INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

Molecular Characterization and Transcriptional Regulation of *GCV3*, the *Saccharomyces cerevisiae* gene Encoding the H-protein of the Glycine Cleavage System.

Nagarajan Lakshmanan

A thesis in the
Department of Chemistry and Biochemistry

Presented in Partial Fulfilment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University,
Montreal, Quebéc, Canada

August 1997

ABSTRACT

Molecular Characterization and Transcriptional Regulation of *GCV3*, the *Saccharomyces cerevisiae* gene Encoding the H-protein of the Glycine Cleavage System.

Nagarajan Lakshmanan

YAL044, a gene on the left arm of *Saccharomyces cerevisiae* chromosome one, is shown to code for the H-protein subunit of the multienzyme glycine cleavage system. The gene designation has therefore been changed from *YAL044* to *GCV3* to reflect its role in the glycine cleavage system. *GCV3* encodes a 177 amino acid residue protein with a putative mitochondrial targeting sequence at its amino terminus. Targeted gene replacement shows that *GCV3* is not essential for growth on minimal media. It is, however, essential for growth when glycine serves as the sole nitrogen source. Studies of *GCV3* expression revealed that it is highly regulated. Supplement with glycine, the glycine cleavage system's substrate, induced expression at least 30-fold. In contrast, addition of the C1-metabolic end products repressed expression about 10-fold.

The regulation of glycine cleavage system activity reflects the availability of glycine and the cellular demand for its metabolic products. In addition the glycine cleavage system has been shown to be important for the growth and viability of organisms ranging from microorganisms such as *E. coli* and *S. cerevisiae* to humans. Although, this system is important and its activity highly regulated little was known about the transcriptional regulatory mechanisms that control its activity.

To address this I have examined the transcriptional regulation of the *S. cerevisiae* *GCV3* gene. The results presented here show that at least six different transcriptional activators control *GCV3* expression. These include: an as yet unidentified activator that is partially responsible for its induction by glycine; Gcn4p the transcriptional activator that mediates general amino acid control; Gln3p which is involved in the activation of nitrogen regulated genes; Gcr1p, a transcription factor important for the expression of glycolytic genes; Bas1p/Bas2p which cooperatively mediates the glycine-dependent expression; and an as yet unidentified factor that represses expression regardless of the growth condition. Additional evidence suggests that Rap1p, Nil1p, Acr1p, Ure2p, and Dal80p also regulate *GCV3* regulation.

ACKNOWLEDGEMENTS

I gratefully acknowledge my mentor Dr.R.K.Storms, for his invaluable supervision, patience, friendship, and advise throughout the course of my research. I am indebted to my committee members Drs. Paul Joyce and Adrian Tsang for their invaluable comments.

I would also like to express my thanks to all members of Biology and Chemistry Departments, particularly Drs.Herrington, Gulick, Ibrahim, Newman, Tsang, Cupples, Kornblatt, Turnbull, Joyce, Mangat, and Kapoor for helping me to use their laboratory facilities whenever I needed.

I would like to thank my colleagues Dr. Sophia Ushinsky, Dr. Bruce Williams, Edith Munro, Paul Taslimi, and Tom Downing for their refreshing view points which often helped me to overcome obstacles in my research.

Finally, I would like to thank the following scientists from other institutions: Drs. Fink, Struhl, Emr, Braus, Moye-Rowley, Madhani, Springer, Arndt, Holland, Guarente, Rowen, Melcher, Magasanik, and Hinnebusch for their valuable yeast strains and plasmids.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

List of Abbreviations

5,10-MTHF	N^5N^{10}-methylene tetrahydrofolate
bp	Base pairs
BSA	Bovine serum albumin
C1-units	Activated single-carbon units
FAD	Flavin adenine dinucleotide
FADH	Flavin adenine dinucleotide reduce form
GAP	General amino acid permease
GCRE	General control response element
GCS	Glycine cleavage system
H-protein	Hydrogen carrier protein of the glycine cleavage system
Inr	Transcription initiation element
L-protein	Dihydrolipoamide dehydrogenase
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide reduced
NKH	Non-ketotic hyperglycinaemia
ORF	Open reading frame
P-protein	Glycine dehydrogenase
PCR	Polymerase chain reaction

PLP	Pyridoxal phosphate
PMSF	Phenylmethylsulfonyl fluoride
SCD	Single carbon donor
SD	Synthetic dextrose with 2% glucose
SH	Sulfhydryl group
T-protein	Aminomethyltransferase
TAFs	TBP-associated factors
TBP	TATA binding protein
THF	Tetrahydrofolate
UAS	Upstream activation sequence
UIS	Upstream induction sequence
URS	Upstream repressor sequences
YGSC	Yeast genetic stock centre
YNB	Yeast nitrogen base

INTRODUCTION

The results presented in this thesis describe my analysis of the glycine cleavage system (GCS) of *Saccharomyces cerevisiae*. It begins by describing the identification and cloning of the gene (*GCV3*, Bussey., et al 1995) that codes for the H-protein subunit of the GCS. Once cloned, this gene was used to study the regulation and physiological importance of the GCS.

A.1. Glycine cleavage system

The GCS is a multienzyme complex consisting of four different subunits, the P-, H-, T- and L-proteins (Fig.1). These subunits are coded for by four unique genes (*GCV1, GCV2, GCV3* and *LPD1*). The genes which encode these proteins, their chromosomal locations, the predicted molecular mass of their products and their Genbank accession numbers are listed in Table 1. All the genes encoding the proteins of the glycine cleavage system, except the *LPD1* gene, were identified by searching the DNA sequence of the yeast genome. The functions of the protein products of *GCV1*, *GCV2* and *GCV3* appear to be restricted to their roles in the GCS. In addition to its role in the glycine cleavage system, the L-protein serves a similar function in both the pyruvate dehydrogenase complex, and the α-ketoglutarate dehydrogenase complex.

The GCS is also known as the "glycine decarboxylase" or "serine synthase" system. The biochemistry of glycine cleavage has been studied in organisms ranging from *Escherichia coli* (Okamura-Ikeda et al., 1993) to plants like peas (Macherel et al., 1990)

Fig. 1.

Schematic representation of the glycine cleavage system (GCS) multienzyme complex (taken from Oliver, 1994). P-protein, glycine decarboxylase; H-protein, hydrogen carrier protein; T-protein, aminomethyltransferase; L-protein, dihydrolipoamide dehydrogenase; THF, tetrahydrofolate; PLP, pyridoxal phosphate; SH, sulfhydryl group; NAD, nicotinamide adenine dinucleotide; NADH, nicotinamide adenine dinucleotide reduced form; FAD, flavin adenine dinucleotide, FADH, flavin adenine dinucleotide reduce form. H-protein plays a central role through its lipoamide cofactor, which carries the reaction intermediates between the reactive sites of the P-protein, T-protein, and L-protein.

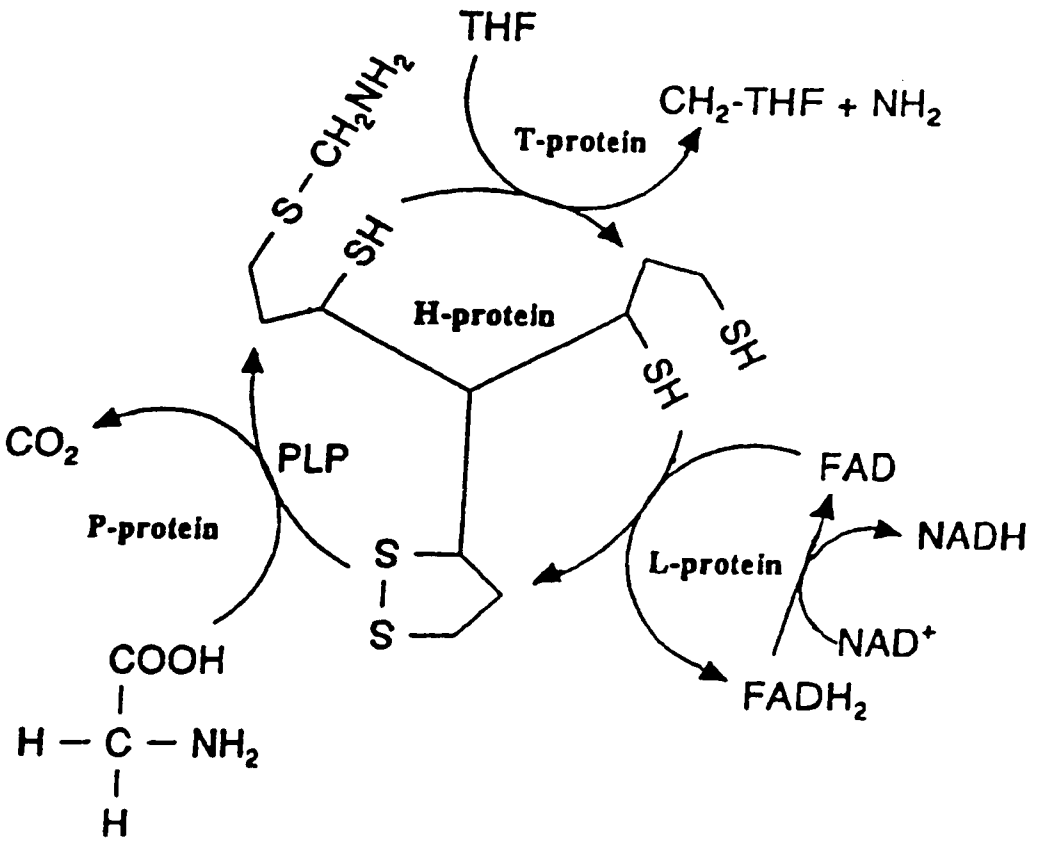

THF
CH2-THF + NH2
T-protein
S – CH2NH2
SH
H-protein
SH
SH
CO2
PLP
P-protein
FAD
NADH
L-protein
NAD+
FADH2
S
S
COOH
H – C – NH2
H

Table 1.

Properties the of *GCV1*, *GCV2*, *GCV3*, and *LPD1* genes.

Gene	Protein and Molecular weight[a]	Chromosome location	Intergenic region[b] (nt)	Reference
GCV1	T 44469	IV	389	McNeil et al., 1997
GCV2	P 114451	XIII	637	Sinclair and Dawes, 1995
GCV3	H 19588	I	211	Nagarajan and Storms, 1997
LPD1	L 54010	VI	863	Sinclair et al., 1996

[a] Molecular weight (da).

[b] Size of the promoter (bp).

and *Arabidopsis thaliana* (Srinivasan and Oliver, 1992), to higher eukaryotes like chickens (Yamamoto et al., 1991), cows (Fujiwara et al., 1990) and humans (Fujiwara et al., 1991). In contrast very little is known about how the glycine cleavage system is regulated, particularly in eukaryotes. The recent identification and cloning of the yeast genes coding for the four polypeptides of the glycine cleavage system, T-protein (McNeil et al., 1997), P-protein (Sinclair and Dawes, 1995), L-protein (Sinclair et al., 1996) and H-protein (this study), facilitate the detailed molecular analysis of the glycine cleavage system and its regulation in a genetically tractable eukaryote.

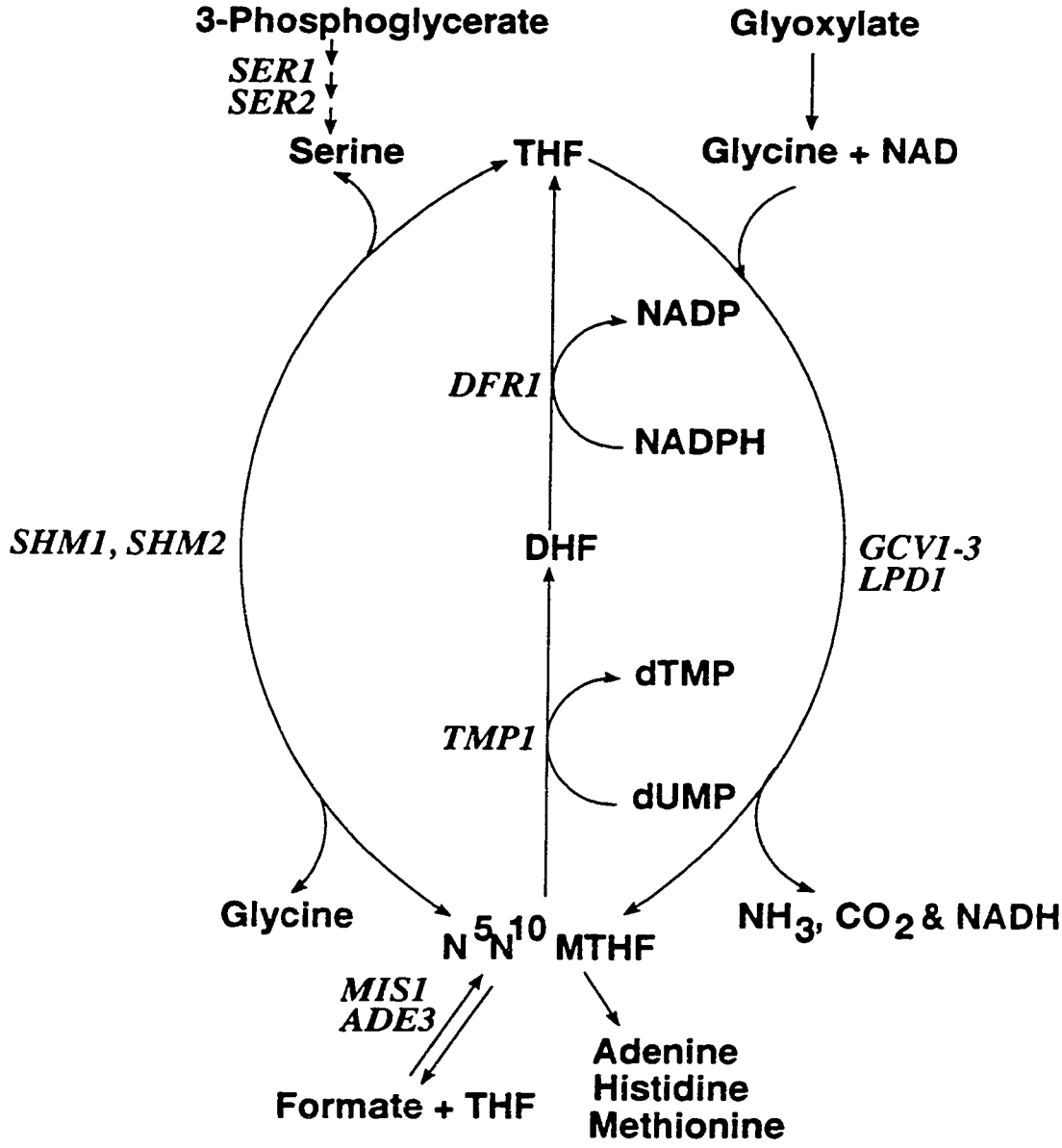
The glycine cleavage system catalyses the oxidative cleavage of glycine into CO_2 and NH_3. The concomitant transfer of methylene carbon unit to THF (tetrahydrofolate) generates the C1-donor 5,10-MTHF (methylene tetrahydrofolate) (Fig. 2). P-protein, a pyridoxal phosphate dependent decarboxylase, catalyses the release of carbon dioxide from glycine and transfers the resulting methylamine intermediate to the lipoic acid prosthetic group of the H-protein. The lipoic acid is reduced as a result of this reaction. T-protein catalyses the release of ammonia from the methylamine H-protein complex and the transfer of a single carbon methylene unit to THF. In the reaction catalyzed by the L-protein the dihydrolipoyl residue of the H-protein is converted to lipoic acid with the coupled reduction of NAD+ to NADH (Oliver, 1994).

A.2. Single Carbon Donor (SCD) biosynthesis

Aside from the glycine cleavage system there are two additional mechanisms for the synthesis of 5,10-MTHF (Fig. 2) (Barlowe and Appling, 1990). One utilizes glycine

Fig. 2.

Selected biochemical reactions involved in the biosynthesis and interconversion of serine, glycine and 5,10-MTHF. Biochemical reactions are designated by the gene or genes that code for the protein(s) that catalyses each reaction. *GCV1*, *GCV2, GCV3* and *LPD1*, (encode the glycine cleavage system); *TMP1*, (thymidylate synthase); *DFR1*, (dihydrofolate reductase); *SHM1*, (mitochondrial glycine hydroxymethyltransferase); *SHM2,* (cytoplasmic glycine hydroxymethyltransferase); *ADE3,* (cytoplasmic C1-tetrahydrofolate synthase); *MIS1,* (mitochondrial C1-tetrahydrofolate synthase); *SER1* (3-phosphoserine transaminase) and *SER2*, (3-phosphoserine phosphatase). dUMP and dTMP represent deoxyuridine 5'-monophosphate and deoxythymidine 5'-monophosphate.


3-Phosphoglycerate
SER1
SER2
Serine
THF
Glyoxylate
Glycine + NAD
NADP
DFR1
NADPH
SHM1, SHM2
DHF
GCV1-3
LPD1
dTMP
TMP1
dUMP
Glycine
N5N10 MTHF
NH3, CO2 & NADH
MIS1
ADE3
Formate + THF
Adenine
Histidine
Methionine

hydroxymethyltransferase and the other utilizes C1-tetrahydrofolate synthase. The multifunctional enzyme C1-tetrahydrofolate synthase catalyzes the interconversion of 5,10-MTHF, 5,10-methenyltetrahydrofolate, 10-formyltetrahydrofolate, and THF with the concomitant production of formate (Barlowe and Appling, 1990; and Pasternack et al., 1992). These C1 compounds in turn are used in the biosynthesis of a number of compounds including adenine, thymidylate and the amino acids histidine, serine, methionine and formylmethionine (Mudd and Cantoni, 1964).

Genetic analysis has also shown that no single mechanism is essential for the production of 5,10-MTHF (McNeil et al., 1996). However, since inactivation of both the glycine hydroxymethyltransferase and glycine cleavage system routes of 5,10- MTHF synthesis renders the growth of *S. cerevisiae* conditional upon supplement with formate, the biosynthesis of formate is apparently dependent upon the interconversions of 5,10-MTHF catalyzed by C1-tetrahydrofolate synthase (Staben and Rabinowitz, 1986). It, therefore, appears that the major role for formate is the shuttling of C1-units between cellular compartments. Consistent with this role is the observation that C1-THF synthase activity is localized to both the cytoplasmic and mitochondrial compartments (Fig. 2) (Pasternack et al., 1992). Indeed the nuclear genome of *S. cerevisiae* encodes the *MIS1* gene for mitochondrial and the *ADE3* gene for cytoplasmic versions of C1-tetrahydrofolate synthase (Shannon and Rabinowitz, 1988). Similarly, there are two genes for glycine hydroxymethyltransferase, one for a cytoplasmic and the other for a mitochondrial version (McNeil et al., 1994). In contrast, for those systems studied to date the glycine cleavage system is localized to only the mitochondrial compartment (Oliver, 1994; Kim and Oliver, 1990; Rebeille et al., 1994).

That the C1-tetrahydrofolate synthase-dependent mechanism is apparently solely responsible for the generation of formate in *S.cerevisiae* makes formate biosynthesis dependent upon 5,10-MTHF. In contrast, for the two other sources of C1-units, serine and glycine, there are at least two biosynthetic pathways (Melcher and Entian, 1992; Melcher et al., 1995). One pathway for serine synthesis involves the conversion of 3-phosphoglycerate to serine by three enzymatic reactions. The last two of these reactions are catalyzed by the *SER1* and *SER2* gene products. A second pathway for serine synthesis is catalyzed by glycine hydroxymethyltransferase and utilizes glycine and 5,10-MTHF as its substrates. Mutants blocked in the synthesis of serine from 3-phosphoglycerate can utilize glycine to provide both C1 units via the glycine cleavage system and serine via the reaction catalyzed by glycine hydroxymethyltransferase (Fig. 2). Glycine can also be synthesized in two ways. One utilizes glycine hydroxymethyltransferase to convert serine into glycine. The other pathway, which apparently derives glycine from glyoxylate, is dependent upon the *GLY1* gene (McNeil et al., 1996). This second route of glycine biosynthesis is apparently inefficient since *ser1* mutants grow very poorly on minimal media (McKenzie and Jones, 1977).

