetric unit. Overall the electron density map was well ordered with the exception of the first eight residues in both molecules. Additionally, residues 19, 298–305, and 414–415 in the first molecule and residues 18–19, 298–305, and 414–418 in the second molecule of the asymmetric unit were disordered. In that the two molecules superimpose with a root-mean-square deviation of 0.47 Å, the following discussion refers only to molecule 1 in the coordinate file.

---

[3] The abbreviations used are: MES, 4-morpholineethanesulfonic acid; AMPPNP, adenosine 5′-(β,γ-imido)triphosphate.

TABLE TWO

Least-squares refinement statistics

Resolution limits (Å)	30.0–2.40
R-factor[a] (overall) (%)/no. reflections	17.4/44350
R-factor (working) (%)/no. reflections	17.2/39861
R-factor (free) (%)/no. reflections	20.5/4489
No. protein atoms	7848
No. hetero-atoms[b]	281
Average B-values (Å^2)	
Protein atoms	39.9
MgAMPPNP	28.9
α-D-Galactose	17.7
Solvents	45.6
Weighted root-mean-square deviations from ideality	
Bond lengths (Å)	0.012
Bond angles (degree)	2.33
Trigonal planes (Å)	0.008
General planes (Å)	0.012
Torsional angles (deg)[c]	17.9

[a] R-factor = $(\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$, where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.
[b] Hetero-atoms include 2 AMPPNP molecules, 2 α-D-galactose molecules, 2 magnesium ions, 2 chloride ions, and 191 water molecules.
[c] Torsional angles were not restrained during the refinement.

Gal1p has overall dimensions of ~72 × 49 × 66 Å and adopts a distinctly bilobal structure with the active site located roughly in the middle of the molecule (Fig. 1a). The N-terminal domain is dominated by a six-stranded mixed β-sheet flanked on one side by an α-helix and on the other side by four α-helices. There are two four-stranded antiparallel β-sheets in the C-terminal domain in addition to 10 α-helices. Pro-236 adopts the *cis* conformation and is located in a surface loop ~18 Å from the active site. The α-helix formed by Ser-170 to Met-187 is situated such that the positive end of its helix dipole moment projects toward the phosphate groups of the nucleotide.

The Active Site—The electron density corresponding to the galactose and MgAMPPNP ligands is well ordered and easily interpretable as shown in Fig. 1b. The magnesium ion is octahedrally coordinated by the β- and γ-phosphoryl oxygens of the AMPPNP, three water molecules, and the carbonyl oxygen of Ser-264. Ligand–metal bond distances range from 2.7 to 3.2 Å, indicating that the metal ion is quite loosely bound. The magnesium ion has a B-value of 34.3 Å^2, whereas the nucleotide and galactose ligands refined with average B-values of 20.4 and 24.7 Å^2, respectively.

A close-up view of the Gal1p active site, formed by residues from both the N- and C-terminal domains, is presented in Fig. 1c. Phe-100, Trp-123, and Phe-174 form a decidedly aromatic crown that encircles the adenine base of the nucleotide. This is similar to that observed for *N*-acetylgalactosamine kinase (GalNAc kinase) (20). The side chains of Asn-95 and Ser-170 serve to anchor the adenine group to the protein via hydrogen bonds to N-1 and N-7, respectively. Tyr-126 abuts the nucleotide ribose, and the loop delineated by Thr-164 to Ser-171 provides three backbone amide hydrogen bonds to the γ-phosphoryl group of AMPPNP. Key residues involved in binding the sugar to the protein include Arg-53, Glu-59, His-60, Asp-62, Asn-213, Asp-217, and Tyr-274. Both Arg-53 and Asp-217 are strictly conserved among galactokinases, and it is believed that these residues may play critical roles in catalysis (21). Seven ordered water molecules lie within 3.7 Å of the galactose and MgAMPPNP ligands.