A.3. Non-ketotic hyperglycinaemia

The GCS plays a physiologically important role. In humans, non-ketotic hyperglycinaemia (NKH) is an inborn error of metabolism in humans that results in defective glycine cleavage system activity. This genetic disease causes profound neurological impairment. An effective therapy for treating this disorder has not been found. People who

suffer from this disease accumulate 10- to 30-fold elevated levels of glycine in plasma, urine, and particularly cerebrospinal fluids (Hamosh et al., 1995). It is not well understood how excess glycine leads to irreversible neurological impairment. Neurophysiological evidence indicates that the excess glycine could act as an inhibitory neurotransmitter (Tada, 1992; 1993; Toone, 1992; 1994). This disease typically presents itself in the newborn and is fatal without intervention. Consequently, most couples with a pregnancy known to be at risk for NKH request prenatal diagnosis.

Other presentations of NKH are recognized: notably a transient neonatal type and a late-onset type (Tada and Hayasaka, 1987). Rare survivors of the classical neonatal type of non-ketotic hyperglycinaemia have profound intellectual impairment and suffer from seizures that are difficult to control (Hamosh et al., 1995).

GCS mutants also adversely affect *S. cerevisiae* (this study). Using an experimentally tractable system to study the physiological consequences of defects in the GCS should provide important insights into the NKH in humans and perhaps suggest effective treatments.

A.4. Transcriptional machinery

In eukaryotes there are three RNA polymerases. One of these, RNA polymerase II, is responsible for the transcription of the protein coding genes (Struhl, 1996b). The transcriptional machinery which transcribes these protein coding genes is complex. In addition to RNA polymerase II it involves an extensive set of auxiliary factors. These auxiliary factors include the general transcription factors (proteins required for the expression of all pol II genes), and a series of more specific transcription factors (factors that

are required for one or a limited subset of pol II genes). These specific transcription factors are often called "gene specific" transcription factors (Struhl, 1989; 1993; Guarente, 1987).

For most protein encoding genes RNA polymerase II and the general transcription factors form a complex that is assembled on the promoter region at a site called the TATA box (Struhl, 1995). The TATA box is usually 30 to 50 base pairs upstream of the site where transcription initiates (called the Inr site). The "gene specific" transcription factors bind to DNA sequences that are upstream of the TATA box. The upstream sites are gene specific and can be hundreds or even thousands of base pairs away from the TATA box (Guarente, 1987; 1995; Struhl, 1989).

A.4.1. "Gene specific" transcription factors

"Gene specific" transcription factors can be divided into two classes, transcriptional activators and transcriptional repressors. Essentially all the "gene specific" transcription factors harbour a domain that can bind to DNA in a sequence specific manner. In addition transcriptional activators usually harbour an acidic domain that can activate transcription (Struhl 1996a). "Gene specific" transcription factors also often harbour a region or regions that can respond to specific regulatory signals so that their own activity can be regulated. For this, signal transduction pathways monitor intracellular and extracellular conditions and ultimately interact with the "gene specific" transcription factors to regulate their activity. These pathways begin with the generation of some internal or external signal and ultimately impinge on the activator and/or repressor proteins at specific sites or regions on the proteins themselves. Alternatively a regulatory protein may be controlled at the level of its translation

or transcription.

"Gene specific" transcriptional activators regulate transcription by interacting with the general transcription factors assembled at the TATA box. How do transcriptional activators which can be bound to promoter sites thousands of base pairs away from the TATA box function over such large distances? Although several models have been put forward, the most widely accepted model involves DNA looping. By looping out the intervening DNA, transcriptional regulators communicate by direct physical contact with the rest of the transcriptional machinery (Ptashne, 1988; Rippe et al., 1995; Struhl, 1993; 1996a; Stargell and Struhl, 1996).

A.4.2. "General Transcription" factors

Most promoters have the sequence TATA located about 30 bp upstream of the Inr. The TATA-box is the binding site for TATA binding protein (TBP). TBP when bound to the TATA-box has 9 additional proteins associated with it. TBP together with these associated proteins called TBP-associated factors (TAFs) constitute the complex TFIID (transcription factor D for RNA polymerase II). TFIID has a mass of about 800 kD. The binding of TFIID facilitates the binding of at least 6 additional general transcription factors and RNA polymerase II. These bind in the following order TFIIA, TFIIB, TFIIF and RNA polymerase II, TFIIE and finally TFIIJ and TFIIH. Together these 7 general transcription factors and RNA polymerase II when assembled at a TATA-box constitute a basal promoter complex. A basal complex cannot initiate transcription *in vivo*. The initiation of transcription requires the interaction of the "basal complex" with gene specific transcriptional activators

bound at UAS sites.

A.5. Yeast Promoter Elements

RNA polymerase II, the "general transcription" factors, and the "gene specific" transcription activators and repressors influence gene expression by binding directly to promoter DNA sequences or by interacting with another protein that has promoter region binding activity. The binding of these "general" and "gene specific" proteins to promoters requires the presence of DNA sequences to which they can bind.

Yeast promoter elements that facilitate the binding of proteins are typically 10-30 bp in length. These promoter elements include the upstream activating sequences or UAS for the binding of gene specific transcriptional activators, upstream repressor binding sequences (URS) for the binding of gene specific repressors, the TATA-box which is the binding site for the general transcription factor TBP, and the transcription initiation element (Inr) (Struhl, 1995). Recently, the existence of another element the UIS (upstream induction sequence) was reported (van Vuuren et al, 1991). UAS and UIS elements positively affect gene expression. In contrast, the URS is a negative regulatory element. Similar to mammalian enhancers the upstream elements function in an orientation independent manner.

Upstream elements are usually the major determinant of a particular gene's transcriptional regulation in response to various physiological conditions. Genes subject to a common control mechanism (coordinately regulated genes) contain upstream elements that are similar in DNA sequence, whereas non-coordinately regulated genes contain upstream elements with different DNA sequences. For example, all those genes sensitive to the

general amino acid control system (discussed in detail below) contain one or more copies of the consensus sequence 5'-TGACTC-3' in their promoter region. In contrast, genes sensitive to nitrogen regulation contain the consensus sequence 5'-GATAA-3 (discussed in detail below), whereas the glycolysis pathway genes contain the consensus sequence 5'-CTTCC-3'.

The Inr is located near the actual mRNA start site. Many different sequences can function as an Inr. Normally a gene's Inr is located 25 to 30 bp downstream from the TATA element and has no strong consensus sequence. Therefore the DNA sequence requirement for a yeast Inr is poorly understood (Struhl, 1995). In contrast to UAS and TATA elements, the initiator element is relatively unimportant for determining the rate of transcription initiation (Struhl, 1995).

A.6. Transcriptional regulation is complex

Since the transcription of a gene is dependent upon the interaction of the basal complex with "gene specific" transcription factors, it is the "gene specific" regulatory proteins assembled at the upstream regulatory sites that control a gene's expression. The binding of "gene specific" regulatory proteins to a promoter requires the presence of sequence specific UAS, URS and UIS elements. It is the binding of "gene specific" regulatory proteins to these cis-acting elements that mediates the transcriptional control of a particular gene. Depending upon the physiological conditions, a trans-acting factor can be synthesized, activated, or inactivated to mediate the transcription or repression of a particular gene or group of genes (Nasmyth and Shore, 1987). Transcriptional regulation appears to be very complex for a number of reasons (Struhl, 1989; 1995; 1996b; Herschbach and Johnson,

1993; Kunzler et al., 1996). These can be summarized as follows: (i) More than one protein can bind to the same promoter region. For example, the consensus sequence recognized by the transcriptional activator Gcn4p (5'-TGACTC-3') has been shown to be included within the binding site used by at least six regulatory proteins. These are Gcn4p, Bas1p, Bas2p, Yap1p, Yap2p and Acr1p. The relevant information concerning these proteins are elaborated under the GCRE binding proteins section below. (ii) More than one protein is often required for a single DNA binding event. For example, both Hap2p and Hap3p are required to form a *CYC1* promoter-protein complex (Olesen et al., 1987), while Gcn4p interacts with DNA as a dimer (Sellers and Struhl, 1989). (iii) Multiple proteins are necessary for the activation or repression of gene transcription. For example, Bas1p and Bas2p (discussed in more detail below) bind to distinctly different DNA sequences but act in a cooperative fashion to activate transcription. Further, neither protein can activate transcription on its own (Tice-Baldwin et al., 1989; Daignan-Fornier and Fink, 1992). (iv) The same protein can activate or repress transcription depending upon the gene and/or physiological condition. For example, Rap1p enhances the Bas1p/Bas2p and Gcn4p dependent transcriptional activation of *HIS4* (Devlin et al., 1991). The same protein also functions as a repressor in the promoters of all ribosomal protein genes (Nasmyth and Shore, 1987). (v) Biochemical and genetic evidence indicate that the modulation of chromatin structure can greatly influence transcription (Struhl, 1996b; Wilson et al., 1996). (vi) Coactivators or supplementary regulatory molecules can modify the activity of a transcription factor (Guarente, 1995; Struhl, 1995; 1996b). Thus, other molecules whose activity and/or levels fluctuate in response to changing physiological states can regulate the activity of a transcription factor.

A.7. General amino acid control or General control

Yeast and bacterial amino acid biosynthetic genes are differentially regulated in response to the availability of amino acids in the growth medium. Bacteria can turn off the amino acid biosynthetic genes if the amino acids are available in the growth medium. Under similar conditions yeast cells maintain significant levels of transcription from amino acid biosynthetic genes. This is often referred to as basal level expression (Tice-Baldwin, et al., 1989; Daignan-Fornier and Fink, 1992).

Another difference between bacteria and yeast is the increased expression of a large number of yeast amino acid biosynthetic genes in response to starvation for a single amino acid. This increased expression of this large battery of genes is called "General Amino Acid Control" or "General Control" (Hinnebusch, 1988; 1990; 1992). Increased amino acid biosynthetic gene expression in response to amino acid starvation is mediated by increased production of Gcn4p. Gcn4p synthesis is regulated at the translational level (Hinnebusch, 1992). Amino acid starvation leads to a 30-fold increase in Gcn4p synthesis (Hinnebusch, 1992). In addition to the amino acid biosynthetic genes, Gcn4p can also mediate increased transcription of several nucleic acid and single carbon donor pathway genes (Stotz et al., 1993; Daignan-Fornier and Fink, 1992 and this study). Besides *GCN4*, there are at least eight *GCN* genes that cooperatively mediate the general control response. Mutations in *GCN4* can impair the general control response (Hinnebusch, 1992; Mösch et al., 1990).

A set of 13 *GCD* gene products is required to maintain repression during non-amino acid starvation conditions. The *GCD* gene products are believed to negatively regulate Gcn4p synthesis at the translational level. Recessive mutations in these *GCD* genes results

in the constitutive derepression of genes subject to general control (Hinnebusch, 1992).

A.8. The general amino acid control response element (GCRE) and GCRE binding proteins.

The structural genes whose transcription is subject to general control contain one or more copies of a short upstream element, which is approximately 12 bp in length. It contains a highly conserved six base pair core sequence (5'-TGACTC-3') which serves as a recognition site for Gcn4p. This element acts as a UAS and is referred to as a GCRE or Gcn4p Recognition Element (Hinnebusch, 1988; Hope and Struhl, 1985; Hope et al., 1988; 1987; Hill et al., 1986; Arndt and Fink, 1986; Arndt et al., 1987; Tice-Baldwin et al., 1989). This short nucleotide sequence located upstream of the start codon of a number of amino acid biosynthetic genes mediates transcriptional activation through Gcn4p binding (Hope and Struhl, 1985; Hill et al., 1986; Arndt and Fink, 1986).

In some promoter contexts, the GCRE hexanucleotide sequence has been shown to overlap the binding sites for one or more of the trans-acting factors Bas1p, Bas2p, Acr1p, Yap1p and Yap2p (Arndt et al,. 1987; Moye-Rowley et al., 1989; Vincent and Struhl, 1992; Harshman et al., 1988; Wu et al., 1993; Daignan-Fornier and Fink, 1992). This can result in two proteins competing for the same site. For example, the *TRP4* gene, during simultaneous phosphate and amino acid starvation, is not induced apparently because in low phosphate Bas2p interferes with Gcn4p binding to the *TRP4* GCRE located between -258 and -225 (Braus et al., 1989). All the proteins that can bind to sites that overlap GCRE sites do not necessarily bind to all GCREs. For example, sequences adjacent the GCREs hexanucleotide

core are also important for binding these six DNA binding proteins. Therefore, GCRE context affects the influence that a particular transcription factor has on its utilization as a regulator of gene expression.

Context may also be important because the activity of one transcription factor is dependent upon the binding of another protein. For example, the function of Bas1p depends upon Bas2p binding. The expression of some of these transcription factors is regulated in response to the physiological conditions. For example, during amino acid starvation Gcn4p expression increases about 30-fold. When present at induced levels, Gcn4p would compete more effectively for GCRE binding sites.

Bas1p is homologous to the oncoprotein Myb (Tice-Baldwin et al., 1989). Bas2p, also known as Pho2p and Grf10p, is a homeodomain type protein (Daignan-Fornier and Fink, 1992). Gcn4p is a leucine zipper protein (Agre et al., 1989) that shares homology with the jun oncoprotein and the human activation factor AP-1 (Mösch et al., 1990; Arndt et al., 1987). Bas1p and Bas2p jointly stimulate the Gcn4p-independent or "basal level" transcription of *HIS4* and purine biosynthetic pathway genes (Tice-Baldwin et al., 1989; Daignan-Fornier and Fink, 1992). "Basal level" expression usually refers to the expression observed under non-inducing conditions. The first demonstration that both "general control" and "basal control" were important for the regulation of a single gene was demonstrated using the *HIS4* gene (Devlin et al., 1991). The general control system increases *HIS4* transcription upon starvation for an amino acid. In the absence of amino acid starvation the "basal control" system regulates *HIS4* transcription (Devlin et al., 1991).

Bas2p cooperatively activates gene transcription by interacting with other DNA

binding proteins. Therefore, Bas2p is involved in several very different types of transcriptional control. For example, Bas2p cooperates with Bas1p to regulate *HIS4* and several adenine biosynthetic genes (Daignan-Fornier and Fink, 1992). Bas2p also functions in cooperation with Pho4p to induce *PHO5, PHO10, PHO11, PHO81* and *PHO84* in response to phosphate starvation (Johnston and Carlson, 1992; Shao et al., 1996). In yet another case, Bas2p and Swi5p bind to form a stable complex *in vitro* and cooperate to regulate *HO* expression (Brazas and Stillman, 1993). Interestingly, the transcriptional control by the Bas1p/Bas2p protein pair is regulated, although it is called "basal control" (a term that suggests constitutive expression). For example, the Bas1p/Bas2p dependent regulation of *HIS4* transcription is down regulated in response to adenine supplement (Trice-Baldwin et al., 1989).

A.9. Nitrogen regulation

Depending upon the quality and abundance of the available source of nitrogen, *S. cerevisiae* adjusts its expression of the enzymes for utilizing the various nitrogen sources. The regulatory mechanisms underlying this control are in large part executed at the level of transcription (Wiame et al., 1985). For example, in the presence of "rich nitrogen" sources such as glutamine, asparagine, and ammonia the genes encoding enzymes required for the utilization of "poor nitrogen" sources such as arginine, proline and urea are expressed at very low levels. This phenomenon is referred to as nitrogen catabolite repression (Bysanti et al., 1991; Coffman et al., 1994; 1995; 1996; Rai et al., 1995; Cooper et al., 1982) or nitrogen regulation (Blinder et al., 1995; 1996; Stanbrough and Magasanik, 1995; 1996; Stanbrough

et al., 1995; Magasanik, 1992).

To date five *S. cerevisiae* proteins, Gln3p, Ure2p, Dal80p, Nil1p (Gta1p), and Gap1p have been shown to be involved in nitrogen regulation (Cunningham and Cooper, 1991; 1993; Rai et al., 1995; Cunningham et al., 1994; 1996; Mineheart and Magasanik, 1991). Gln3p and Nil1p are positive regulators whereas Ure2p and Dal80p are negative regulators. Intracellular levels of glutamine and glutamate regulate the activity of Gln3p and Nil1p respectively. For example, in response to increased intracellular concentration of glutamine Ure2p inactivates Gln3p (Coffman et al., 1994; 1996). Similarly, in the presence of excess intracellular glutamate Nil1p is inactivated (Stanbrough et al., 1995; Courchesne and Magasanik, 1988). The mechanism involved in Nil1p inactivation is not clear. Dal80p is a repressor of transcription which is activated in the presence of a preferred nitrogen source. Dal80p, Gln3p and Nil1p appear to compete for the same upstream regulatory element (5'-GATAA-3'). The outcome of this competition reflects the nitrogen source(s) available (Cunningham et al., 1994; Bysanti et al., 1991; Coffman et al., 1996).

Gap1p is a general amino acid permease (GAP). It selectively facilitates the transport of several amino acids into the cell. These amino acids are capable of inducing the enzymes required for their own utilization as a nitrogen source. The *GAP1* gene is highly sensitive to Gln3p-dependent regulation (Wiame et al., 1985).

A.10. Gcr1p mediated transcriptional activation

Gcr1p, also known as the glycolytic regulator is required for the efficient transcription of glycolytic genes in *S. cerevisiae* (Santangelo and Tornow, 1990; Baker, 1986; 1991;

Clifton and Fraenkel, 1981; Holland et al., 1987; Huie et al., 1992). The glycolytic enzymes constitute a major fraction of the soluble protein (30 to 60%) in yeast. Most of the highly abundant mRNA species code for glycolytic enzymes. Therefore the glycolytic enzymes are among the most highly expressed in *S. cerevisiae* (Chambers et al., 1989; 1995). Gcr1p binds to the sequence 5'-CTTCC-3' (Baker, 1991; Huie et al., 1992). Gcr1p functions with Gcr2p as a complex. Gcr1p binds the target DNA sequence and binds Gcr2p. Gcp2p can activate transcription (Uemura and Jigami, 1992).

Efficient expression of glycolytic genes also requires the binding of Rap1p to its site, 5'-RTRCACCCANNCMCC-3 (Drazinic et al., 1996; Chambers et al., 1989; 1995; Santangelo and Tornow, 1990; Bitter et al., 1991). High level glycolytic gene expression depends upon the synergistic action of Rap1p and Gcr1p (Bitter et al., 1991). Although there are a number of potential Gcr1p binding sites in the promoter regions of all glycolytic genes, it is likely that only those sites in reasonable proximity to a Rap1p binding site are functionally important (Huie et al., 1992). Recently, Scott and Baker (1993) proposed that the binding of Gcr1p is critical for the activation of glycolytic gene transcription. Other factors that bind adjacent to Gcr1p binding sites facilitate or modulate Gcr1p binding *in vivo* (Chambers et al., 1995).

A.11. Experimental Approach

The objectives of this investigation were to establish the functional role of the *GCV3* gene and to delineate the general features of its regulation. I have shown that *GCV3*, a gene identified by the systematic sequencing of chromosome one (Bussey et al., 1995), codes for

an H-protein that is essential for glycine cleavage. Studies of *GCV3* expression revealed that it is induced by glycine and repressed by the metabolic products that require C1-units for their synthesis. In addition, I have established a detailed analysis of the transcriptional regulation of *GCV3*. The results presented show that *GCV3* transcription is: (1) subject to Gln3p-dependent nitrogen regulation, (2) under Gcn4p-dependent general amino acid control, (3) under Bas1p and Bas2p dependent basal control, and (4) regulated by Gcr1p, a transcription factor important for the regulation of glycolytic genes. Further the upstream sequences that respond to glycine, Gln3p, Gcn4p and Gcr1p were localized within the *GCV3* promoter.

MATERIALS AND METHODS

B.1. Radioactive Tracer Compounds, Restriction Enzymes, and Oligonucleotides.

1[^{14}C]-glycine, 2[^{14}C]-glycine, γ-[^{32}P]-ATP and α [^{35}S]-dATP were purchased from ICN. Tetrahydrofolate was synthesized from folic acid using the previously described method (O'Dell, 1947). Restriction endonucleases, DNA modifying enzymes and ribonucleases were purchased from Bio/Can, New England Biolab, Gibco BRL, MBI Fermentas, and Boehringer Mannhein. The sequences of the oligonucleotides used in this study are presented in (Table 2) and they were purchased from Immunocorp, GIBCO BRL, BioCorp, and Sheldon Biotech.

B.2. Media and Culture Conditions.

S. cerevisiae rich medium (YPD) consisted of 2% Bacto-peptone (Difco), 1% yeast extract (Difco) and 2% dextrose (BDH). Three types of minimal media were routinely used. The first was standard SD (synthetic dextrose) consisting of 0.175% yeast nitrogen base without amino acids and ammonium sulphate (Difco), 2% dextrose, and 0.5% ammonium sulphate as the nitrogen source (Sigma). The second was the same minimal medium described above (SD) supplemented with 10 mM glycine as an inducer (SD + 10 mM glycine). The third minimal medium consisted of everything included in SD except ammonium sulphate. In place of ammonium sulphate various amino acids were separately included as the sole nitrogen source. The most commonly used alternative nitrogen source was 250 mM glycine (SD + 250 mM glycine) (Nagarajan and Storms, 1997). In addition to glycine, 0.1% of proline, glutamine, glutamate, (Courchesne and Magasanik, 1988) or 0.1%

Table 2.

Oligonucleotides used in this study

OLIGONUCLEOTIDE[a]	SEQUENCE (5'to 3')
GH1	ATGGATCCTGCTTCATGTGGAGATTCC
GH2	TCCCAAGCTTCACAACCATTACATACA
GH3	GGGACCTCGCAATTTTGACACTTCCGTAAGGAG TCATTCAGCGGCGGAGGATTCCGGTTTCTTTG
GH4	TGCATAAGCATAAGTGTACACGTTGAGTTTATTG TTTTATTTCCCCTACCGTCATTATAGAAATC
GH5	CGCATGTGTAGGGTCTTCC
GH6	GCATAAGTGTACACGTTGAG
GH7	ATACCCGGGATTCACTGCTTGACAGG
GH8	CATGGATCCAAAGCGAAATAGAATGC
GH-186	TACGCATGTGTAGGGTCTTCC
GH-165	TTCAACTTGCAAATCAATGT
GH-155	AAATCAATGTTTTATCTATC
GH-140	CTATCATTCTGTGCTATAGGG
GH-128	GCTATAGGGACCTCGCAATT
GH-111	ATTTTGACACTTCCGTAAGG
GH-96	TAAGGAGTCATTCAGCGGCG
GH-86	TTCAGCGGCGGAGTCACTTTC
GH-62	TCTTTCTTCTATATATACAGA

OLIGONUCLEOTIDE[a]	SEQUENCE (5'to 3')
GH-128A[b]	GCTATAGGGACCTCGCAATTTTtACACTTCCGTAAG GcGTaATTCAGCGGCGGcGTaACTTTCG
GH-128B[b]	GCTATAGGGACCTCGCAATTTTtACACTTCCGTAAG GAGTCATTCAGCGGCGGAGTCACTTTCG
GH-128C[b]	GCTATAGGGACCTCGCAATTTTGACACTTCCGTAA GGcGTaATTCAGCGGCGGAGTCACTTTCG
GH-128D[b]	GCTATAGGGACCTCGCAATTTTGACACTTCCGTAA GGAGTCATTCAGCGGCGGcGTaACTTTCG
GH-128E[b]	GCTATAGGGACCTCGCAATTTTGACACTTCCGTAA GGcGTaATTCAGCGGCGGcGTaACTTTCG
GH-128F[b]	GCTATAGGGACCTCGCAATTTTtACACTTCCGTAAG GAGTCATTCAGCGGCGGcGTaACTTTCG
GH-128G[b]	GCTATAGGGACCTCGCAATTTTtACACTTCCGTAAG GcGTaATTCAGCGGCGGAGTCACTTTCG
GH-156	CAAATCAATGTTTTATCTATCATTCTGTGCTATAGG GACCCGCA
GH-176	TAGGGTCTTCCTTCAACTTGCAAATCAATGTTTTAT CTATCATTC
LacZ1	GGATGTGCTGCAAGGCGATTA
LacZ2	CCCAGTCACGACGTTGTAAAACG
2μ1	GAATTCTCATGTTTGACAGC
2μ2	GTATCGTACAGTAGACGGAG
UR1	CTAGGATGAGTAGCAGCACG

[a] All the GH oligos hybridize to the *GCV3* gene; LacZ1 and LacZ2 oligos hybridize to the *E.coli LacZ* gene; 2μ1 and 2μ2 hybridize to *S.cerevisiae* 2μ circle (replication origin)

[b] Mutated GCREs are underlined and the mutated bases represented with small letters.

asparagine (Coffman et al., 1994) was also used individually in place of ammonium sulphate.

The effects of C1-end products on *GCV3-lacZ* expression were tested using SD + 10 mM glycine medium supplemented with adenine (40 μg/ml), histidine (20 μg/ml), methionine (20 μg/ml), serine (5 mM), and formate (10 mM). For sporulation studies, liquid pre-sporulation and sporulation media were prepared as described previously (Sherman, 1991). Amino acids and nucleotides were supplemented as needed minus those required for plasmid selection (Sherman, 1991; Ausubel et al., 1989). Solid media were prepared with 2% agar. All *S.cerevisiae* cultures were grown at 30°C.

The *E.coli* strains DH5α (Hanahan et al., 1983) and MC1066 (Casadaban et al., 1983) were used for plasmid manipulation, propagation and selection. *E. coli* strains were grown in either minimal medium M9 (50 mM Na_2HPO_4, 22 mM KH_2PO_4, 19 mM NH_4Cl , 17 mM NaCl, 1 mM $MgSO_4$, 0.1 mM CaCl2, O.2% glucose, 20 μg/ml thiamine-HCl, and 50 μg/ml of the required amino acids), or rich LB medium (1% N-Z amine type A, 0.5% yeast extract (Difco), 0.2% glucose). Solid media were prepared with 2% agar. *E.coli* cells were grown at 37°C. Ampicillin was included at 100 μg/ml when selecting for plasmid transformants.

In experiments performed with strain S150-2B (Table 3) the *his3* mutation was complemented with pRS313 (Sikorski and Hieter, 1989), so that histidine supplement of the medium was not required. Strains lacking general amino acid control and basal transcriptional control, (CY614, CY945, and 5014) need histidine supplement for their growth. Therefore, pLNGCV3-211 expression studies using strains CY614, CY615, CY945, CY946 (Devlin et al., 1991), and 5014 were supplemented with histidine.

Table 3.

Yeast strains used in this study

Strain	Genotype	Reference
1095-302C	*MATa ser1 cyh4 ade2 his3 leu2 ura3 his4 thr4 gal MAL2*	YGSC
DBY745	*MATα ade1-100 leu2-3 leu2-12 ura3-52 Gal+*	YGSC
3751	*MATα ade1-100 leu2-3 leu2-12 ura3-52 gcv3::URA3 Gal+*	Nagarajan and Storms, 1997
3771	Diploid obtained by crossing DBY745 and 1095-302C	Nagarajan and Storms, 1997
3773	Diploid obtained by crossing 3751 and 1095-302C	Nagarajan and Storms, 1997
3792A	*MATa leu2 ade thr4 ura3*	Nagarajan and Storms, 1997
3792B	*MATα ser1 leu2 his3 ade gcv3::URA3*	Nagarajan and Storms, 1997
3792C	*MATa ser1 leu2 his3 ade ura3*	Nagarajan and Storms, 1997
3792D	*MATα leu2 ade thr4 gcv3::URA3*	Nagarajan and Storms, 1997
4049	*MATa Δura3 leu2*	Nagarajan and Storms, 1997
4070	*MATa Δura3 leu2 ser1::URA3*	Nagarajan and Storms, 1997
4404	*MATa Δura3 leu2 gcv3::URA3*	Nagarajan and Storms, 1997
RH1385	*MATa Δura3*	Mösch et al., 1990
RH1378	*MATa Δura3 gcd2-1*	Mösch et al., 1990
RH1408	*MATa ura3-52 gcn4-103*	Mösch et al., 1990

Strain	Genotype	Reference
3634	*MATa Δura3 leu2::URA3*	Nagarajan and Storms, 1997
3640	*MATa Δura3 gcd2-1 leu2::URA3*	Nagarajan and Storms, 1997
3646	*MATa ura3-52 gcn4-103 leu2::URA3*	Nagarajan and Storms, 1997
5013	*MATα gcn4-2 bas1-2 bas2 ura3-52 leu2-3, 112*	This study
CY614	Same as 5013 but transformed with YCp50	Devlin et al., 1991
CY615	Same as 5013 but transformed with wild type *GCN4* on YCp50	Devlin et al., 1991
CY945	Same as 5013 but with BAS2 transformed with YCp50	Devlin et al., 1991
CY946	Same as 5013 but transformed with wild type *BAS1* on YCp50	Devlin et al., 1991
5014	Same as 5013 but trasformed with wild type *BAS1* on YCp50	This study
PM38	*MATα ura3-52 leu2-3, 112*	Minehart and Magasanik, 1991
PM81	*MATα ura3-52 leu2-3, 112 gln3Δ6::URA3*	Minehart and Magasanik, 1991
S150-2B	*MATa leu2-3, 112 his3 Δtrp1-289 ura3-52*	Scott and Baker, 1993
HBY4	*MATa leu2-3, 112 his3 Δtrp1-289 ura3-52 gcr1Δ::HIS3*	Scott and Baker, 1993

YGSC, Yeast Genetic Stock Centre.
3792A to D are from a single diploid.

B.3. *S. cerevisiae* Strains and Strain Constructions.

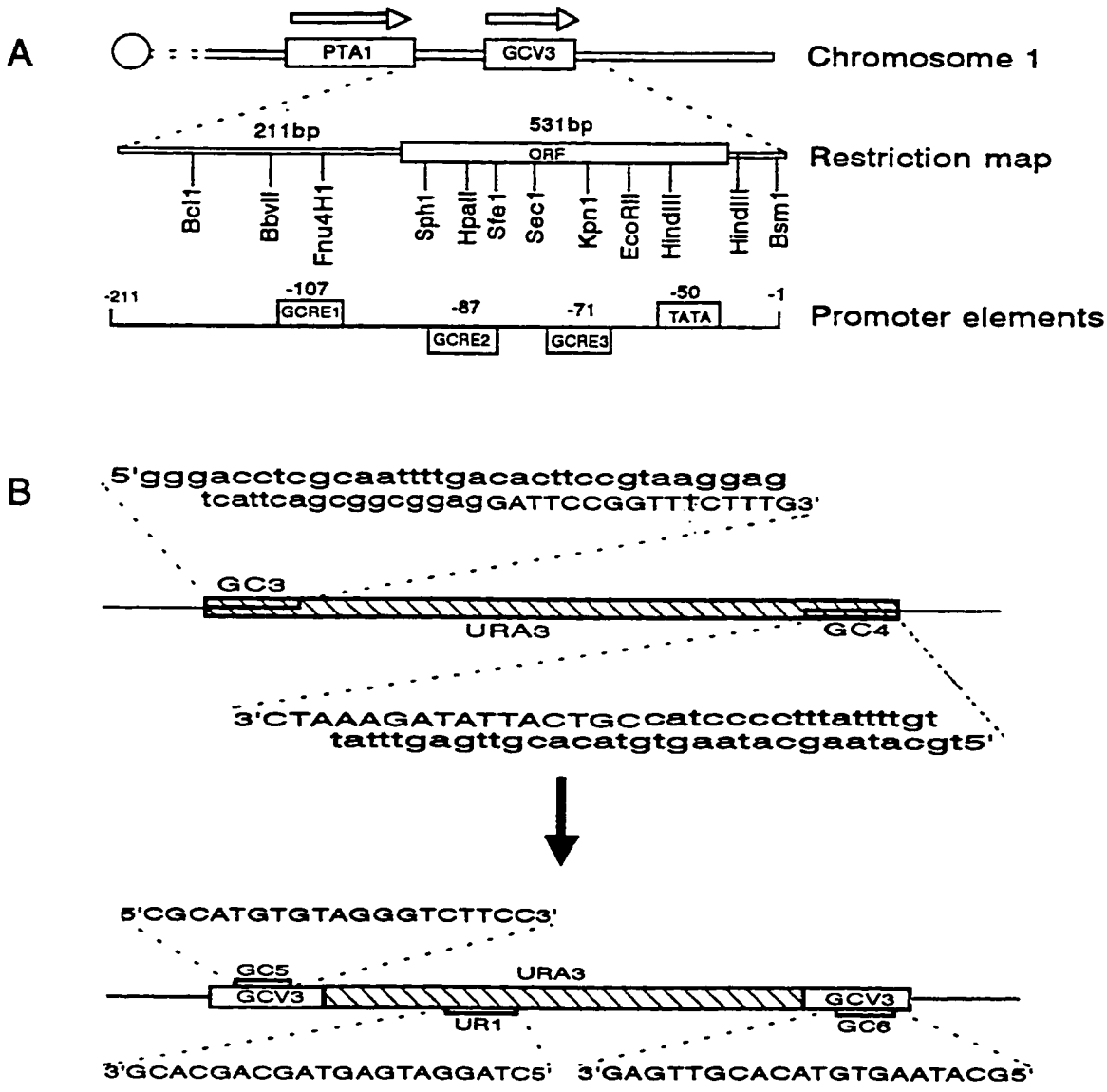
The strains used in this study are described in Table 3. The isogenic strains 3634, 3640 and 3646 are *leu2::URA3* derivatives of the strains RH1385, RH1378 and RH1408 (Mösch et al., 1990), respectively. These *ura3-52* strains were converted to *leu2::URA3* using the one step gene replacement method (Rothstein, 1983). This was accomplished by transformation with the 1.6 kbp *Bgl*II fragment from pNKY85 (Alani et al., 1987).

Strain 4049 was derived from strain 3634 by selecting for uracil auxotrophs (Boeke et al., 1984). Strain 4070, a *ser1::URA3* derivative of 4049, was constructed by gene replacement using the 2.0 kbp *Bst*EI to *Pvu*II fragment of pKM140 (Melcher et al., 1995).

The *GCV3* gene of strains DBY745 and 4049 was replaced with *URA3* to generate strains 3751 and 4404 respectively. The method used, a PCR based modification of the gene replacement method (Baudin et al., 1993), is outlined in (Fig.3). In summary, the complete *URA3* gene was amplified by PCR from S288C genomic DNA using oligonucleotides GC3 and GC4. The PCR product generated is a 1098 bp fragment extending from 182 bp upstream of the *URA3* translation start codon to 96 bp beyond the translation stop codon. In addition, the amplified fragment is flanked at both ends by 45 base pair sequences that are identical to sequences flanking the *GCV3* locus. The upstream end of the *URA3* fragment has 45 bp that are identical to the sequence from nucleotides -73 to -117 relative to the *GCV3* ORF. The 3' end of the *URA3* fragment has a 45 bp extension that is identical to the sequence from 97 to 141 base pairs downstream of the *GCV3* stop codon. Transformation of *S. cerevisiae* with this PCR product can result in the replacement of 665 bp from the *GCV3* locus with a 1098 bp fragment encoding *URA3*. Several Ura$^-$ transformants of strains

Fig. 3.

Schematic representation of the *GCV3* gene and its disruption with *URA3*. *(A)* Schematic representation showing the *GCV3* region on the left arm of *S. cerevisiae* chromosome I, a partial restriction map of the *GCV3* region and the location of potential transcription control signals in the *GCV3* promoter region. Open arrows indicate the direction of transcription of the upstream gene (*PTA1*) and of *GCV3*. GCRE1, 2 and 3 and TATA indicate potential general control response elements and basal transcription complex assembly sites. The open circle represents the centromere. (*B*) Schematic of the PCR based methods used to generate (shown above the arrow) and characterize (shown below the arrow) the *gcv3::URA3* mutants used in this study. The thick hatched line indicates the *URA3* gene. Oligonucleotides used for PCR amplification (GC3, GC4, GC5, GC6 and UR1) are depicted. For GC3 and GC4 uppercase letters represent *URA3* sequences (the 3' ends correspond to the nucleotides -166 and +96 relative to the first base of the *URA3* start and stop codons), lowercase represent *GCV3* sequences (the 3' ends correspond to nucleotides -73 and +50 relative to the start and stop codons of *GCV3*). GC5 and GC6 are identical to the coding and noncoding strands adjacent to the *GCV3* locus. The 3' nucleotides of these oligos correspond to nucleotides -184 and +90 relative to the *GCV3* ORF. The 3' end of UR1 is 21 nucleotides into the *URA3* ORF.


A
PTA1
GCV3
Chromosome 1
211bp
531bp
ORF
Restriction map
Bcl1
BbvII
Fnu4H1
Sph1
HpaII
Sfe1
Sec1
Kpn1
EcoRII
HindIII
HindIII
Bsm1
-211
-107
GCRE1
-87
GCRE2
-71
GCRE3
-50
TATA
-1
Promoter elements
B
5'gggacctcgcaattttgacacttccgtaaggag
tcattcagcggcggagGATTCCGGTTTCTTTG3'
GC3
URA3
GC4
3'CTAAAGATATTACTGCcatccccttattttgt
tatttgagttgcacatgtgaatacgaatacgt5'
5'CGCATGTGTAGGGTCTTCC3'
GC5
GCV3
URA3
GCV3
UR1
GC6
3'GCACGACGATGAGTAGGATC5'
3'GAGTTGCACATGTGAATACG5'

DBY745 and 4049 were screened using the PCR method depicted in (Fig. 3) to identify Ura$^-$ transformants that had the desired gene replacement. Strain 5014, is a derivative of CY614, in which the YCp50 plasmid has been eliminated by repeatedly growing on a rich medium, followed by retransforming with the plasmid pCP291 bearing BAS1, which was originally isolated from CY946 (Devlin et al., (1991).

B.4. *S. cerevisiae* and *E. coli* Transformation and Genetic Methods.

Standard *S.cerevisiae* genetic techniques such as mating, isolation of diploids, sporulation, tetrad analysis and complementation were performed as described previously (Sherman, 1991a; Ausubel et al., 1989). Transformation was performed according to the fast colony procedure (Geitz et al., 1992). *E.coli* transformation was routinely carried out by electroporation using a Bio-Rad gene pulser and the method recommended by the supplier.

B.5. DNA isolation and manipulation.

S. cerevisiae genomic DNA was isolated essentially as described (Ausubel et al., 1989). Methods for the manipulation of DNA were performed as described by Sambrook et al., (1989). DNA sequencing was performed using the USB Sequenase kit and the method supplied by the manufacturer.

B.6. PCR amplifications.

PCR amplifications (Saiki et al., 1988) were performed using a Hybaid thermal

reactor. For PCR based screening using *S. cerevisiae* cells as the source of template the following method was used. About one-half of a large colony was suspended in 0.5 ml water. Cells ($5x10^4$ to $7.5x10^4$ depending upon the strain) were transferred to a microfuge tube and the volume adjusted to 60 µl with water. The cell suspensions were boiled for 5 min, immediately frozen in liquid nitrogen then boiled again for another 5 min. Following vigorous vortexing for 30 seconds the cell debris was removed by centrifuging at 2200 g for 2 min. This solution was used as the template source in a 100 µl PCR reaction.

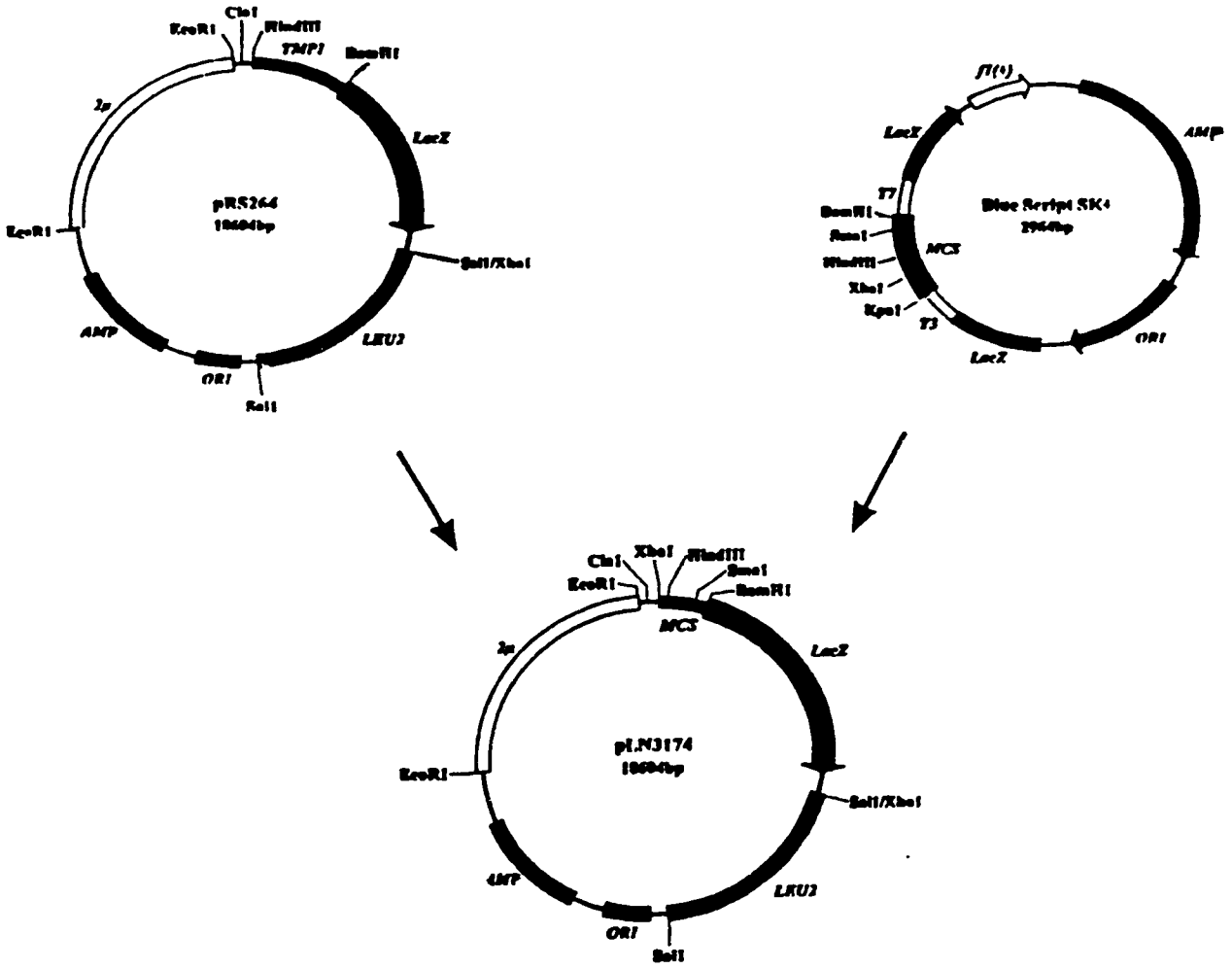
B.7. Plasmids and Plasmid Constructions.