Similarity of Gal1p to other Galactokinases—The three-dimensional structures of two bacterial galactokinases and the human form of the enzyme have been determined previously (21–23). These enzymes show amino acid sequence similarities of ~45% to Gal1p. Additionally, the molecular architecture of human GalNAc kinase has been defined recently, and this enzyme demonstrates 50% sequence similarity to Gal1p. For the sake of simplicity, only the structures of yeast Gal1p and the human forms of galactokinase and *N*-acetylgalactosamine kinase will be compared.

Human galactokinase and Gal1p superimpose with a root-mean-square deviation of 1.3 Å for 355 structurally equivalent α-carbons (Fig. 2a). The N terminus of each enzyme wraps around the main body of the molecule quite differently, such that they are separated by ~38 Å as can be seen in Fig. 2a. The two proteins begin to align at Lys-28 in Gal1p and Leu-12 in galactokinase. There are several small insertions or deletions between these two proteins in the N-terminal domain. The first major difference occurs at Ala-289 in Gal1p and Gly-251 in galactokinase where there is a 45-residue insertion in Gal1p. The region defined by Ala-289 to Met-342 in Gal1p contains a loop and an additional α-helix (Leu-310 to His-322). In galactokinase, the polypeptide chain connecting Gly-251 to Gln-259 adopts a random coil conformation. This region is labeled "A" in Fig. 2a. The second area where the two polypeptide chains differ considerably occurs at Leu-350 in Gal1p and Arg-267 in galactokinase as indicated by the label "B" in Fig. 2a. Here there is a 33-residue insertion in Gal1p, which folds into two additional α-helices (Val-359 to Leu-366 and Arg-370 to Tyr-377) and an extended loop (Leu-378 to Lys-389). Following these insertions, the polypeptide chains for Gal1p and galactokinase are quite similar with, again, several small insertions and deletions. The C termini of these enzymes are separated by ~3 Å.

Gal1p is more similar in size to GalNAc kinase (528 *versus* 458 amino acids, respectively). Accordingly, these two enzymes correspond with a root-mean-square deviation of 1.2 Å for 411 structurally equivalent α-carbons (Fig. 2b). The first major difference between these two molecules occurs at Lys-139 in Gal1p and Gly-123 in GalNAc kinase. In Gal1p there is an additional α-helical turn (Pro-143 to Arg-145). The second major difference between these two enzymes occurs at Tyr-317 in Gal1p and Leu-275 in GalNAc kinase. Here there is an 18-residue insertion in Gal1p as indicated by the label "A" in Fig. 2b. Unlike the difference noted between Gal1p and human galactokinase, Gal1p and GalNAc kinase correspond well in the region labeled "B" (Fig. 2b).

One of the surprises in the Gal1p structure is the manner in which the AMPPNP moiety is accommodated in the active site. In both human galactokinase and GalNAc kinase, the γ-phosphoryl group of the nucleotide is situated within 3.1 to 3.5 Å of the sugar C-1 hydroxyl group that is ultimately phosphorylated in the reaction catalyzed by these enzymes. The situation is different in Gal1p where the γ-phosphoryl group of the nucleotide lies at ~6 Å from the sugar hydroxyl. This distance is clearly incompatible for catalysis and suggests that the nucleotide analog is bound to Gal1p in a nonproductive manner. A comparison of the nucleotide conformations observed in Gal1p and GalNAc kinase is depicted in Fig. 2c. In GalNAc kinase, Lys-234 forms a hydrogen bonding interaction with the bridging nitrogen of AMPPNP and the carbonyl oxygen of Met-232. The side chain of the structurally equivalent Lys-266 in Gal1p, however, swings toward the sugar and the binding site for the magnesium ion. As a consequence, the AMPPNP ligand is shifted in the active site and the phosphoryl groups adopt quite different dihedral angles relative to those observed in GalNAc kinase. Clearly the ligand binding pocket of Gal1p is large enough to accommodate such changes in nucleotide conformation without major structural perturbations of the surrounding active site residues.