The Bluescript SK+ (Short et al., 1988) multiple cloning site (*Hin*dIII to *Bam*H1 fragment) was cloned into the backbone of the multicopy shuttle vector pRS264 (Ord et al., 1988), prepared by *Hin*dIII and *Bam*HI digestion. This replaces the *Hin*dIII to *Bam*HI fragment of pRS264. The resulting plasmid was designated as pLN3174 (Fig. 4).

The first *GCV3-lacZ* fusion plasmid, pLNGCV3-211, was constructed as follows (Fig. 4). The *GCV3* promoter DNA was generated by PCR amplifying pLNGCV3 DNA using primers GH1 and GH2. The PCR product was end filled with Klenow polymerase (from Bio/Can and Boehringer Mannheim) followed by digestion with *Bam*HI and cloning into the backbone of pRS264 (Ord et al., 1988) which had been generated by *Sma*I and *Bam*HI digestion. This construct places *lacZ* expression under the control of the *GCV3* promoter. The second construct, pLNGCV3, is identical to pLNGCV3-211 except that the *GCV3-lacZ* portion has been replaced by the complete *GCV3* gene. *GCV3* DNA, prepared

Fig. 4.

Construction of pLN3174 and various *GCV3-lacZ* fusion plasmids. Yeast and *E.coli* DNA are schematically represented by boxes. *2μ, S. cerevisiae* 2μ plasmid sequences; *LacZ,* coding region of *E. coli lacZ* gene; *LEU2, S. cerevisiae LEU2* gene; *ori,* pBR322 origin of replication; *AMP,* pBR322 ß-lactamase gene. The plasmid names and the sizes in base pairs are denoted in the center of the circle. For constructing pRS3174, the Bluescript SK+ multiple cloning site (*Hin*dIII to *Bam*HI fragment) was cloned into the backbone of pRS264 prepared by *Hin*dIII and *Bam*HI. Various portions of the *GCV3* truncated and mutant promoter DNAs prepared by PCR amplification were cloned onto the backbone of pRS3174, which was prepared by sequential digestion with *Sma*I and *Bam*HI, followed by phosphatase treatment.


ClaI
EcoRI
HindIII
TMP1
BamHI
2μ
LacZ
pRS264
EcoRI
SalI/XbaI
AMP
LEU2
ORI
SalI
f1(+)
LacZ
AMP
T7
BamHI
SmaI
MCS
HindIII
XbaI
KpnI
T3
LacZ
ORI
Blue Script SK+
XbaI
HindIII
SmaI
BamHI
ClaI
EcoRI
MCS
LacZ
2μ
pLN3174
EcoRI
SalI/XbaI
AMP
LEU2
ORI
SalI

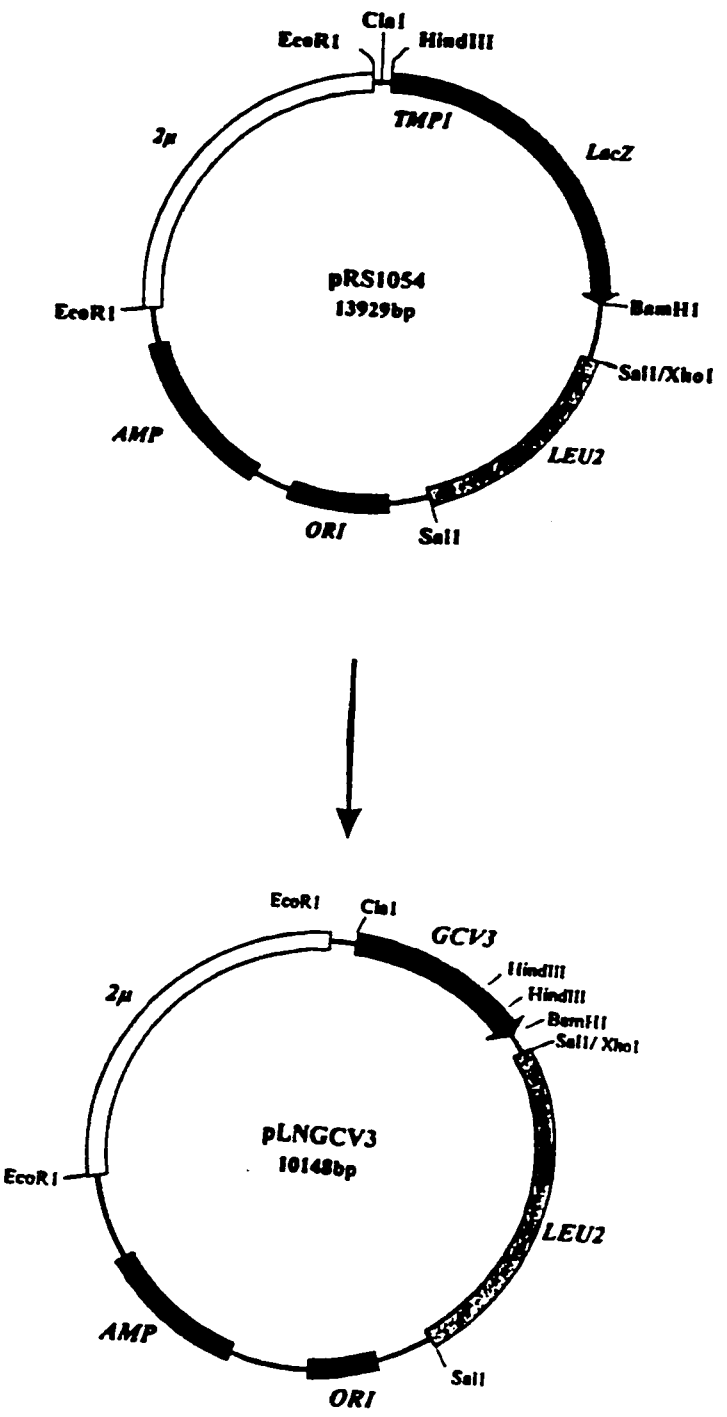
by PCR amplification from S288C genomic DNA using oligos GH7 and GH8, was digested with *Bam*HI and *Sma*I and cloned into the backbone of multicopy shuttle vector pRS1054 (Fig. 5).

B.8. *GCV3* promoter deletion analysis.

The ten *GCV3-lacZ* fusion plasmids depicted in Figure 6 were constructed as follows. The backbone of pLN3174 was prepared by sequential digestion with *Sma*I then *Bam*HI. The various *GCV3* promoter regions were prepared by PCR amplification using pLNGCV3, (Nagarajan and Storms, 1997) a yeast shuttle vector harbouring the entire *GCV3* gene, as the template DNA. The ten different *GCV3* promoter fragments used for these constructs were all generated using oligonucleotide GH1 (5'-ATGGATCCTGCTTCATGTGGAGATTCC-3') as the downstream oligo (Table 2). Proceeding from the 5' end this oligo, contains two filler nucleotides, a *Bam*HI site from nucleotides 3 to 8 and a 19 base stretch that is identical to nucleotides +8 through -11 (relative to the start codon) of the noncoding strand of *GCV3*. In addition to primer GH1 each fragment was generated using a different upstream oligo that was also 20 nucleotides long and identical to various portions of the *GCV3* coding strand (oligos GH-211 through GH-62, Table 2). These 10 oligos begin at their 5' ends with nucleotide, -211, -186, -165, -155, -140, -128, -111, -96, -86 and -62 respectively. The 10 PCR products, after end-filling with Klenow polymerase followed by digestion with *Bam*HI, were ligated into the previously prepared pLN3174 backbone. The 5' end of construct pLNGCV3-211 begins just after the stop codon of *PTA1* (O'Connor and Peebles, 1992) and

Fig. 5.

Construction of pLNGCV3. PCR amplified complete *GCV3* DNA was digested with *Sma*I and *Bam*HI and cloned onto the backbone of pRS1054 which was prepared by digestion with *Hin*dIII, end filling with Klenow polymerase, digestion with *Bam*HI and finally phosphatase treatment. *2μ, S. cerevisiae 2μ* plasmid sequences; *LacZ,* coding region of *E. coli lacZ* gene; *LEU2, S. cerevisiae LEU2* gene; *ori,* pBR322 origin of replication; *AMP,* pBR322 ß-lactamase gene.


ClaI
EcoR1
HindIII
TMP1
2μ
LacZ
pRS1054
13929bp
EcoR1
BamH1
SalI/XhoI
AMP
LEU2
ORI
SalI
EcoR1
ClaI
GCV3
HindIII
HindIII
BamHI
SalI/XhoI
2μ
pLNGCV3
10148bp
EcoR1
LEU2
AMP
SalI
ORI

this is the longest construct.

Restriction endonucleases *Hin*dIII, *Bam*HI, and *Cla*I were routinely used for screening and identification of potential recombinant plasmids (Fig. 6, Fig. 7, and Fig. 8). Subsequently, the entire *GCV3* promoter portions were sequenced to identify three independent isolates of each one of the desired plasmids (Fig. 6, Fig. 7, and Fig. 8). LacZ1, LacZ2, 2μ1, and 2μ2 oligos (Table 2) were routinely used for DNA sequencing.

B.9. GCRE mutant plasmid construction .

pLNGCV3-128A, pLNGCV3-128B, pLNGCV3-128C, and pLNGCV3-128D constructs are identical to pLNGCV3-128 (Fig. 7) except for mutations in GCREs. These plasmids were constructed by PCR amplifying pLNGCV3-128 using GH-128A, GH-128B, GH-128C, and GH-128D and GH1 oligos (Table 2). The 5' ends of oligos GH-128A to GH-128D start at -128 bp relative to the start codon of the *GCV3* gene. The PCR products were treated with Klenow polymerase followed by digestion with *Bam*HI and cloned into the backbone of pLN3174 prepared as described in section B.7. The resulting construct pLNGCV3-128A contains mutations in all three GCREs (Fig. 7). The remaining constructs (pLNGCV3-128B, pLNGCV3-128C, and pLNGCV3-128D) contain mutations in GCRE-107, GCRE-87, and GCRE-71 respectively (Fig. 7).

pLNGCV3-gcre-156 and pLNGCV3-gcre-176 plasmids were constructed by PCR amplifying pLNGCV3-128A using primers GH-156, GH-176, and GH1 (Table 2). The 5' ends of oligos GH-156 and GH-176, start at -156 bp and -176 bp respectively (relatively to the first base pair of the *GCV3* start codon). These PCR products were end filled with

Fig. 6.

***GCV3-lacZ*-fusion constructs** - All *GCV3-lacZ* fusion plasmids used in this study are identical except for the *GCV3* portion that they harbour in the pRS3174 backbone. The *GCV3* portion of the longest construct, pLNGCV3-211, starts immediately after the last bp of the stop codon of the upstream gene, *PTA1* (O'Connor, 1992). The numbers with arrows indicate the position of the 5'- ends of the *GCV3* portions relative to the *GCV3* start codon (+1 for A). All the fusion genes have *GCV3* portions that extend for varying distances into the upstream region of *GCV3* from nucleotide +8 at their 3'- ends. The plasmid designations are depicted on the left.

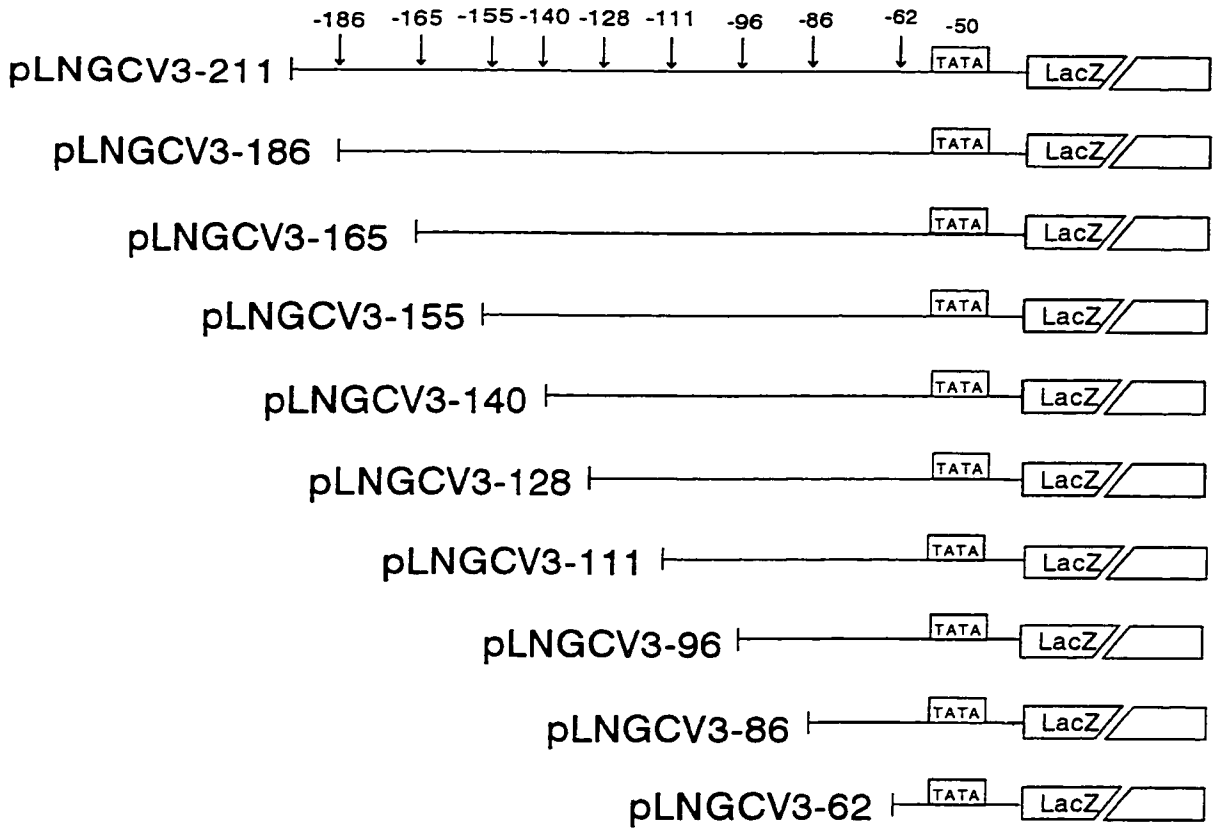

-186
-165
-155
-140
-128
-111
-96
-86
-62
-50
pLNGCV3-211
pLNGCV3-186
pLNGCV3-165
pLNGCV3-155
pLNGCV3-140
pLNGCV3-128
pLNGCV3-111
pLNGCV3-96
pLNGCV3-86
pLNGCV3-62
TATA
LacZ

Fig. 7.

Plasmids constructed for *in vivo* GCRE analysis.- All these constructs are identical except for their GCRE mutations. GCRE elements are represented as boxes. The mutant GCREs are indicated with crossed boxes. The arrows indicate the orientation of the consensus sequence. The nucleotide position numbers are marked relative to the first base pair of the start codon. Plasmid designation is depicted on the left. All these constructs harbour the glycine response element (UAS_{Gly}) upstream of GCREs.

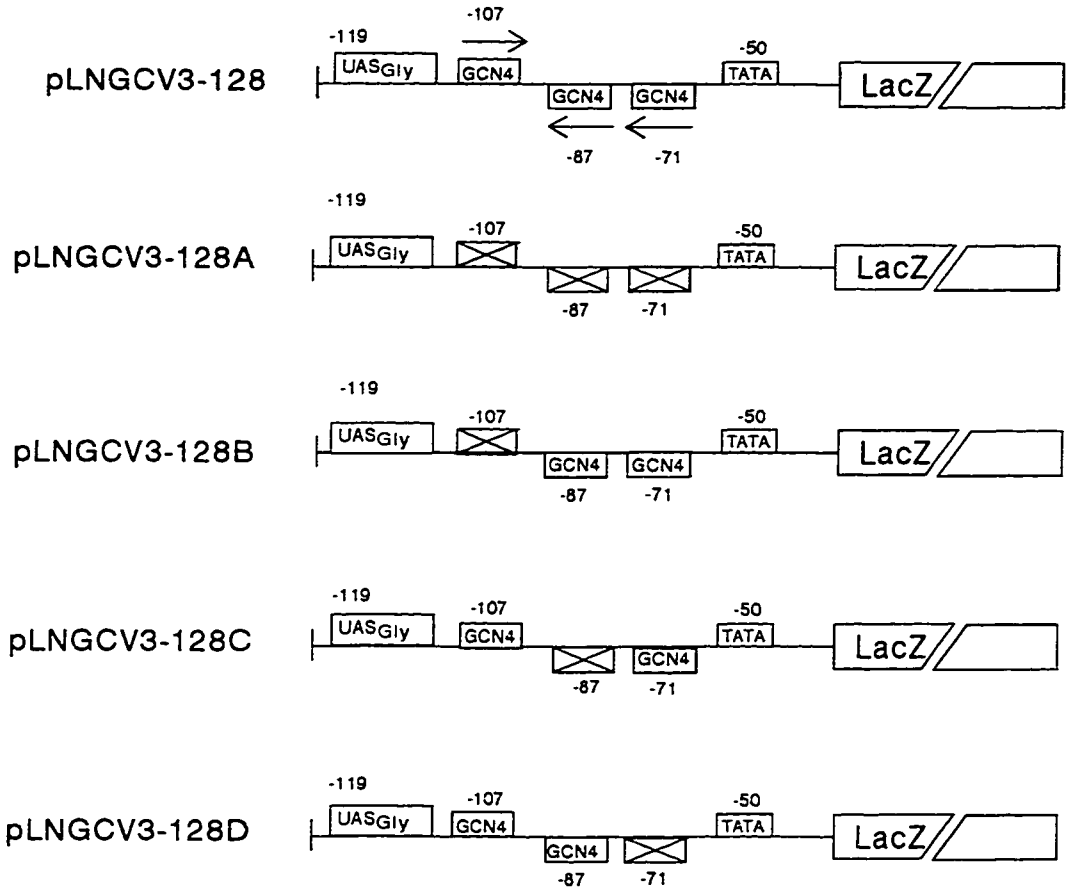
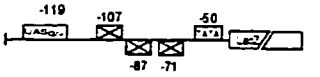
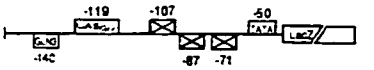
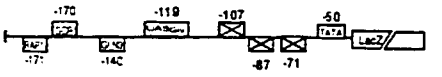

pLNGCV3-128
-119
-107
-50
UASGly
GCN4
GCN4
GCN4
TATA
LacZ
-87
-71
pLNGCV3-128A
-119
-107
-50
UASGly
TATA
LacZ
-87
-71
pLNGCV3-128B
-119
-107
-50
UASGly
TATA
LacZ
GCN4
GCN4
-87
-71
pLNGCV3-128C
-119
-107
-50
UASGly
GCN4
TATA
LacZ
GCN4
-87
-71
pLNGCV3-128D
-119
-107
-50
UASGly
GCN4
TATA
LacZ
GCN4
-87
-71

Fig. 8.