Homology Model for Gal3p—The similarity between yeast galactokinase (Gal1p) and the transcriptional inducer of the *GAL* genes (Gal3p) is extensive and occurs over the lengths of each protein. This high level of sequence similarity (~90%) allows for a homology model of Gal3p to be

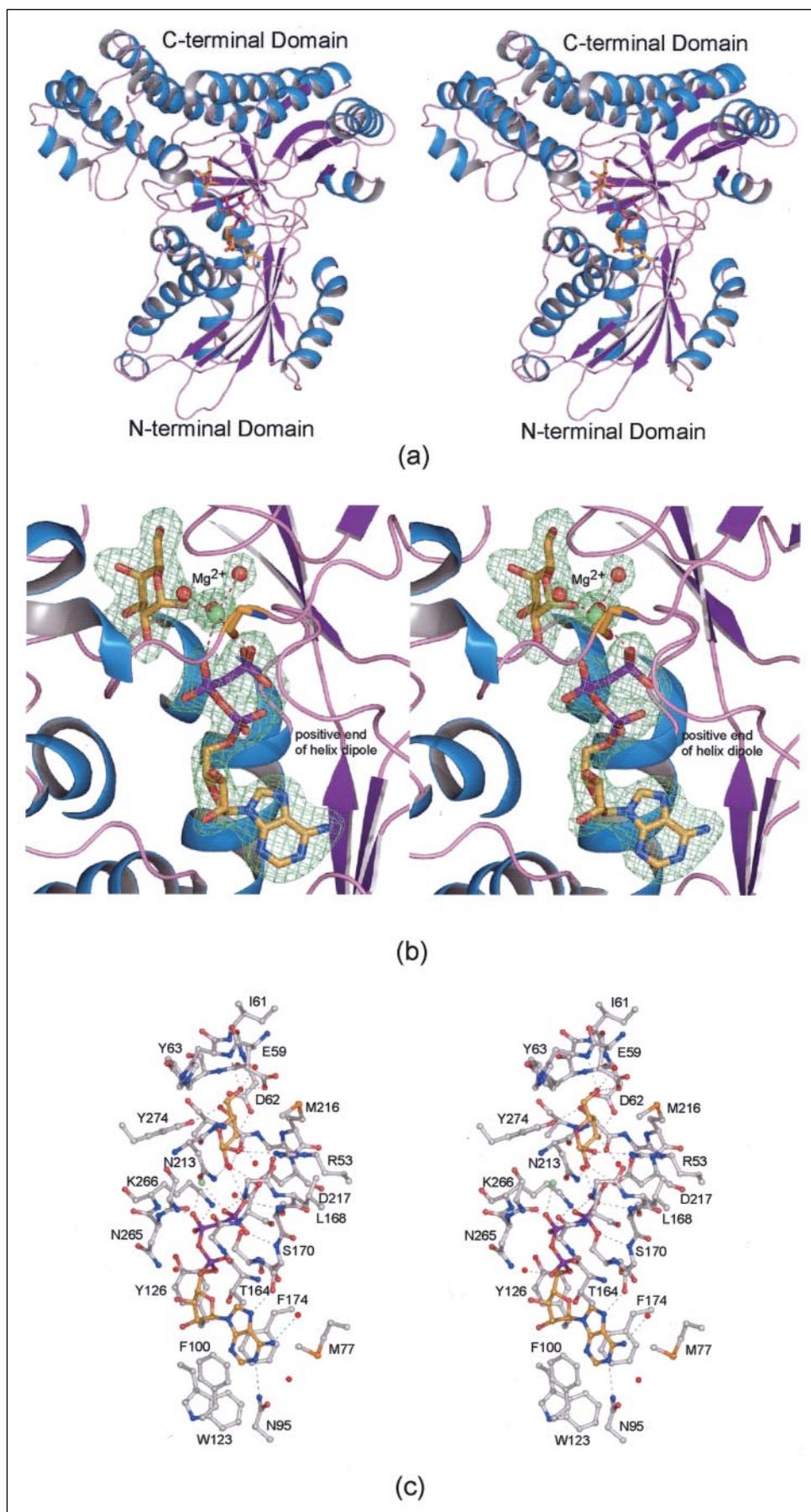FIGURE 1. **The structure of Gal1p.** *a*, *ribbon* representation of Gal1p with the AMPPNP and α-D-galactose ligands displayed in a *ball-and-stick* representation. The electron density corresponding to MgAMPPNP and galactose is displayed in *panel b*. The map, contoured at 3 σ, was calculated with coefficients of the form $F_o - F_c$, where F_o is the native structure factor amplitude and F_c the calculated structure factor amplitude from the model lacking coordinates for MgAMPPNP, galactose, and the three water molecules that form part of the coordination sphere of the metal. The ribose of the nucleotide adopts the $C_{3'}$-endo pucker, and the adenine ring is in the *anti*-conformation. *c*, a close-up view of the active site within ~3.7 Å of the ligands. The *black dashed lines* indicate potential hydrogen-bonding interactions. Water molecules are depicted as *red spheres*, and the position of the magnesium ion is indicated by the *green sphere*.
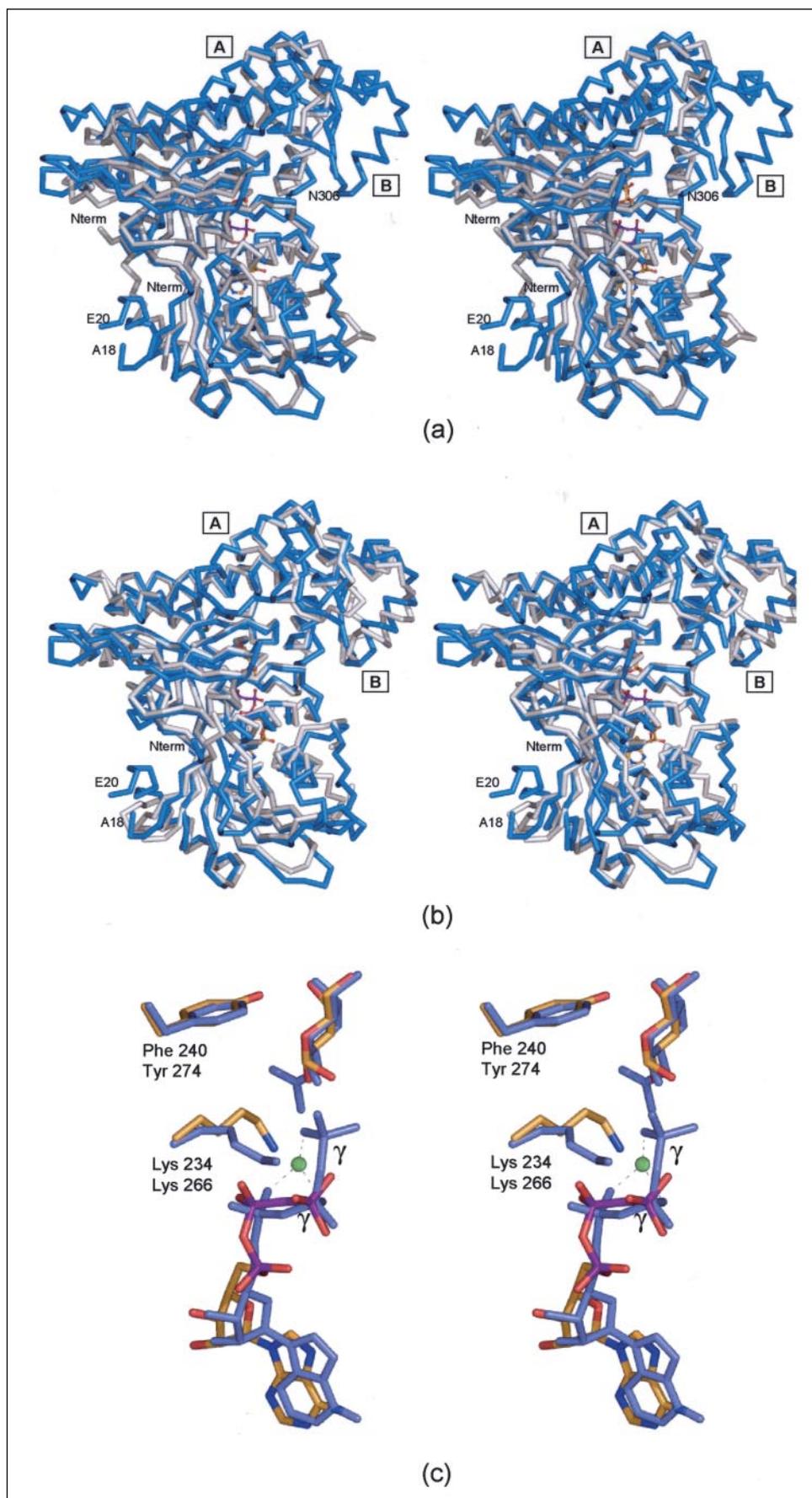
FIGURE 2. **Comparison of Gal1p to human galactokinase and GalNAc kinase.** *a*, a superposition of the polypeptide chains for Gal1p (in *blue*) and human galactokinase (in *gray*). X-ray coordinates for human galactokinase were determined in this laboratory (Protein Data Bank accession no. 1WUU). The two regions that differ significantly between these two enzymes are labeled *A* and *B*. The view is approximately a 180° rotation about the vertical axis from that shown in Fig. 1*a*. *b*, a superposition of the structures for Gal1p (in *blue*) and GalNAc kinase (in *gray*). Note that in this case, the two molecules differ significantly only in the region labeled *A*. Coordinates for GalNAc kinase were also determined in this laboratory (Protein Data Bank accession no. 2A2D). *c*, the different conformations observed for the AMPPNP nucleotides in Gal1p (*yellow* bonds) *versus* GalNAc kinase (*blue* bonds) are highlighted. The *top* and *bottom* residue labels refer to GalNAc kinase and Gal1p, respectively.
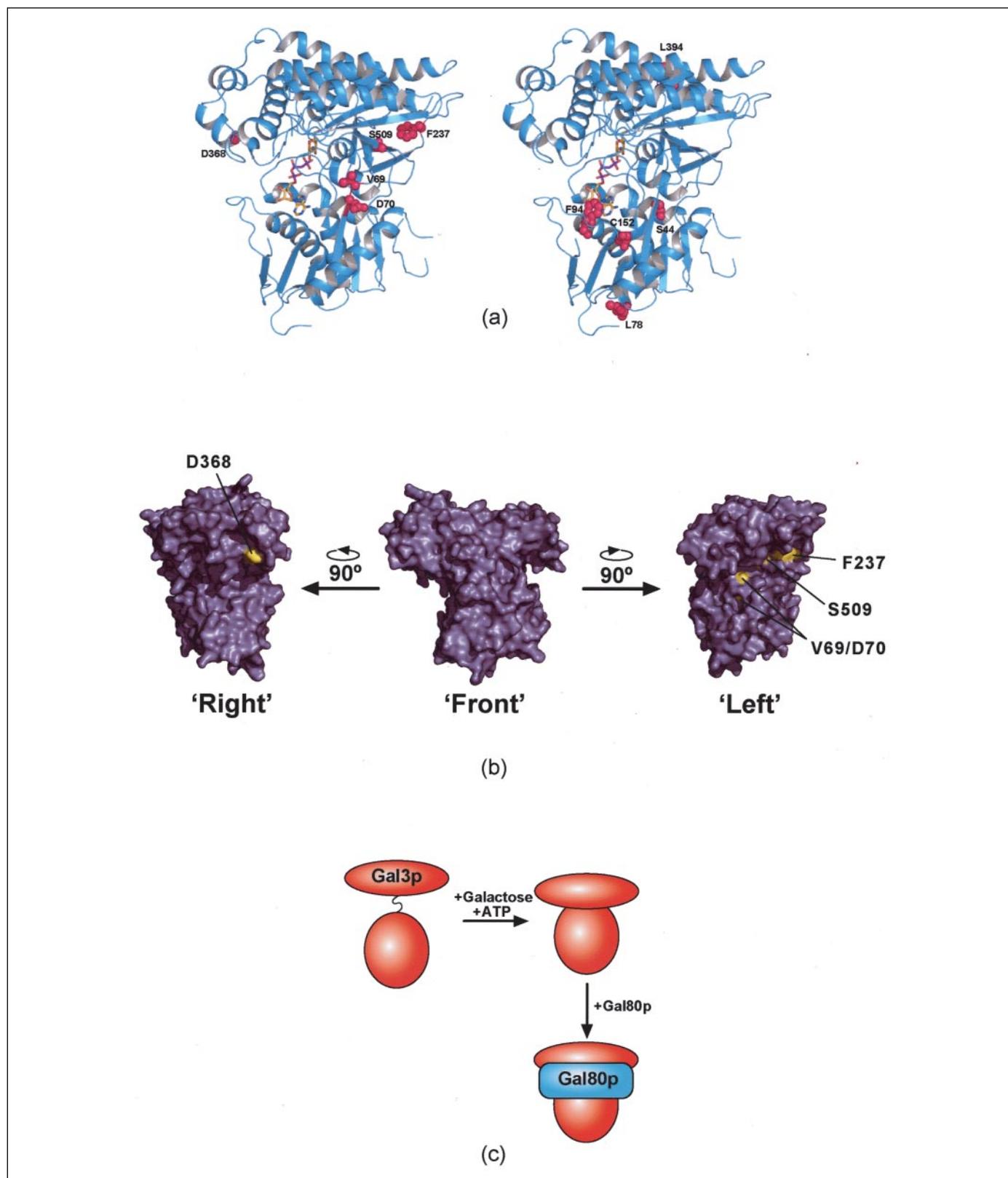
FIGURE 3. **Homology model for Gal3p. Based on the structure of Gal1p, a Gal3p homology model was constructed.** *a*, the locations of the Gal3p constitutive mutations (*magenta*) mapped onto the model of Gal3p are shown in the *left panel*. In the *right panel*, the known constitutive mutations in the *K. lactis* Gal1p (*magenta*) are mapped onto the model of Gal3p. *b*, a surface rendering of the Gal3p is depicted with the *front* of the molecule showing the same orientation as in *a*. The constitutive mutations are highlighted in *yellow* on models that have been rotated 90° around the vertical axis, either clockwise or counterclockwise. *c*, a potential model for the interaction of Gal3p and Gal80p in the presence of galactose and ATP. Galactose and ATP bind to Gal3p to induce the formation of a "rigid" protein structure in which the N- and C-terminal domains of the protein are held together. Gal80p binds at the interface of these domains and interacts with Gal3p at a minimum of two separate points located on opposite sides of the protein. As a result of this interaction, Gal80p is now rendered incapable of inhibiting the transcriptional activation properties of Gal4p.