Nitrogen and glycolytic regulatory regions can function independently. The boxes represent cis-acting elements. The nucleotide positions of these cis-acting elements are marked relative to the first base of the *GCV3* start codon. Crossed boxes represent mutated GCREs. The plasmid designations are depicted on the left. All these constructs are identical with regard to their GCRE mutations. The constructs pLNGCV3-gcre-156 and pLNGCV3-gcre-176 are derived from pLNGCV3-128A. pLNGCV3-gcre-156 contains additional information for nitrogen regulation (GLN3). Besides nitrogen regulation the pLNGCV3-gcre-176 construct also contains upstream information for glycolytic regulation (GCR1 and RAP1). On the right the ß-gal activity of these constructs in a wild-type strain (3634) measured under three different physiological conditions is presented in Miller units.

	ß-Gal activity (Miller Units)		
	SD	SD + 10 mM Glycine	SD + 250 mM Glycine
pLNGCV3-128A 	0	0	0
pLNGCV3-gcre-156 	0	0.15	0.76
pLNGCV3-gcre-176 	0.15	0.28	0.95

Klenow polymerase followed by digestion with *Bam*HI and cloned into the backbone of pLN3174 prepared as described earlier in section B.7. Similar to pLNGCV3-128A, the resulting constructs contain mutations in all three GCREs but the 5' end of pLNGCV3-gcre-156 extends from -156 bp which includes a Gln3p binding site located at -140 bp (Fig. 8). Similarly the 5' end of the pLNGCV3-gcre-176 extends from -176 bp, and includes two upstream elements a Gcr1p binding site at position -170 of the coding strand and a Rap1p binding site at position -171 of the non-coding strand (Fig. 8).

B.10. ß-galactosidase assays.

ß-galactosidase assays were performed essentially as described previously (Storms et al., 1984) with a few modifications. At least 4 ml of mid log phase (OD_{600} 0.2 to 0.25) cultures were harvested and immediately frozen in liquid nitrogen for 15 to 20 min. Then the cells were thawed at room temperature for 5 min, followed by freezing in liquid nitrogen for another 10 minutes. To obtain consistent results I found that it was necessary to subject the cells to at least two rounds of freezing and thawing. Immediately after thawing the second time the cells were resuspended in 1 ml Z-buffer (Miller, 1972) containing 3 % Brij-35 (BDH diagnostics). The cell suspension was allowed to equilibrate in this buffer for 5 min at 28 °C. Then 0.2 ml O-nitro-phenyl-ß-D-galactopyranoside (ONPG) 4 mg/ml (Miller, 1972) was added, the tubes were vortexed gently, and the incubation continued until a faint yellow colour was detectable. The assays were terminated by the addition of 0.5 ml of 1 M sodium bicarbonate. The samples were chilled on ice until the cell debris settled. The levels of O-nitro-phenyl production were measured spectrophotometically at OD_{420}. The ß-

galactosidase activities were calculated as Miller units (Miller, 1972). All ß-gal activity values are averages obtained from at least three independent experiments. The values did not vary by more than ± 5%.

B.11. Cell extract preparation for GCS assays.

Extracts for glycine cleavage activity assays were prepared as follows: 100 ml cultures, grown to mid log phase (OD_{600} = 0.25) in SD + 10 mM glycine, were harvested by centrifugation and washed twice with cold sterile water. After freezing in liquid nitrogen the cell pellets were stored at -70°C until needed. To prepare extracts, the cells were suspended in 1 ml of cell extraction buffer (25 mM potassium phosphate pH 7.4, 2 mM EDTA, 10 mM benzamidine, 10 mM ß-mercaptoethanol and 1 mM PMSF). One-half volume of glass beads (acid-washed and baked) was added to the cell extraction mixture, and the cells were disrupted using a Silamat agitator (five 30 sec bursts separated by 1 min on ice). After cell debris was removed by centrifugation (5000 rpm for 5 min using a Beckman JA17 rotor) the supernatant was carefully recovered and subjected to three cycles of freezing in liquid nitrogen and thawing to break down the mitochondrial membranes as described in Bourguignon et al., (1988). The protein concentration was determined (Bradford, 1976) using bovine serum albumin as the standard.

B.12. Glycine cleavage activity assays.

Glycine cleavage assays were performed essentially as described (Meedel and Pizer, 1974) with a few modifications. The 0.5 ml assay mixture consisted of 20 mM potassium

phosphate buffer pH 7.4., 2 mM dithiothreitol, 1 mM tetrahydrofolate, 2.5 mM pyridoxal phosphate, 2 mM NAD^+ and enzyme (2 mg crude cell extract). The glycine cleavage reaction was initiated by the addition of 66 µl of [2-^{14}C]-glycine (1 µCi/µmol). After incubation at 30°C for 45 min, the reaction was terminated by the sequential addition of 0.3 ml of 1 M sodium acetate pH 4.5, 0.2 ml of 0.1 M formaldehyde and 0.3 ml of 0.4 M dimedon in 50% ethanol. The stopped reaction was placed at 65°C for 5 min and then placed on ice for at least 5 min. Next 5 ml of toluene was added, followed by vigorous vortexing and then centrifugation in a clinical centrifuge for 5 min. A portion (3 ml) of toluene layer (upper phase) containing the glycine cleavage system product $^{14}CH_2$-THF was transferred to a scintillation vial containing 5 ml of Ecolite (ICN) and radioactivity was measured using an LKB, RackBeta, liquid scintillation spectrophotometer.

B.13. Gcn4p and Bas2p preparation.

Gcn4p and Bas2p were produced in AR68 (*E.coli*) using the pAB100 and pME1135 vectors, respectively. These two vectors contain the *GCN4* and *BAS2* genes, respectively, under the control of an inducible lambda P_L promoter. The AR68 strain contains a temperature sensitive c1857 lambda repressor and a mutation in the *htpR* gene. The AR68 strains harbouring these vectors were grown at 30°C, and the proteins were overproduced by raising the temperature to 43.5°C for 70 min with gentle agitation. From the crude cell extract the Gcn4p was partially purified on a phosphocellulose column as previously described by Arndt and Fink (1986). Similarly, the Bas2p was also purified as previously described by Mösch et al. (1990).

B.14. DNA fragments used for gel retardation assays.

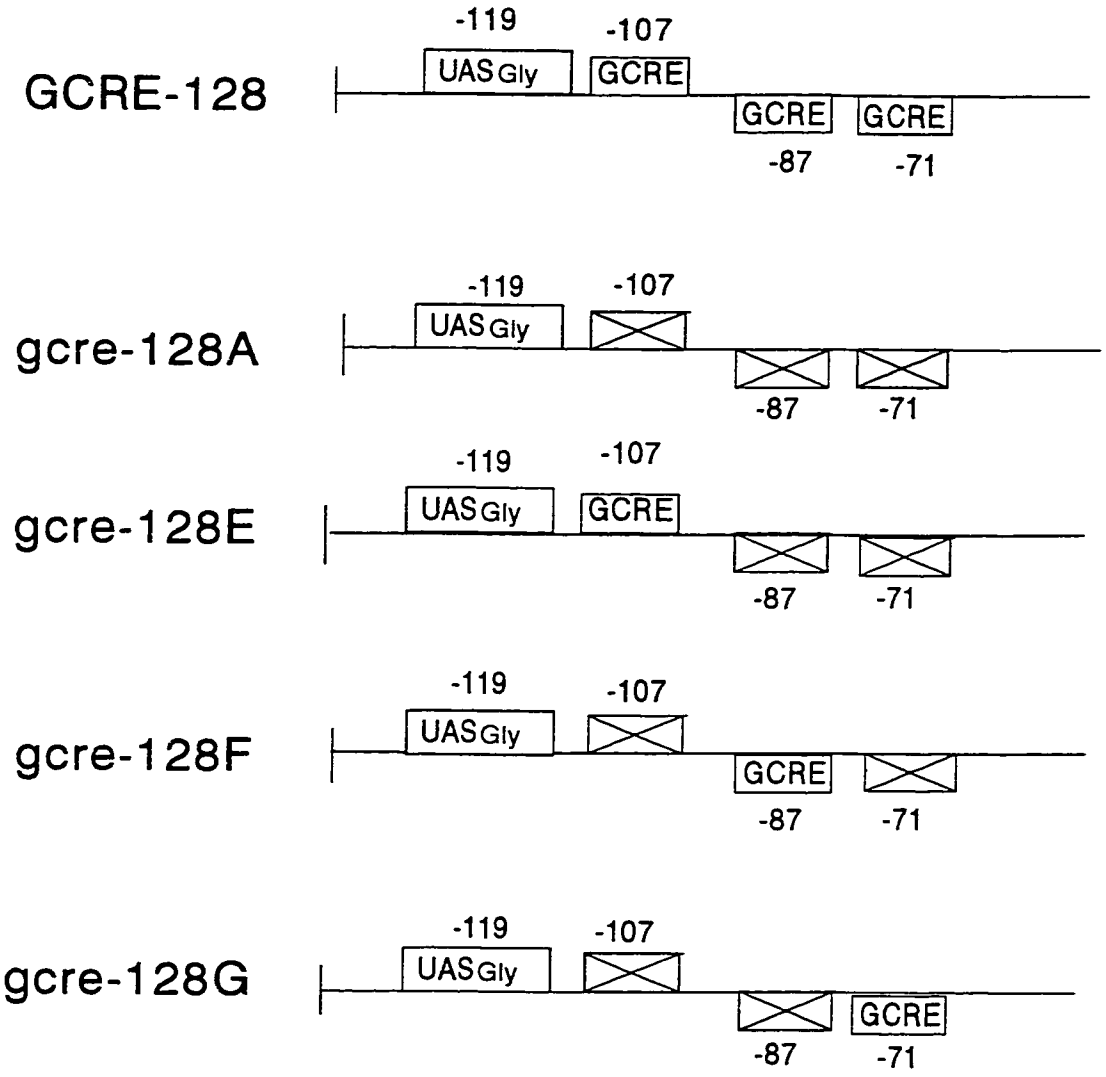
The same PCR fragments used for constructing pLNGCV3-128 (GCRE wild-type) and pLNGCV3-128A (GCRE mutant) were also used in the gel retardation experiments. In addition to these two PCR products, gcre-128E, gcre-128F, and gcre-128G (Fig. 9) PCR fragments were also generated using GH-128E, GH-128F, GH-128G, and GH1 oligos (Table 2) using pLNGCV3-128 as a template. These three PCR fragments were used to test the Gcn4p and Bas2p binding ability of each GCRE independently, along with a combination of GCRE mutations (Fig. 9). For example, to test the protein binding ability of GCRE located at -107 bp, the other two GCREs located at -87 bp and -71 bp were mutated as shown in oligo GH-128E (Table 2). To test the protein binding ability of GCRE located at -87 bp, the GCREs located at -107 bp and -71 bp were mutated as shown in oligo GH-128F (Table 2). Similarly, to test the DNA binding ability of GCRE located at -71 bp, the GCREs located at -107 bp and -87 bp were mutated as shown in oligo GH-128G (Table 2).

B.15. Gel retardation assays.

The gel retardation assays were performed as described earlier by Paravicini et al., (1989) with few modifications. PCR amplified DNA fragments (5 picomoles) were routinely end labelled with γ-[^{32}P] ATP as described (Sambrook et al., 1989). Unbound γ-[^{32}P] ATP was removed by passing the labelled oligos over G-50 columns. The DNA-protein binding reaction mixtures contained 20 mM HEPES buffer (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid; pH 7.5), 0.15 mM EDTA, 1 mM DTT, 5 mM $MgCl_2$, 20 mM KCl, 8 % glycerol, 150 µg of poly (dI-dC) per ml, and 8 ng of γ-[^{32}P] ATP labelled DNA-probe in a

Fig. 9.

PCR products used *in vitro* GCRE analysis. The boxes represent cis-acting elements. The numbers indicate the position of these elements relative to the first base of the *GCV3* start codon. All of these PCR products are identical except for their mutations in GCREs. The mutated GCREs are indicated with crossed boxes. GCRE-128 (wild-type) and gcre-128A (GCRE mutant) are the same PCR products used for constructing pLNGCV3-128 and pLNGCV3-128A respectively (see Fig. 7). gcre-128E, gcre-128F, and gcre-128G contain two mutant and wild type GCRE element. These were prepared for DNA-protein binding studies.


GCRE-128
-119
-107
UAS Gly
GCRE
GCRE
GCRE
-87
-71
gcre-128A
-119
-107
UAS Gly
-87
-71
gcre-128E
-119
-107
UAS Gly
GCRE
-87
-71
gcre-128F
-119
-107
UAS Gly
GCRE
-87
-71
gcre-128G
-119
-107
UAS Gly
GCRE
-87
-71

20 µl volume. The binding reaction was initiated by the addition of a DNA-binding protein (3.2 µg of Gcn4p or 3 µg of Bas2p) and the reaction mixtures were incubated at 25°C for 25 minutes. Then 2µl of 0.5% Xylene Cyanol FF was added and the reaction mixtures were separated on a 6% native polyacrylamide gel at 4°C. The electrophoresis apparatus was modified to allow for circulation of the TE buffer (10 mM Tris, pH 7.4 at 21°C, 1 mM EDTA) during electrophoresis. Before loading the samples, the gels were pre-run until the conductivity was invariant with time. To prevent raising the gel temperature due to low conductivity, 16 volt cm^{-1} was maintained. The gels were run until Xylene Cyanol FF migrated one third of the total length of the gel. Afterwards, the gels were fixed for 10 min in a 10 % methanol, 10 % acetic acid solution. Then the gel was transferred to Whatman 3MM paper, and dried for 45 min using a Bio-Rad gel drier. Autoradiograms were generated by exposing Fuji X-ray film 035010 with Du Pont Cronex intensifier screens to the gel for 3 to 5 hr at -70°C.

B.16. Gcn4p and Bas1p binding optimization

The PCR product used for constructing pLNGCV3-128 was used for Gcn4p and Bas2p binding optimization experiments (Fig. 10 and Fig. 11). The previously described (section B.15) DNA-protein binding reaction mixtures were supplemented with differing amounts of *E.coli* produced Gcn4p. Each of the following lanes contains 2-fold more Gcn4p than the preceding lane (lane 1, 0.1 µg through lane 6, 3.2 µg) (Fig. 10). The retardation pattern indicated that optimum retardation required at least 3.2 µg Gcn4p per reaction (Fig. 10, lane 6). A similar experiment was performed with Bas2p (Fig. 11) and the reaction

Fig. 10.

Optimization of GCN4p binding to GCRE-128. Lanes 1 to 7 contain, 5 ng of γ-[^{32}P] ATP labelled GCRE-128 probe. Varying amounts of Gcn4p were included for determining the level of Gcn4p required to cause optimal retardation. The reaction mixtures were supplemented with increasing amounts of Gcn4p, starting from lane 1, 0.1 μg; lane 2, 0.2 μg; lane 3, 0.4 μg; lane 4, 0.8 μg; lane 5, 1.6 μg; and lane 6, 3.2 μg. Lane 7 contains the radiolabelled GCRE-128 probe alone. Lane 6 containing 3.2 μg Gcn4p was considered to be the optimal level of Gcn4p required for causing optimal retardation.

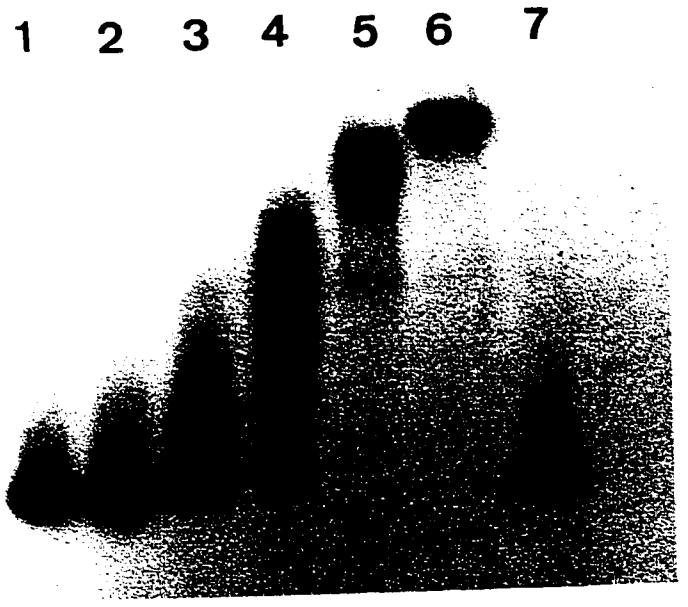
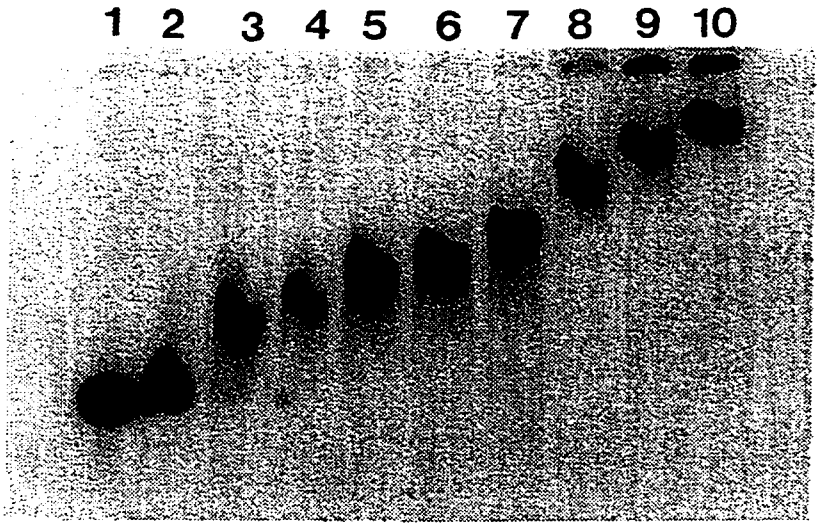

1 2 3 4 5 6 7

Fig. 11.

Bas2p Binding Optimization. Radiolabelled GCRE-128 PCR product shown in Fig. 9 was used to standardize the Bas2p binding. All the lanes contained 5 ng of γ-^{32}P labelled GCRE-128 probe. Lane 1 was not supplemented with any protein. Starting from lane 2 the reaction mixtures contained an increasing amount of Bas2p: lane 2, 0.5 μg; lane 3, 1 μg; lane 4, 1.5 μg; lane 5, 2 μg; lane 6, 2.5 μg; lane 7, 3 μg; lane 8, 3.5 μg; lane 9, 4 μg; and lane 10, 4.5 μg. Lane 7 containing 3 μg Bas2p was considered to be the optimal level of Bas2p required to cause optimal retardation.


1 2 3 4 5 6 7 8 9 10

mixtures starting from lane 2 contained 0.05 μg more of Bas2p than the preceding lane (lane 2, 0.5 μg through lane 10, 4.5 μg). Based on the retardation pattern it was found that 3.0 μg (Fig. 11, lane 7) of Bas2p was necessary for optimal retardation. Therefore, further experiments involving Gcn4p were performed with 3.2 μg of protein and similarly experiments involving Bas2p were performed with 3.0 μg of protein.

B.17. Growth rates.

Growth rates of individual yeast strains were determined by inoculating with a portion of fresh overnight cultures into 100 ml of liquid SD medium (initial ODs at 600 nm were <0.02) with appropriate supplements. The cultures were constantly shaken (350 rpm) at 30°C. Growth was monitored at 2 to 4 hr intervals by measuring turbidity at 600 nm to a maximum of 24 hr. The *gcv3* null mutant (4404) and mutants lacking basal transcriptional control (CY614, CY615, CY945, and 5014) grew slowly in the presence of 10 mM glycine. Therefore, for these strains growth was followed to a maximum of 44 hr. Growth on plates was followed for at least 5 days at 30°C.

B.18. Glycine transport assays.