built on the basis of the Gal1p structure. Such a model is presented in Fig. 3*a*. A number of mutants of Gal3p have been isolated that give rise to a transcriptionally altered phenotype. These include noninducible mutants that are defective in their interaction with Gal80p and constitutive mutants that have a reduced requirement for galactose and ATP to interact with Gal80p (5, 6). The noninducible mutants (6) are poorly defined and, presumably, could arise from a number of defects in protein structure or function. The *GAL3* constitutive mutants are, however, more interesting because they represent a gain of function. The presence of these mutations *in vivo* results in the expression the *GAL* genes in the absence of galactose, and *in vitro* the mutant proteins form a complex with Gal80p that is not dependent on the presence of galactose, although some of the mutants do require ATP to interact with Gal80p (5). Four *GAL3* constitutive mutants have been identified previously: V69E/D70V, F237Y, D368V, and S509P or S509L (5). In the Gal3p homology model depicted in Fig. 3*a* (*left panel*), each of these maps to the interface between the N- and C-terminal domains of Gal3p. Residues 69/70, 237, and 509 lie within 16 Å of each other along a groove on one side between the two domains, which is highlighted in Fig. 3*b*. Residue 368 is located on the opposite face of the protein but still at the interface of the domains. Additionally, a number of constitutive mutants in the *Kluyveromyces lactis* Gal1p protein, which functions both as a galactokinase and a transcriptional inducer, have been identified (24, 25). The equivalent residues at the sites of these mutations are not always conserved in the *S. cerevisiae* protein, but where they are they are indicated in Fig. 3*a* (*right panel*). The *K. lactis* constitutive mutations are more dispersed on the homology model and appear to fall into two categories. The first (including the residues equivalent to Asp-293 and Asn-294) fall at the interface of the N- and C-terminal domains of Gal3p. This region is disordered in the Gal1p model. Other constitutive *K. lactis* Gal1p mutants (*e.g.* equivalent residues to Ser-44, Phe-94, and Cys-152) occur close to the nucleotide-binding site and may mimic the state of the protein when the nucleotide is bound. Other constitutive mutants (*e.g.* the residues equivalent to Leu-78 and Leu-394) occur at locations distant to either the N- and C-terminal interface or the ligand binding sites. The structural and functional consequences of these mutations are not readily explained.

On the basis of our structural data for *S. cerevisiae* Gal1p and the corresponding model for Gal3p, we propose a model for the galactose- and ATP-dependent interaction with Gal80p (Fig. 3*c*). To date, the structures of galactokinases have been solved only in the presence of either (or both) of the small molecule ligands. For Gal3p, the binding of both ligands represents the state in which the protein is competent for interaction with Gal80p. Attempts to crystallize galactokinases in the absence of ligands have proved to be unsuccessful to date in this laboratory. We therefore suggest that in the absence of galactose and ATP, the N- and C-terminal domains of Gal3p, as in its galactokinase counterparts, are flexible with respect to each other. Upon binding of the ligands, a more rigid structure is adopted, and the interface between the two structural domains of the protein forms the binding site(s) for Gal80p. That Gal1p has an ordered, sequential mechanism provides evidence for a ligand-induced conformational change in this molecule (26). It is possible that parts of Gal80p fit into the groove between the two domains, the so-called "lips" as suggested by Menezes *et al.* (25) in their model of the *K. lactis* Gal1p based on the structure of the rather poorly conserved mevalonate kinase from *Methanococcus jannaschii* (27). However, we favor a model in which Gal80p contacts Gal3p at multiple points around the interface, highlighted in Fig. 3*b* between the domains of Gal3p. We suggest that the majority of the constitutive Gal3p mutations hold the protein in the rigid conformation that enables