Glycine transport assays were performed as previously described (Sinclair and Daws, 1995) with a few modifications. Yeast strains 3634 (wt), 4404 (*gcv3*), CY614 (*bas1*, *bas2*, and *gcn4*), CY615 (*bas1*, and *bas2*), CY945 (*gcn4*, and *bas1*), and CY946 (*gcn4*) were employed in this study. These strains were grown in SD media to mid log phase (OD_{600} = 0.2) at 30°C. An aliquot (5 ml) of each culture was harvested and the cells were washed

twice with medium A, (0.1% glucose, 0.17% yeast nitrogen base without amino acids and ammonium sulphate, and 10 mM potassium phosphate buffer, pH 7.4). The cells were resuspended in 1 ml of the same medium and starved for nitrogen and partially for carbon by incubating with gentle agitation for 2 hr at 30°C. Then the cells were harvested and resuspended in medium A to a cell density of $1x10^7$ cells/ml. The cell density was adjusted using a haemocytometer. While adjusting the cell density the cell suspensions were kept on ice. To initiate the assay, duplicates of 100 µl of cell suspension were added to 100 µl of 2 mM glycine (containing 0.5 µCi of [1-^{14}C] glycine) in medium A. The reaction mixtures were overlaid with 100 µl oil mixture (containing eight parts di-n-butylphthalate plus two parts of di-iso-octyl phthalate) and incubated at 30°C. The assays were stopped every 10 min by pelleting the cells into the oil mixture by centrifugation in a microcentrifuge for 30 sec. After carefully removing the top oil layer the cell pellets were immediately frozen in liquid nitrogen. The assays were allowed to continue to a maximum of 120 min. The cell pellets were washed three times with medium A (without glycine) and intracellular glycine was extracted with 400 µl of 1.5% Triton X100 for 12 hr and 400 µl of 5% cold perchloric acid for 1 hr. After centrifugation in a microcentrifuge for 15 min, 800 µl of supernatant was added to 5 ml scintillant Ecolite (ICN) and the radioactivity was measured using a LKB, RackBeta, liquid scintillation counter.

B.19. Determination of intracellular glycine levels or *in vivo* glycine degradation assay.

SD grown mid log phase (OD_{600} = 0.2) yeast cultures (5 ml) were harvested, washed once in medium A (described in section B.18) and resuspended in the same medium. The

cells were starved for nitrogen and partially for carbon in medium A for 2 hr at 30°C. Then the cells were harvested and resuspended in the same medium adjusting the cell density to $1x10^7$ cells/ml using a haemocytometer. To initiate the glycine uptake, two identical duplicate sets of 100 µl aliquots of the above cell suspensions were added to 100 µl of 2 mM glycine (containing 0.5 µCi of [1-^{14}C] glycine) in medium A. Then the reaction mixtures were overlaid with 100 µl of oil mixture (containing eight parts of di-n-butylphthalate plus two parts of di-iso-octyl phthalate). The reaction mixtures were incubated at 30°C with gentle agitation. The glycine uptake was stopped after 30 min by pelleting the cells into the oil layer. The aqueous and oil layers were carefully removed and the cell pellets were washed with medium B (medium A supplemented with 0.5 mM pyridoxal phosphate and 0.5 mM NAD for facilitating *in vivo* glycine cleavage reaction). One set of identical samples was subjected to intracellular glycine extraction as described earlier. Another set of samples continued incubating with 200 µl of medium B at 30°C for another 8 hr with gentle agitation. At the end of incubation the cells were harvested and the intracellular glycine was extracted as described earlier. The supernatant was examined to ensure that the intracellular glycine was not leaching into the medium during the prolonged incubation in medium B.

RESULTS

C.1. *GCV3* Codes for a Glycine Cleavage H-protein and is Essential for Glycine Cleavage Activity.

A BlastX search (Altschul et al., 1990) revealed that the putative protein encoded by *GCV3* had significant similarity to the glycine cleavage system H-protein from organisms ranging from bacteria to mammals (Fig. 12). The degree of identity observed varied from a high of 38% (*E.coli*) to a low of 31% (*Pisum sativum)*. In addition to its overall homology to other H-proteins, Gcv3p also harbours the highly conserved lipoate attachment site signature surrounding LYS^{109} found in all H-proteins characterized to date (Fig. 12). The presence of a lipoic acid signature suggests that Gcv3p is mitochondrial since the enzymatic attachment of lipoic acid occurs in the mitochondrial compartment (Fujiwara et al., 1991). Consistent with this possibility the N-terminal 54 residues of *GCV3* encode a putative mitochondrial targeting sequence (Fig. 12). The lipoic acid binding signature, significant similarity to other H-proteins and a mitochondrial targeting signal strongly suggested that *GCV3* encoded a glycine cleavage system H-protein.

To determine whether the *GCV3* encoded H-protein was necessary for glycine cleavage system activity, I performed glycine cleavage assays on cell extracts prepared from a wild-type strain (3634) and an isogenic derivative (4404) harbouring a *gcv3::URA3* null mutation. The assay was also performed on extracts prepared from strain 4070, a *ser1::URA3* derivative of 3634, to determine whether expression increased when cells could not utilize serine to meet their C1-unit requirements. The amount of activity present in strain 3634 was very similar to that present in extracts prepared from *E. coli*. Further, no glycine

Fig. 12.

Gcv3p aligned with other known glycine cleavage system H-proteins. The alignment was generated using the Clustal program (Higgins and Sharp, 1988). Asterisks represent residues that are identical in all 8 H-proteins. A period indicates residues within the alignment where only conserved amino acid replacements have occurred. Species designations are given on the right side of the alignment. Also listed (right of the last set of rows) is the percent identity with the *S. cerevisiae* H- protein. Underlined residues represent a putative mitochondrial targeting signal. The solid triangle flanked by two vertical arrows indicates lysine-109, the lipoamide attachment site and the conserved surrounding region.

```
MKHTTSTMLRTTRLWTTRMPAVSKLFLRNSSGNALNKNKLPFLYSSQGPQ 50  S.cerevisiae
MALRVVRSVRALLCTLRAVPLPAAPCPPRPWQLGVGAVR-TLRTGPALLS 49  Homo sapiens
MALRAVRSVRAAVGGLRAISAPSAPCLPRPWGLRAGAVR-ELRTGPALLS 49  Bos taurus
MAWLVLRRLGPVL---------APRCPRLSLRPQVPAVR-RLGTGSLLLS 40  Gallus gallus
MALRM--WASSTANALKLSSS-SRLHLSPTFSIS----R-CFSN---VLD 39  Pisum sativum
MALRM--WASSTANALKLSSSVSKSHLSP-FSFS----R-CFST---VLE 39  Arabidopsis thaliana
MALRM--WASSTANALRLSSA-TRPHFSP---LS----R-CFSS---VLD 36  Flaveria chloraefolia
M----------------------------------------SNVPA  6  Escherichia coli
*

AVRYTSQHEWIAVHQDKTAFVGITKYATDSLGDATYVELPEVGTEISQGE 100 S.cerevisiae
VRKFTEKHEWVTTENGIG-TVGISNFAQEALGDVVYCSLPEVGTKLNKQD 98  Homo sapiens
VRKFTEKHEWVTTENGVG-TVGISNFAQEALGDVVYCSLPEVGTKLNKQE 98  Bos taurus
ARKFTDKHEWISVENGIG-TVGISNFAQEALGDVVYCSLPEIGTKLNKDD 89  Gallus gallus
GLKYAPSHEWVKHEGSVA-TIGITDHAQDHLGEVVFVELPEPGVSVTKGK 88  Pisum sativum
GLKYANSHEWVKHEGSVA-TIGITAHAQDHLGEVVFVELPEDNTSVSKEK 88  Arabidopsis thaliana
GLKHANSHEWVKHEGSVA-TIGITDHAQDHLGEVVFVDLPETGGSVTKAT 85  Flaveria chloraefolia
ELKYSKEHEWLRKEADGTYTVGITEHAQELLGDMVFVDLPEVGATVSAGD 56  Escherichia coli
  . .  ***.  . . .  .**.  * . **. .. .*** . ...  .
     ↓  ▼          ↓
SLGSIESVKSASEIYQPADGTVEEINTNLEENPGVVNEDPMGDGWLVKMK 150 S.cerevisiae
EFGALESVKAASELYSPLSGEVTEINEALAENPGLVNKSCYEDGWLIKMT 148 Homo sapiens
EFGALESVKAASELYSPLSGEVTEINKALAENPGLVNKSCYEDGWLIKMT 148 Bos taurus
EFGALESVKAASELYSPLTGEVTDINAALADNPGLVNKSCYQDGWLIKMT 139 Gallus gallus
GFGAVESVKATSDVNSPISGEVIEVNTGLTGKPGLINSSPYEDGWMIKIK 138 Pisum sativum
SFGAVESVKATSEILSPISGEIIEVNKKLTESPGLINSSPYEDGWMIKVK 138 Arabidopsis thaliana
GFGAVESVKATSDVNSPISGEIVEVNSKLSETPGLINSSPYEDGWMIKVK 135 Flaveria chloraefolia
DCAVAESVKAASDIYAPVSGEIVAVNDALSDSPELVNSEPYAGGWIFKIK 106 Escherichia coli
. .  ****..*..  * .*.. ..*  *...*...*..   .**. *..

LGEGVNVEQVEGLMSLEQYEKTLVHDD       U12980          177 S.cerevisiae
LSNPSELDELMSEEAYEKYIKSIE--E       M69175    34%   173 Homo sapiens
FSNPSELDELMSEEAYEKYIKSIE--E       M58361    33%   173 Bos taurus
VEKPAELDELMSEDAYEKYIKSIE--D       M64401    32%   164 Gallus gallus
PTSPDELESLLGAKEYTKFCEEEDAAH       X53656    31%   165 Pisum sativum
PSSPAELESLMGPKEYTKFCEEEDAAH       M82921    37%   165 Arabidopsis thaliana
PSNPSELDSLMGAKEYT----------       Z37520    34%   152 Flaveria chloraefolia
ASDESELESLLDATAYEALLED----E       M57690    38%   129 Escherichia coli
 ..  ... . .  . .
```

Table 4.

Glycine cleavage activity

Strain	**Enzyme Activity (pmol/min/mg protein)**
3634 (wt)	106
4070 (*ser1*)	157
4404 (*gcv3::URA3*)	0
DH5α	168

cleavage activity was detected in the *gcv3::URA3* mutant (Table 4). Therefore, *GCV3* encodes an essential component of the glycine cleavage system. Consistent with this, a search of the *S.cerevisiae* genome did not identify a second H-protein gene. That glycine cleavage activity was found to be about 1.5 times higher in the *ser1* strain (Table 4) may reflect the fact that a *ser1* mutant would be more dependent upon the glycine cleavage system for its supply of C1-units.

C.2. *GCV3* is Required to Utilize Glycine as a Nitrogen Source but not for Growth on Minimal Media.

Once it was established that *GCV3* was essential for glycine cleavage system activity in cell extracts, I wanted to study its role *in vivo*. Since *S. cerevisiae* can utilize glycine as a nitrogen source (Sinclair and Dawes, 1995; Sinclair et al., 1996), the effect of a *gcv3::URA3* mutation on using glycine as a nitrogen source was assessed. Figure 13 shows that the utilization of glycine is dependent upon a functional *GCV3* gene.

The *gcv3::URA3* mutant, like its wild-type parent, grows irrespective of the availability of glycine and serine (Fig. 14). Therefore *GCV3* is not essential for growth on SD. *GCV3* is required for growth at wild-type rates, since the *gcv3::URA3* strain, with a doubling time of 3.6 h in minimal medium, grew significantly slower than its wild-type parent (Table 5). It was unexpected but informative to find that glycine was toxic to the *gcv3::URA3* mutant and increased its doubling time from 3.6 h to 10.5 h. Finally, formate supplement enabled the *gcv3::URA3* strain to grow at the same rate as the wild-type parent (2.6 h).

Table 5.

Growth rate of wild-type, *ser1*, and *gcv3::URA3* strains grown in SD.

Strain	Generation time (hr)			
	No supplement	**with serine**	**with glycine**	**with formate**
3634 (wt)	2.7	2.6	2.6	2.5
4070 (*ser1*)	>30	3.0	3.8	14
4404 (*gcv3::URA3*)	3.6	3.5	10.5	2.6

Fig. 13.

***GCV3* is required for the utilization of glycine as a nitrogen source.** Growth on minimal media; *Left panel (*minimal medium with 250 mM glycine as the nitrogen source); *right panel* (minimal medium with 250 mM glycine and ammonium sulphate). *Plate sector A* (wild-type strain 2661); *plate sector B* (*gcv3::URA3* strain 3751); and *plate sector C* (strain 3751 transformed with pLNGCV3).

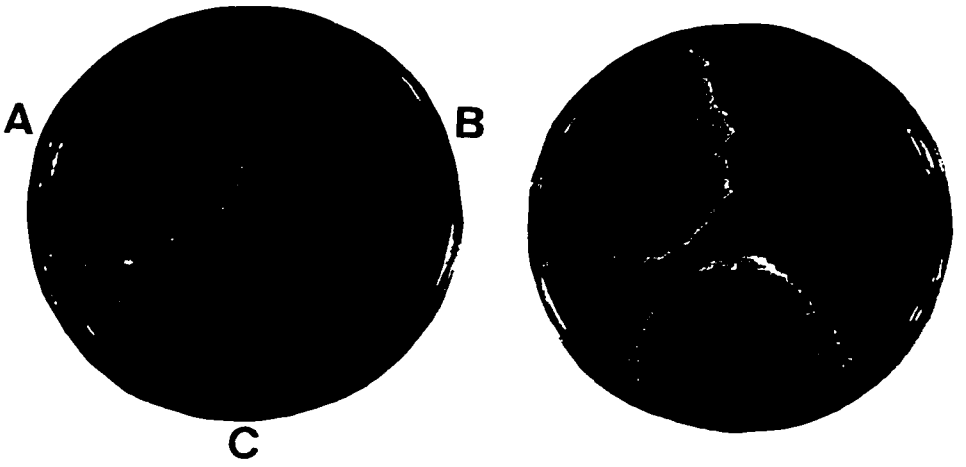
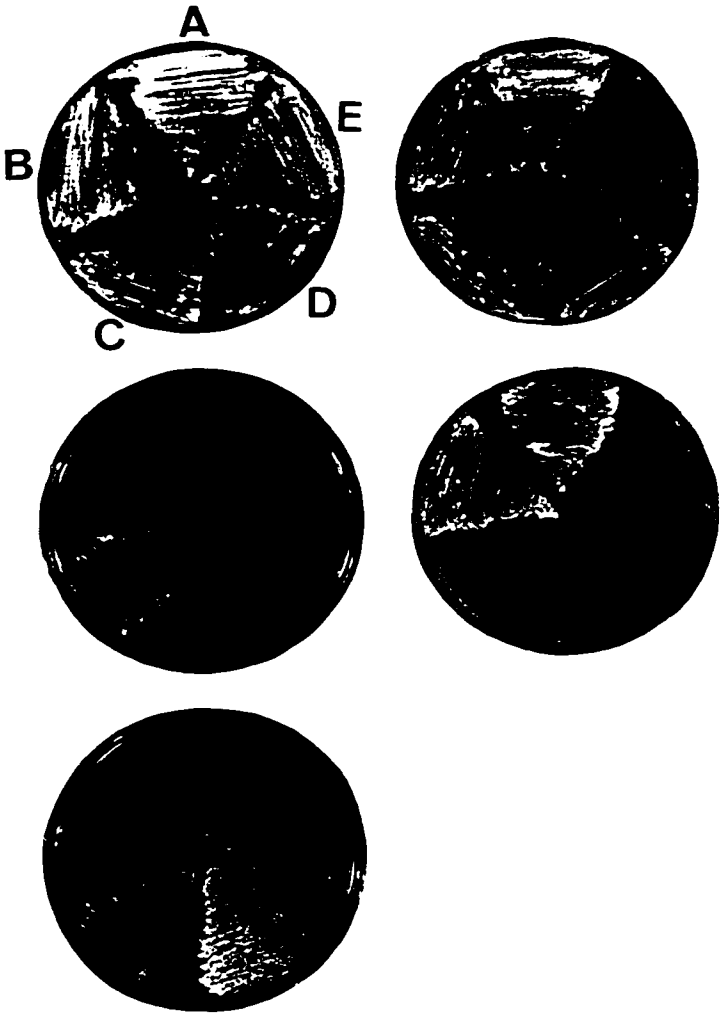

A
B
C

Fig. 14.

Glycine can be used in a glycine cleavage system dependent fashion to meet cellular demand for serine. Growth requirements of strains; *plate sector A* (strain DBY745); *plate sector B* (3751 *gcv3::URA3*); *plate sector C* (1095-302C *ser1*); *plate sector D* (3792B *ser1 gcv3::URA3*); and *plate sector E* (3792B harbouring plasmid pLNGCV3). The YNBD plates shown are supplemented with; *row 1 left* (glycine and serine); *row 1 right* (serine), *row 2 left* (glycine); *row 2 right* (no supplement); and *row 3* (glycine and formate).


A
B
C
D
E

C.3. *De Novo* Glycine Biosynthesis Cannot Meet the Growth Requirements of a *ser1* Mutant.

Most serine is derived from 3-phosphoglycerate (Fig. 2). Furthermore, although a *ser1* mutant (strain 4070) grows very slowly without supplement (> 30 h doubling time), it grows quite well with a 3.8 h doubling time if supplemented with glycine (Table 5). That glycine enables a *ser1* mutant to grow at near normal rates shows that *de novo* glycine biosynthesis cannot meet cellular demand for glycine in a *ser1* background. Apparently cellular demand for glycine, and the metabolic products serine and the C1-donor 5,10-MTHF exceeds the capacity of *ser1* strains to synthesize glycine. These results support the previous finding that supplement with serine or glycine enables *ser1* mutants to grow at near normal rates (McKenzie and Jones, 1977; Mortimer and Hawthorne, 1966).

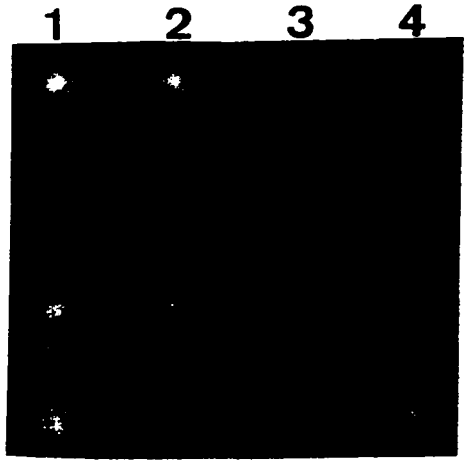
C.4. C1-Tetrahydrofolate Synthase is the Major and Perhaps only Route for Formate Synthesis in *S. cerevisiae*.

There are three main routes for the generation of activated one carbon units in *S. cerevisiae*. These pathways which obtain their C1-units from serine, glycine and formate utilize glycine hydroxymethyltransferase, the glycine cleavage system and C1-tetrahydrofolate synthase respectively (Nagarajan and Storms, 1997). That the double mutant (*ser1 gcv3::URA3*) grew very slowly relative to strains harbouring either one of these mutations (Fig. 15), suggested that the *ser1* and *gcv3::URA3* mutations combined to adversely affect growth on rich medium. To test whether this was due to loss of the glycine cleavage system dependent synthesis of 5,10-MTHF, I compared the growth of wild-type,

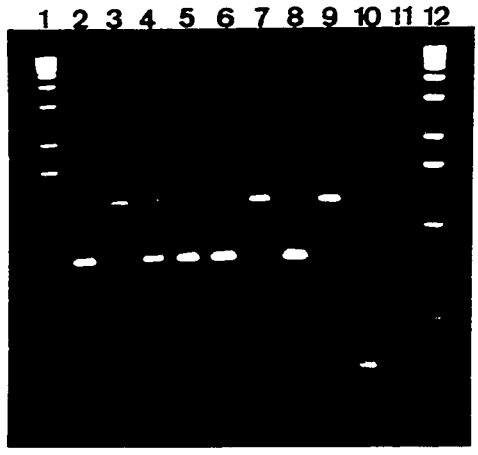
Fig. 15.

Characterization of *GCV3* null mutants. (A) Tetrad analysis of strain 3773, a strain heterozygous for *ser1* and *gcv3::URA3*. Columns 1 & 2 show the products obtained from tetratype and non-parental ditype tetrads respectively. The four products shown in column two were designated strains 3792A, 3792B, 3792C, & 3792D (Table 3). Columns 3 and 4 show the meiotic products of two tetrads from strain 3771. (B) PCR based characterization of strains 3792A, 3792B, 3792C and 3792D. Agarose gel electrophoresis of DNA samples; *lanes 1 and 12* (1 kbp DNA ladder); *lanes 2 through 9* (PCR products generated using oligonucleotides GC5 and GC6); *lane 10* (PCR product generated using oligonucleotides GC5 and UR1). The template DNAs used were; *lane 2* (strain 2661); *lane 3* (strain 3751); *lane 4* (diploid 3773); *lane 5* (diploid 3771), *lane 6* (strain 3792-A); *lane 7* (strain 3792B); *lane 8* (strain 3792C); *lane 9* (strain 3792D); and *lane 10* (strain 3792B). Lane 11 does not contain a DNA sample.