the binding of Gal80p. From a structural point of view, it is not clear if all of the *K. lactis* Gal1p constitutive mutants follow this pattern. It is, of course, possible that *S. cerevisiae* Gal80p interacts somewhat differently with Gal3p than *K. lactis* Gal80p does with Gal1p. Indeed, it has previously been shown that *K. lactis* Gal80p is unable to fully functionally substitute for the *S. cerevisiae* protein. When expressed in place of *S. cerevisiae* Gal80p, the *K. lactis* protein is capable of repressing the transcriptional activity of Gal4p but is refractory to induction by Gal3p (7).

The *GAL* genetic switch is often considered a paradigm for the control of gene expression in eukaryotes (28). The x-ray structure of Gal1p reported here provides a new and important molecular scaffold for further biochemical investigations into the mode of action of Gal3p and its interactions with Gal80p. At present, nothing is known regarding which parts of the Gal3p and Gal80p polypeptide chains interact with one another. The model presented here is the simplest possible explanation whereby all of the constitutive mutations act in the same manner by locking Gal3p into the required conformation for Gal80p binding. This is unlikely to be the case, however, and there are bound to be differences in the manner in which some of the Gal3p mutations cause the constitutive phenotype. It is conceivable, for example, that Gal80p binds to only one face of Gal3p and that mutations in the other face are required for Gal3p to adopt the rigid conformation (Fig. 3*c*) required for this interaction to occur. Research in our laboratories is currently aimed at more fully addressing these issues, which are critical for understanding this exquisitely regulated genetic switch.