A



B



ser1, *gcv3::URA3* and *ser1 gcv3::URA3* mutants on media supplemented with different combinations of glycine, serine and formate (Fig. 14). Unlike the *ser1* mutant the double mutant could not be rescued by glycine supplement. This suggested that glycine enabled the *ser1* mutant to grow at normal rates, because it served as a precursor for both serine and 5,10-MTHF synthesis. Since formate can also serve as a precursor for 5,10-MTHF synthesis the ability of formate supplement to support growth was also tested. Formate and glycine enable the *ser1 gcv3::URA3* strain, 3792B, to grow at normal rates (Fig. 14). That formate supplement enabled strain 3792B to grow well suggests that under these growth conditions formate synthesis is dependent upon C1-tetrahydrofolate synthase. In contrast the *de novo* synthesis of formate in *E. coli* can meet cellular demand for activated C1-units (Knappe, 1987).

C.5. *GCV3* is Induced by Glycine and Repressed by the C1-Metabolic End Products.

Enzyme activity is often regulated at the level of gene expression in response to a substrate or demand for a metabolic product. To study *GCV3* regulation strains 3634 (wt) and 4070 (*ser1*), were used (Fig. 16A). Strain 4070 was included to test whether *GCV3* expression was further elevated when demand for C1-units derived from glycine via the glycine cleavage system was increased. *GCV3* expression was measured using the *lacZ* reporter gene on pLNGCV3-211 (Fig. 4). A dose response curve revealed that pLNGCV3-211 expression was induced by glycine and that it reached a plateau at about 10 mM glycine for the *SER1* strain and at 20 mM for the *ser1* mutant (data not shown). Glycine induced pLNGCV3-211 expression 30-fold in the *SER1* strain (Fig. 16A). In the *ser1* mutant

expression was induced only 10-fold; however, both the expression without glycine supplement and fully induced expression were higher in the *ser1* strain (Fig. 16A).

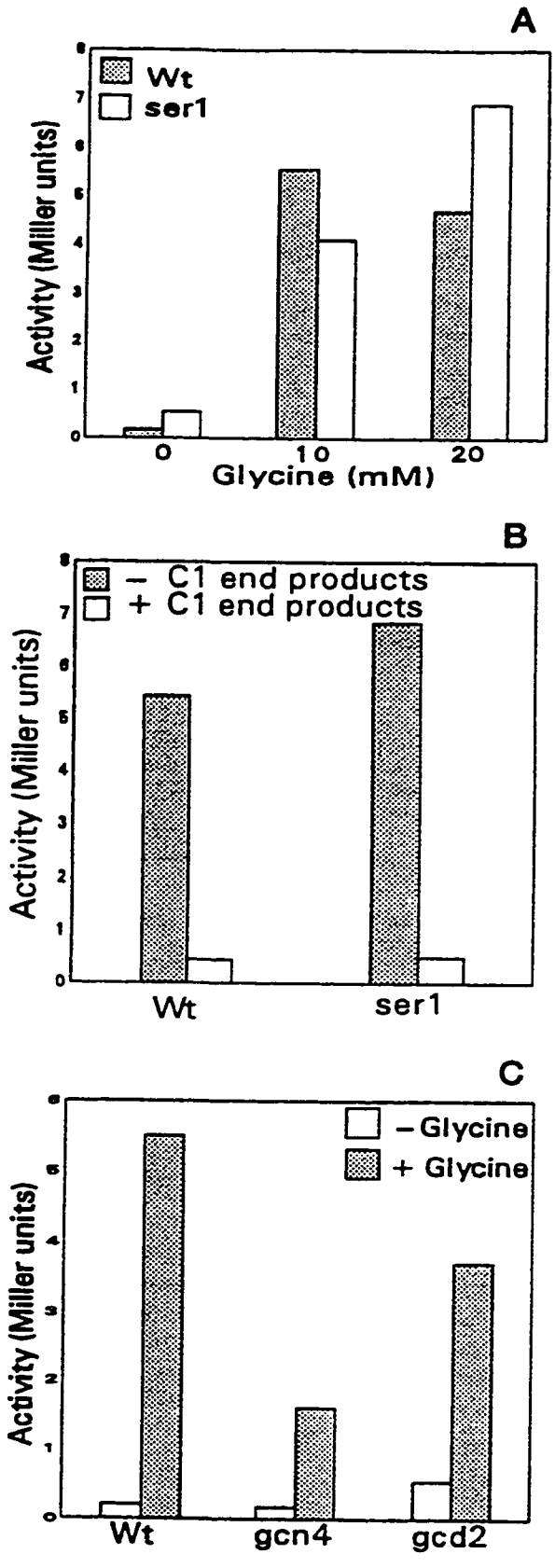
Since the glycine cleavage system generates 5,10-MTHF, *GCV3* expression might be repressed when cellular demand for C1-units is low. Indeed, supplement with the metabolic products that require C1-units for their syntheses repressed pLNGCV3-211 expression about 10-fold (Fig. 16B).

C.6. General Control of *GCV3* Expression.

The presence of three GCRE elements upstream of the *GCV3* open reading frame (Fig. 3), suggested that *GCV3* is regulated by the general amino acid control system. To test this the expression of the pLNGCV3-211 reporter gene was examined in strains 3634 (wt), 3640 (*gcd2*) and 3646 (*gcn4*) (Fig. 16C). If a gene is regulated by the general amino acid control system, its expression increases in response to amino acid starvation because the concentration of the transcriptional activator Gcn4p increases (Struhl and Davis, 1981; Hinnebusch and Fink, 1983; Arndt, and Fink, 1986). Here I have used strains 3646 (*gcn4*), which does not express any Gcn4p, and 3640 (*gcd2*), which constitutively expresses induced levels of Gcn4p, to test whether pLNGCV3-211 expression is subject to general control (Mösch et al., 1990). pLNGCV3-211 expression in the *gcd2* strain, which mimics amino acid starvation conditions, is 2.5 fold higher than in the *gcn4* mutant which mimics nonstarvation conditions. This represents a general control response that falls within the range reported for other genes that are subject to general control (Hinnebusch, 1992). A similar effect was obtained when histidine starvation was induced by addition of 3-amino

Fig. 16.

Regulation of *GCV3* expression. (*A*) Effect of glycine supplement on levels of ß-galactosidase expressed by the isogeneic strains 3634 (wild-type) and 4070 (*ser1*) harbouring pLNGCV3-211. (*B*) Effect of the C1-metabolic end products on pLNGCV3-211 expression in strains 3634 and 4070. Cells were grown in SD + 10 mM glycine or SD +10 mM glycine and the C1-metabolic end products adenine (40 µg/ml), histidine (20 µg/ml) methionine (20 µg/ml), serine (5 mM) and formate (10 mM) as indicated. (*C*), Effect of the general control system on *GCV3-lacZ* expression in the isogenic strains 3634 (wild-type), 3646 (*gcn4-103*) and 3640 (*gcd2-1*).


A
Wt
ser1
Activity (Miller units)
0
1
2
3
4
5
6
7
8
0
10
20
Glycine (mM)
B
– C1 end products
+ C1 end products
Activity (Miller units)
0
1
2
3
4
5
6
7
8
Wt
ser1
C
– Glycine
+ Glycine
Activity (Miller units)
0
1
2
3
4
5
6
Wt
gcn4
gcd2

triazole (Hope et al., 1988; Wek et al., 1992) (data not shown).

C.7. The *GCV3* promoter is complex.

pLNGCV3-211 expression studies using the intact *GCV3* promoter revealed that expression was induced by amino acid starvation (general control) and glycine, and repressed by the C1-end products (see Fig.16). This suggested that the *GCV3* promoter region might contain several regulatory elements. To identify the upstream elements that control *GCV3* expression a set of nine promoter deletion mutants (pLNGCV3-211 through pLNGCV3-62) was constructed (Fig. 6). These promoter constructs place the *lacZ* coding region under the control of different portions of the *GCV3* promoter. All nine have *GCV3* portions extending from nucleotide +8 at their 3' ends to nucleotide -211 for the longest construct through -186, -165, -155, -140, -128, -111, -96, -86, and -62 for the shortest. Once the plasmids shown in Figure 6 had been constructed and their sequence verified by DNA sequencing, at least three independent constructs of each deletion plasmid (Fig. 6) were transformed into the wild-type yeast strain 3634. These transformants were then assessed for their ability to express ß-gal activity when grown in SD and SD + 10 mM glycine.

The results shown in Table 6 can be summarized as follows. First, several upstream regions contribute to the regulation of *GCV3* expression. Expression by the pLNGCV3-186 and pLNGCV3-211 transformants is essentially the same. Therefore, the 25 bp upstream of -186 bp does not contain information important for expression under the two growth conditions tested. Deleting an additional 21 bp (construct pLNGCV3-165) resulted in a 2-fold decrease in expression in both SD and SD + 10 mM glycine. Therefore this region has

information that affects *GCV3* expression.

Removing an additional 10 bp (construct pLNGCV3-155) did not further alter expression significantly, suggesting that this region does not contain regulatory information. However, transformants harbouring both constructs pLNGCV3-140 and pLNGCV3-128 expressed 5-fold more activity in SD and 2-fold more activity in SD + 10 mM glycine than did the pLNGCV3-155 transformants. This suggested that a regulatory element that overlapped the region between -155 bp and -140 bp negatively affected expression in these growth conditions.

Relative to construct pLNGCV3-128, pLNGCV3-111 expressed more ß-galactosidase in SD, but less activity in the presence of 10 mM glycine. This suggests that information important for *GCV3* regulation overlaps this 17 base pair region. Deleting an additional 15 bp generated construct pLNGCV3-96. Transformants harbouring this plasmid expressed 1.6-fold and 2.1-fold more activity in SD and SD + 10 mM glycine than did the pLNGCV3-111 transformants. Therefore, the region between -111 bp and -96 bp had a negative effect on expression. pLNGCV3-86 expressed drastically reduced levels of ß-gal regardless of the availability of glycine. A further decrease in expression was also observed when the results with pLNGCV3-86 and pLNGCV3-62 were compared. Therefore, information that positively affects *GCV3* expression is found in both the regions between -96 bp and -86 bp and the region between -86 bp and -62 bp.

These results (Table 6) suggest that at least six regulatory elements are found in the *GCV3* promoter. Furthermore, under the growth conditions tested; three of these have a

Table 6.

Expression of *GCV3* promoter deletion constructs in a wild-type strain (3634).

	ß-Gal activity (Miller Units)	
Plasmid[a]	**SD**	**SD + 10 mM Glycine**
pLNGCV3-211	0.24	5.75
pLNGCV3-186	0.25	5.64
pLNGCV3-165	0.13	3.35
pLNGCV3-155	0.11	3.27
pLNGCV3-140	0.54	6.52
pLNGCV3-128	0.57	6.56
pLNGCV3-111	0.88	1.73
pLNGCV3-96	1.48	3.74
pLNGCV3-86	0.04	0.35
pLNGCV3-62	0	0

[a] The plasmid name indicates the nucleotide deletion position relative to the first base of the start codon.

positive effect on expression in both SD and SD + 10 mM glycine, two have a negative effect on expression in both these media and the last element negatively affects expression in SD and positively affects expression in SD + 10 mM glycine (Fig. 17).

C.8. Localization of cis-acting sequences for general amino acid control.

GCV3 is subject to regulation by the general amino acid control system (Fig. 16). To localize the sequences for general amino acid control the plasmids depicted in Figure 6 were transformed into strains 3640 (*gcd2*), which expresses increased levels of Gcn4p, and 3646 (*gcn4*) which does not express Gcn4p (Table 7 and Table 8). Plasmid pLNGCV3-111 and all the larger plasmids respond to general amino acid control (i.e., were expressed at 3 to 4-fold higher levels in the *gcd2* strain than in *gcn4* strain) (Table 7 and 8). In contrast, the general control response of the shorter fusion, pLNGCV3-96, was significantly reduced (by about 2-fold). The two shortest fusion genes did not respond to general amino acid control, although the shortest fusion was not expressed at detectable levels (Table 6, 7 and 8). Since pLNGCV3-111 responds to general amino acid control, the information for general amino acid control is within -111 bp of the start codon. This region contains three potential general control response elements (GCRE) (Mösch et al., 1991; Mösch et al., 1990) located at -77 bp, -87 bp, and -107 bp (Fig. 17). Similar results were obtained when these same transformants were grown with and without 10 mM glycine, although *GCN4*-dependent induction of *GCV3* was less pronounced with glycine.

Fig. 17.

Summary of the functional regions (upstream elements). The horizontal line represents the *GCV3* promoter region extending from nucleotide -1 to -211. The arrows above the horizontal line represent positions of the deletions in the *GCV3* promoter. Boxes represents potential cis-acting upstream elements. Horizontal arrows indicates the orientation of the consensus sequence. The numbers near the boxes indicate the positions of the upstream elements relative to the *GCV3* start codon. Plus (+) and minus (-) signs in the three rows below the horizontal line indicate whether the region depicted enhances or represses expression under the physiological conditions depicted on the right. SD, minimal media; SD + 10 mM glycine, minimal media supplemented with 10 mM glycine as an inducer; SD + 250 mM glycine, minimal media supplemented with 250 mM glycine as the sole nitrogen source.

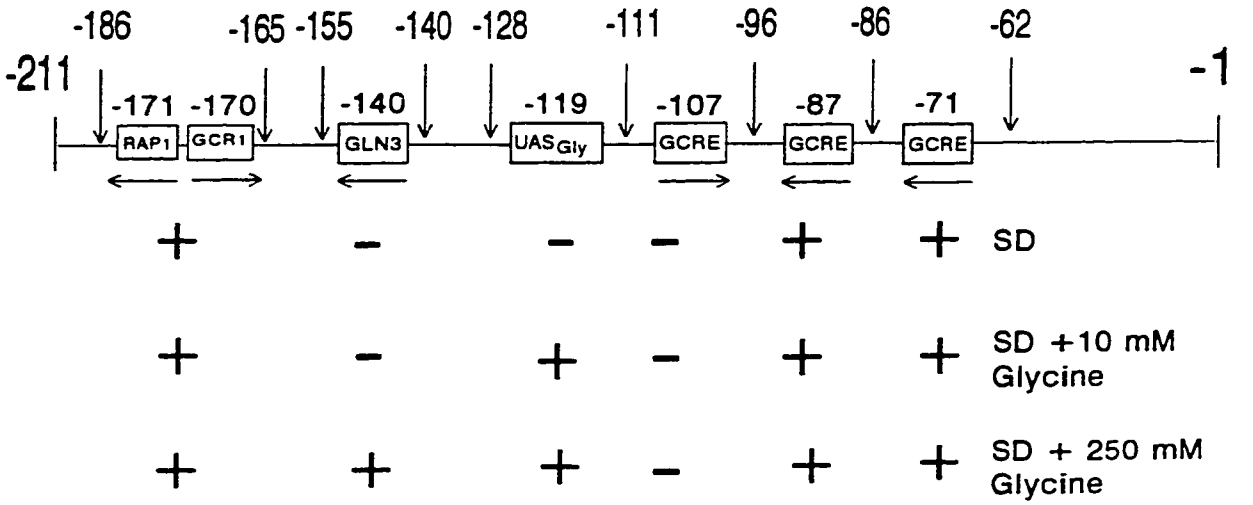

-186
-165
-155
-140
-128
-111
-96
-86
-62
-211
-1
-171
-170
-140
-119
-107
-87
-71
RAP1
GCR1
GLN3
UASGly
GCRE
GCRE
GCRE
+ - - - + + SD
+ - + - + + SD +10 mM Glycine
+ + + - + + SD + 250 mM Glycine

Table 7.

Expression of *GCV3* promoter deletion constructs in SD.

	ß-gal Activity (Miller units)		
Plasmid[a]	**3646** ***gcn4***	**3640** ***gcd2***	**General control response[b]**
pLNGCV3-211	0.15	0.48	+
pLNGCV3-186	0.13	0.47	+
pLNGCV3-165	0.09	0.21	+
pLNGCV3-155	0.07	0.19	+
pLNGCV3-140	0.13	0.61	+
pLNGCV3-128	0.16	0.58	+
pLNGCV3-111	0.35	1.15	+
pLNGCV3-96	0.63	1.46	+
pLNGCV3-86	0.01	0.01	-
pLNGCV3-62	0	0	-

[a] The plasmid names indicates the nucleotide deletion position relative to the first base of the start codon.

[b] + sign represents general control response, - sign represents abolished general control response.

Table 8.

Expression of *GCV3* promoter deletion constructs in SD + 10 mM glycine.

	ß-gal activity (Miller units)		
Plasmid[a]	**3646 *gcn4***	**3640 *gcd2***	**General control response[b]**
pLNGCV3-211	1.76	3.81	+
pLNGCV3-186	1.71	3.85	+
pLNGCV3-165	0.98	2.37	+
pLNGCV3-155	1.08	2.42	+
pLNGCV3-140	1.42	2.45	+
pLNGCV3-128	1.40	3.97	+
pLNGCV3-111	0.53	2.14	+
pLNGCV3-96	1.46	3.18	+
pLNGCV3-86	0.21	0.22	-
pLNGCV3-62	0	0	-

[a] The plasmid name indicates the nucleotide deletion position relative to the first base of the start codon.

[b] + sign represents general control response, - sign represents abolished general control response.

C.9. Expression of promoter deletion constructs when glycine is the sole nitrogen source.

To localize the upstream information that responded to glycine the expression of ß-galactosidase by the wild-type strain (3634) harbouring each of the deletion constructs was assessed when grown in SD + 250 mM glycine as the sole nitrogen source. These growth conditions were used to monitor the effect of glycine (a poor nitrogen source) without the potentially repressing effects of NH_4 (a preferred nitrogen source). At least three interesting differences were observed between the pattern of gene expression in the presence of glycine as the sole nitrogen source (Table 9) and the pattern observed under the previously tested growth conditions (Table 6).

First, when glycine was the sole nitrogen source the wild-type strain harbouring pLNGCV3-140 expressed 3-fold less activity than the pLNGCV3-155 transformant; however, when growth was in the presence of NH_4 or NH_4 and 10 mM glycine the shorter construct expressed about 5-fold and 2-fold more activity respectively. These results suggest that the region between -155 bp and -140 bp enhances expression when glycine is the sole nitrogen source (Table 9) but represses expression when NH_4 is present (Table 6). Second, pLNGCV3-111 expressed about 30% of the level obtained with the next longest construct (Table 9). This reduction was also produced by supplementing SD + 10 mM glycine (Table 6), but was not observed in SD. Therefore the 17 bp region between -128 bp and -111 bp positively affects expression in the presence of extracellular glycine (Table 6 and 9). Finally, when glycine served as the sole nitrogen source the effect of general control was masked (Table 9). The mechanism that overrides the general control response in the presence of 250

Table 9.

Expression of *GCV3* deletion constructs in SD+ 250 mM glycine as the sole nitrogen source.

	ß-Gal Activity (Miller Units)		
Plasmid[a]	**3634 Wt**	**3646 *gcn4***	**3640 *gcd2***
pLNGCV3-211	7.24	2.75	4.16
pLNGCV3-186	7.36	2.81	4.19
pLNGCV3-165	6.07	2.16	3.18
pLNGCV3-155	5.96	2.37	2.85
pLNGCV3-140	2.12	0.98	1.18
pLNGCV3-128	2.18	1.22	1.36
pLNGCV3-111	0.57	0.34	0.38
pLNGCV3-96	0.94	0.27	0.33
pLNGCV3-86	0.03	0.02	0.03
pLNGCV3-62	0	0	0

[a] The plasmid name indicate the nucleotide deletion position relative to the first base of the start codon.

mM glycine is not clear.

C.10. The nitrogen regulation system is important for *GCV3* expression.

Removing the 15 bp region between -140 bp and -155 bp caused expression to increase 5-fold in SD, and 2-fold in SD + 10 mM glycine (Table 6). However, when glycine was the sole nitrogen source, deleting this 15-bp region reduced expression by 65% (Table 9). The major difference in the growth conditions used for generating the data in these two tables was that the media used for Table 6 included the preferred nitrogen source NH_4SO_4, whereas for Table 9 glycine was used. This suggested that *GCV3* is subject to regulation by the nitrogen regulation system (NRS).

That NRS was important for *GCV3* regulation was verified in two ways. First, since the NRS regulates gene expression in response to the quality of the nitrogen source, I asked whether varying the quality of the nitrogen source affected *GCV3* expression. Second, I asked whether *GCV3* regulation was dependent upon Gln3p, a transcription factor known to be important for gene regulation by the NRS. For this pLNGCV3-211 was transformed into the wild-type strain PM38 and an isogenic derivative with defective NRS (PM81). The ß-gal activity expressed by transformants of the wild-type strain (PM38) grown with the poor nitrogen source proline were 6 to 8 times higher than the activities expressed when a rich nitrogen source (glutamine, glutamate, asparagine or NH_4S0_4) was used (Table 10). Similarly when glycine (also a poor nitrogen source) was used 14 times as much ß-gal was expressed (Table 10). That expression in media containing proline, a poor nitrogen source, was much higher than the expression with a rich nitrogen source is consistent with *GCV3*

being regulated by the NRS.