REFERENCES

1. Holden, H. M., Rayment, I., and Thoden, J. B. (2003) *J. Biol. Chem.* **278,** 43885–43888
2. Reece, R. J. (2000) *Cell. Mol. Life Sci.* **57,** 1161–1171
3. Sherman, J. R., and Adler, J. (1963) *J. Biol. Chem.* **238,** 873–878
4. Lohr, D., Venkov, P., and Zlatanova, J. (1995) *FASEB J.* **9,** 777–787
5. Blank, T. E., Woods, M. P., Lebo, C. M., Xin, P., and Hopper, J. E. (1997) *Mol. Cell. Biol.* **17,** 2566–2575
6. Suzuki-Fujimoto, T., Fukuma, M., Yano, K. I., Sakurai, H., Vonika, A., Johnston, S. A., and Fukasawa, T. (1996) *Mol. Cell. Biol.* **16,** 2504–2508
7. Zenke, F. T., Engles, R., Vollenbroich, V., Meyer, J., Hollenberg, C. P., and Breunig, K. D. (1996) *Science* **272,** 1662–1665
8. Bhat, P. J., and Hopper, J. E. (1992) *Mol. Cell. Biol.* **12,** 2701–2707
9. Platt, A., Ross, H. C., Hankin, S., and Reece, R. J. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 3154–3159
10. Adams, B. G., and Dalbec, J. M. (1977) *Biochem. Biophys. Res. Commun.* **74,** 1348–1354
11. Rotman, B., and Spiegelman, S. (1953) *J. Bacteriol.* **66,** 492–497
12. Bhat, P. J., Oh, D., and Hopper, J. E. (1990) *Genetics* **125,** 281–291
13. Platt, A., and Reece, R. J. (1998) *EMBO J.* **17,** 4086–4091
14. Yount, R. G., Babcock, D., Ballantyne, W., and Ojala, D. (1971) *Biochemistry* **10,** 2484–2489
15. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* **53,** 240–255
16. Cowtan, K., and Main, P. (1998) *Acta Crystallogr. Sect. D* **54,** 487–493
17. Roussel, A., Fontecilla-Camps, J. C., and Cambillau, C. (1990) *Acta Crystallogr. Sect. A* **4,** 66–67
18. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A* **43,** 489–501
19. Timson, D. J., Ross, H. C., and Reece, R. J. (2002) *Biochem. J.* **363,** 515–520
20. Thoden, J. B., and Holden, H. M. (2005) *J. Biol. Chem.* **280,** 32784–32791
21. Thoden, J. B., Timson, D. J., Reece, R. J., and Holden, H. M. (2005) *J. Biol. Chem.* **280,** 9662–9670
22. Thoden, J. B., and Holden, H. M. (2003) *J. Biol. Chem.* **278,** 33305–33311
23. Hartley, A., Glynn, S. E., Barynin, V., Baker, P. J., Sedelnikova, S. E., Verhees, C., de Geus, D., van der Oost, J., Timson, D. J., Reece, R. J., and Rice, D. W. (2004) *J. Mol. Biol.* **337,** 387–398
24. Vollenbroich, V., Meyer, J., Engels, R., Cardinali, G., Menezes, R. A., and Hollenberg, C. P. (1999) *Mol. Gen. Genet.* **261,** 495–507
25. Menezes, R. A., Amuel, C., Engels, R., Gengenbacher, U., Labahn, J., and Hollenberg, C. P. (2003) *J. Mol. Biol.* **333,** 479–492
26. Timson, D. J., and Reece, R. J. (2002) *Biochimie* (*Paris*) **84,** 265–272
27. Yang, D., Shipman, L. W., Roessner, C. A., Scott, A. I., and Sacchettini, J. C. (2002) *J. Biol. Chem.* **277,** 9462–9467
28. Sellick, C. A., and Reece, R. J. (2005) *Trends Biochem. Sci.* **30,** 405–412