Further support for the nitrogen-regulation of *GCV3* is provided by the results obtained with the nitrogen regulation defective *gln3* strain, PM81. This strain unlike the *GLN3* strain did not express pLNGCV3-211 at high levels when proline was the nitrogen source (Table 10). That the *gln3* mutation reduces expression with glycine as the nitrogen source shows that the elevated expression of *GCV3* in response to glycine is in part Gln3p-dependent.

GCV3 transcription in the presence of proline and glycine is apparently activated by Gln3p. In contrast with ammonia as nitrogen source Gln3p apparently represses *GCV3* expression. This result is similar to that reported previously for expression of the general amino acid permease gene *GAP1* (Stanbrough et al., 1995). These authors suggest that in the presence of ammonia Gln3p is present in a form that retains DNA binding activity but cannot activate transcription. Therefore, in ammonia Gln3p represses transcription.

The above result suggest that Gln3p is involved in the regulation of *GCV3*. The NRS apparently controls *GCV3* expression through Gln3p-binding to the sequence 5'-GATAA-3' located at base pair -140. The *gln3* allele did not reduce expression as dramatically when the nitrogen source was glycine as it did when it was proline. Perhaps the role of Gln3p is masked or replaced by another transcriptional activator(s) in the presence of glycine. The pLNGCV3-128 construct expressed six times more ß-gal activity than did the pLNGCV3-155 when NH_4SO_4 was used as the nitrogen source, suggesting that the Gln3p binding site is required for repressed expression in response to a rich nitrogen source (Table 11).

Table 10.

Activation of transcription of GCV3 by Gln3p in response to various nitrogen sources.

	ß-gal Activity (Miller Units)	
Nitrogen[a] source	**PM38** ***GLN3***	**PM81** ***gln3***
Glycine	108.31	65.01
Proline	67.64	12.50
Glutamine	11.51	21.5
Glutamate	9.48	1.42
Asparagine	11.94	9.59
Ammonia	7.51	47.53

[a] The compounds indicated were used as the sole nitrogen source

C.11. Two physically and functionally distinct regions, UAS_{GATAA} and UAS_{Gly}, enhance *GCV3* expression in response to glycine.

Although *GCV3* expression is induced about 30-fold by glycine (Nagarajan and Storms, 1997) (Fig. 16) only two of the six upstream regions that are important for regulation are responsible for the induction by glycine. One of these was inactivated by deleting the region between -155 bp and -140 bp and the other by deleting the region between -128 bp and -111 bp (Table 6 and 9). Removing the most upstream of those two regions caused 3-fold decreased expression in SD + 250 mM glycine but increased expression 5-fold in SD. It, therefore, appeared that the region between -155 bp and -140 bp regulated expression in response to the nitrogen source. Furthermore, since many genes required for the utilization of nitrogen sources are regulated in a Gln3p-dependent fashion via the UAS_{GATAA} (Coffman et al., 1996; Blinder et al., 1996) the presence of a UAS_{GATAA} consensus sequence within this region (Fig. 17) suggested that *GCV3* is subject to Gln3p-dependent regulation.

The presence of a functional UAS_{GATAA} was verified in two ways. First, the ability of constructs pLNGCV3-140 (a plasmid missing this UAS) and pLNGCV3-155 (a plasmid harbouring UAS_{GATAA}) to respond to nitrogen source dependent regulation were compared (Table 11). These results show that regulation in response to the nitrogen source requires this region. Second I asked whether the ability to express *GCV3* in a nitrogen source dependent fashion was mediated by Gln3p. For this these plasmids were transformed into strains PM38 (*GLN3*) and PM81 (*gln3*). Response to whether asparagine or proline is used as the nitrogen source is Gln3p-dependent (Table 11).

Table 11.

Localization of the Gln3p dependent transcriptional activation region of *GCV3*.

		ß-gal Activity (Miller Units)			
Plasmid	**Strain[a]**	**NH_4So_4**	**Glycine**	**Proline**	**Asparagine**
pLNGCV3-140	Wt	0.79	1.4	0.48	0.42
pLNGCV3-140	*gln3*	0.86	1.4	0.47	0.46
pLNGCV3-155	Wt	0.47	3.6	2.94	0.38
pLNGCV3-155	*gln3*	1.1	1.3	0.42	0.57

[a] The wild-type strain (*GLN3*) and mutant (*gln3*) strains used were the isogenic strains PM38 and PM81.

Preferred nitrogen sources are **NH_4SO_4** and asparagine.

Non-preferred nitrogen sources are glycine and proline.

ß-gal activities were measured in the presence of these nitrogenous compounds as the sole nitrogen in SD media.

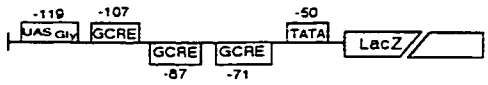
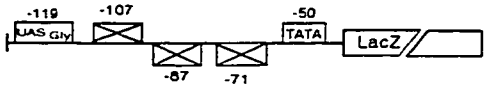
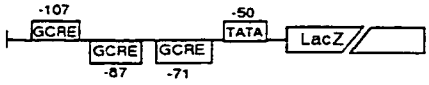
The region between bp -128 and bp -111 is also important for high levels of expression in SD + 250 mM glycine. Deleting this region caused expression to increase slightly in SD medium but reduced expression about 4-fold in SD + 250 mM glycine (Fig 18). A search of the upstream regions adjacent to all 4 glycine cleavage system genes identified an element (5'-RCCTCGM-3') that was present in all 4 promoters. The presence of this element within the 17 bp region that harbours a glycine-responsive element (Fig. 17), suggests that all four GCS genes are coordinately induced by glycine via a transcriptional control system that utilizes the upstream sequence 5'-RCCTCGM-3'.

C.12. UAS_{Gly} dependent *GCV3* regulation.

Although pLNGCV3-111 responded to the presence of Gcn4p (Table 7 and 9), it was not induced by glycine. The next larger construct, pLNGCV3-128, was induced by glycine and responded to Gcn4p (Table 7 and 9). Therefore, distinct elements are responsible for general amino acid control and induction by glycine. Inactivating all three of the GCREs (pLNGCV3-128A) abolished essentially all expression regardless of the growth condition (Fig. 18). Apparently, the glycine response region (-111 to -128) only enhances expression in the presence of another upstream promoter element. Perhaps this region interacts with a transcription factor whose binding or ability to enhance transcription is dependent upon another transcription factor. Transcription factors with activity that depends upon the binding of another regulatory protein have been described previously. For example the Gcr1p-dependent enhancement of glycolytic gene expression is dependent upon Rap1p binding to adjacent sites (Chambers et al., 1995; Tornow et al., 1993). Another example is

Fig. 18.

Identification and function of UAS_{Gly}. pLNGCV3-128, pLNGCV3-128A and pLNGCV3-111 expression in a wild-type strain (3634). ß-gal expression was measured under two different physiological conditions in SD: minimal media containing ammonia and SD + 250 mM glycine as the sole nitrogen source. Crossed boxes represents mutated GCREs. The numbers indicate the nucleotide position of the upstream elements relative to the *GCV3* start codon. The plasmid designations are depicted on the left.

		ß-Gal Activity (Miller Units)	
		Ammonia	Glycine
pLNGCV3-128		0.57	2.2
pLNGCV3-128A		0	0
pLNGCV3-111		0.88	0.57

the transcriptional activator Dal82p. Dal82p binds to an element (UIS_{DAL}) in the promoter regions of *DAL2*, *4* and *7*, and *DUR1*, *2* and *3*. Transcriptional activation that utilizes the UIS_{Dal} is dependent upon the binding of Dal82p; however UIS_{Dal} -dependent activation is only observed when a second transcriptional activator, for example Gln3p, binds at an adjacent site (Yoo and Cooper, 1989).

C.13. Gcr1p a positive regulator of *GCV3*.

Expression studies showed that pLNGCV3-186 expressed 2-fold more activity in SD and 1.7-fold more activity in SD + 250 mM glycine than pLNGCV3-165 (Table 6). I found a potential binding site for the glycolytic regulator Gcr1p at -170 bp (Fig. 17). This suggested that Gcr1p, a transcription factor that activates glycolytic gene expression by binding to 5'-CTTCC-3' sites (Willett et al., 1993; Scott and Baker, 1993), might also be regulating *GCV3* expression. Since transcriptional activation by Gcr1p is apparently dependent upon a proximal Rap1p site (Chambers et al., 1995; Tornow et al., 1993), the presence of a potential Rap1p binding site at -171 bp of the non-coding strand (Fig. 17) further supports the notion that *GCV3* is regulated by Gcr1p.

To test the possibility that expression is regulated by Gcr1p the expression of pLNGCV3-186 and pLNGCV3-165 in the *gcr1* strain HBY4 and its *GCR1* parent was compared. ß-gal amounts expressed by pLNGCV3-186 in the parental strain is at least 2-fold greater than its expression in the *gcr1* mutant both in SD and SD + 250 mM glycine (Table 12). The ability of glycine to induce *GCV3* was not affected by the *gcr1* mutation since expression was induced 7.5-fold and 11-fold in the *gcr1* and *GCR1* strains respectively

(Table 12). In contrast pLNGCV3-165 is expressed at the same level in both the *gcr1* and *GCR1* strains.

C.14. Regulatory activity of regions harbouring potential Gln3p and Gcr1p binding sites.

The inability of pLNGCV3-128A to express ß-gal activity provided an opportunity to assess the regulatory importance of sequences beyond base pair -128 in the absence of potentially masking effects imposed by the GCRE elements. For these studies I used pLNGCV3-gcre-156 and pLNGCV3-gcre-176 which are identical to pLNGCV3-128A except they also include respectively the adjacent 23 bp and 48 bp of upstream information (Fig. 18). pLNGCV3-gcre-156 was constructed to assess a potential Gln3p binding site at -140 bp. Similarly pLNGCV3-gcre-176 was constructed to test the importance of two potential Gcr1p binding sites that are found at -170 bp and -160 bp.

pLNGCV3-gcre-156 directed ß-gal expression was not detectable in SD medium; however, in the presence of 10 mM glycine and glycine as the sole nitrogen source, 0.15 and 0.76 Miller units were expressed (Fig. 8). Including the two potential Gcr1p sites (construct pLNGCV3-gcre-176) resulted in detectable expression in SD (0.15 Miller units) and expression was increased compared with pLNGCV3-gcre-156 in the two other growth conditions (Fig. 8). These results show that the region harbouring a potential Gln3p site and the region with the two GCR1 consensus sites can contribute to *GCV3* regulation.

Table 12.

Activation of transcription of *GCV3* by Gcr1p.

		ß-gal activity (Miller Units)		
Plasmids	**Strains[a]**	**SD**	**SD + 250 mM glycine**	**SD + 10 mM glycine**
pLNGCV3-165	Wt	0.47	4.86	3.14
pLNGCV3-165	*gcr1*	0.48	4.72	3.22
pLNGCV3-186	Wt	1.18	8.96	7.38
pLNGCV3-186	*gcr1*	0.46	5.28	3.17

[a] The wild-type strain (*GCR1*) and mutant (*gcr1*) strains used were the isogeneic strains S150-2B and HBY4.

C.15. *In vivo* and *in vitro* analysis of GCREs.

GCV3 is subject to transcriptional regulation by the general amino acid control system (Fig. 13). The results in Tables 7 and 8 show that the cis-acting regulatory information for general amino acid control is found within -111 bp of the start codon. The DNA sequence of this region contains 3 potential general control response elements (GCRE) (Fig. 17). All three copies conform to the previously reported GCRE consensus sequence (Hinnebusch, 1992; Mösch et al., 1990; 1991; Kunzler et al., 1995).

To assess the regulatory importance of each potential GCRE, the plasmids pLNGCV3-128A, pLNGCV3-128B, pLNGCV3-128C, and pLNGCV3-128D were constructed. All four are identical to pLNGCV3-128 except that one or more of their GCRE elements have been mutated. These mutations altered the sequence of the most upstream (5'-TGACAC-3'), the central (5'-TGACTC-3') and the downstream (5'-TGACTC-3') elements to 5'-TTACAC-3', 5'-TTACGC-3' and 5'-TTACGC-3' respectively. These changes convert a potentially functional GCRE to an element that should not respond to general amino acid control (Mösch et al., 1990; 1991; Kunzler et al., 1995). The construct pLNGCV3-128A has all three GCREs mutated, whereas the constructs pLNGCV3-128B, pLNGCV3-128C, and pLNGCV3-128D each have one potential GCRE element inactivated by mutation (Fig. 7). Strain 3634 harbouring the plasmids pLNGCV3-128A and pLNGCV3-128 was grown in SD + ammonia and SD + 250 mM glycine (Fig. 18). A comparison of pLNGCV3-128 and the construct with all three elements mutated, pLNGCV3-128A, showed that detectable expression from pLNGCV3-128 required at least one of these three potential GCREs.

C.16. *In vivo* assessment of GCRE activity.

Transformants of strains 3634 (wt), 3646 (*gcn4*), and 3640 (gcd2) harbouring the plasmids pLNGCV3-128A, pLNGCV3-128B, pLNGCV3-128C, pLNGCV3-128D and pLNGCV3-128 were grown in three conditions (SD, SD + 10 mM glycine, and SD + 250 mM glycine as the sole nitrogen source) and their expression of ß-gal determined (Table 13, 14 and 15). Expression of pLNGCV3-128 and the construct with all three elements mutated, pLNGCV3-128A, showed that detectable activity required at least one GCRE regardless of the strain or growth condition. Although at least one GCRE was required for detectable expression, mutating them individually produced some unexpected effects on reporter gene expression.

Mutating the most upstream GCRE (pLNGCV3-128B), had only a modest effect or in some cases no effect on regulation by the general amino acid control system. Therefore, the GCRE at -107 bp plays only a modest role in the general amino acid control of *GCV3*. However, mutating this element affected Gcn4p-independent expression. With one exception, pLNGCV3-128B expression was higher than that obtained with pLNGCV3-128 in all three strains whatever the growth condition (the exceptional case was in strain 3646 (gcn4) grown with glycine as the nitrogen source). Therefore, the upstream GCRE or a site overlapping it had a negative effect on expression in all three growth conditions tested (Table 13, 14 and 15). These results support those presented earlier (section C.7.) which found that a repressor site mapped to this region. However, since the ability of this site to repress gene expression was not dependent upon Gcn4p (ie: expression from pLNGCV3-128B was also increased relative to expression from pLNGCV3-128 in the *gcn4* mutant).

Table 13.

Expression of GCRE mutant constructs in SD.

	ß-gal activity (Miller units).			
Plasmid[a]	**3634 wt**	**3646 *gcn4***	**3640 *gcd2***	**General control response[b]**
pLNGCV3-128	0.57	0.12	0.37	+
pLNGCV3-128A	0	0	0	-
pLNGCV3-128B	0.93	0.64	1.57	+
pLNGCV3-128C	0.13	0.098	0.107	-
pLNGCV3-128D	0.31	0.21	0.44	+

[a] pLNGCV3-128A to pLNGCV3-128D contain combination of GCRE mutations (see Fig. 7 for detail).

[b] + sign represents general control response, - sign represents abolished general control response.

Table 14.

Expression of GCRE mutant constructs in SD + 10 mM glycine.

	ß-gal activity (Miller units).			
Plasmid[a]	**3634 wt**	**3646 *gcn4***	**3640 *gcd2***	**General control Response[b]**
pLNGCV3-128	6.52	1.38	3.55	+
pLNGCV3-128A	0	0	0	-
pLNGCV3-128B	9.68	3.49	7.36	+
pLNGCV3-128C	0.28	0.26	0.34	±
pLNGCV3-128D	0.78	0.43	0.58	±

[a] pLNGCV3-128A to pLNGCV3-128D contain combination of GCRE mutations (see Fig. 7 for detail).

[b] + sign represents general control response, - sign represents abolished general control response, ± sign represents poor general control response.

Table 15.

Expression of GCRE mutant constructs in SD + 250 mM glycine served as the sole nitrogen source.

	ß-gal Activity (Miller units)			
Plasmid[A]	**3634 Wt**	**3646 *gcn4***	**3640 *gcd2***	**General control response[B]**
pLNGCV3-128	2.2	1.2	1.4	-
pLNGCV3-128A	0	0	0	
pLNGCV3-128B	2.67	0.86	2.18	+
pLNGCV3-128C	< 0.1	< 0.1	< 0.1	-
pLNGCV3-128D	0.55	0.62	0.60	-

[a] pLNGCV3-128A to pLNGCV3-128D contain combination of GCRE mutations (see Fig. 7 for detail).

[b] + sign represents general control response, - sign represents abolished general control response.

repression apparently acts via another regulatory protein with a binding site that overlaps the upstream GCRE. Mutating the central GCRE dramatically decreased expression in the wild-type strain in all 3 growth conditions (Tables 13, 14, and 15). The effect of mutating this element was much larger when the wild-type strain was grown in the presence of glycine (>20-fold) than it was without glycine (5-fold). Mutating the -87 bp site did not affect expression in the *gcn4* strain when it was grown in SD, but did reduce expression in the *GCN4* and *gcd2* strains that express functional Gcn4p. Therefore the GCRE site at -87 bp contributes in a major way to the Gcn4p-dependent regulation of *GCV3*.

The -87 bp GCRE, or an overlapping element, also plays a regulatory role that is Gcn4p-independent. That is the *gcn4* strain when grown in SD + 10 mM glycine and SD + 250 mM glycine expressed 5-fold and 12-fold more ß-gal from pLNGCV3-128 than from pLNGCV3-128C.

Expression studies with pLNGCV3-128D revealed that the -87 bp GCRE contributed 2-fold, 8-fold and 4-fold to the expression of *GCV3* in the wild-type strain grown in SD, SD + 10 mM glycine, and SD + 250 mM glycine as the sole nitrogen source. Furthermore, although this mutation reduced expression by 50% in the wild-type strain grown in SD, it had the opposite effect (increased expression 2- fold) in the *gcn4* mutant grown under the same conditions.

In SD, the *gcn4* mutant expressed construct pLNGCV3-128D at only 50% of the level expressed by the gcd2 strain. Therefore, the -71 bp site is important for Gcn4p-dependent regulation. However, growth with glycine reduced the role Gcn4p mediated via this site. For example, when this site was mutated the *gcn4* mutant expressed 80% as much ß-gal as did

the *gcd2* mutant in SD + 10 mM glycine and 100% the activity of the gcd2 mutant when growth was with glycine as the sole nitrogen source. That expression is significantly reduced when either the -87 bp or -71 bp sites are mutated under all the conditions tested suggests that both these GCREs are used to regulate *GCV3* expression.

C.17. Gcn4p binds to *GCV3* promoter.

The above results suggest that Gcn4p contributes to the regulation of *GCV3* by binding to the two downstream GCREs. DNA-protein interaction studies were done to directly test the ability of Gcn4p to bind to these sites. The Gcn4p binding conditions, which were optimized using the GCRE-128 probe (Fig. 10), were essentially the same as those developed previously by Mösch et al. (1990). These optimized conditions were used to compare Gcn4p-binding to the gcre-128 and gcre-128A probes. gcre-128A has all three GCRE elements mutated (depicted in Fig. 9). The optimized conditions which produced Gcn4p/GCRE-128 complexes failed to detect Gcn4p/gcre-128A complexes (Fig. 19, lane 7). Therefore, Gcn4p binds to the *GCV3* promoter and binding is dependent upon the presence of at least one GCRE.

To find out whether all three potential GCREs could bind Gcn4p the DNA probes gcre-128E, gcre-128F, and gcre-128G (Fig. 9) were prepared. Each of these contains two mutated general control elements and one normal copy. Protein binding assays performed with each of these probes showed that all three elements contribute to Gcn4p binding (Fig. 19). Since all three sites could contribute to Gcn4p binding, it was unexpected to find that the gcre-128E, gcre-128F, and gcre-128G probes were retarded to the same extent as the

wild-type GCRE-128 probe (Fig. 19). There are at least three possible explanations for these results. One is that Gcn4p can bind all three sites but it can only bind to one site at a time. Alternatively, binding may be cooperative such that Gcn4p binding to one wild-type site facilitates binding to mutated sites. Finally, the DNA mobility shift performed here may be unable to distinguish between the mobility of probe Gcn4p complexes with one, two or three sites occupied. Further studies are needed to distinguish between these possibilities.

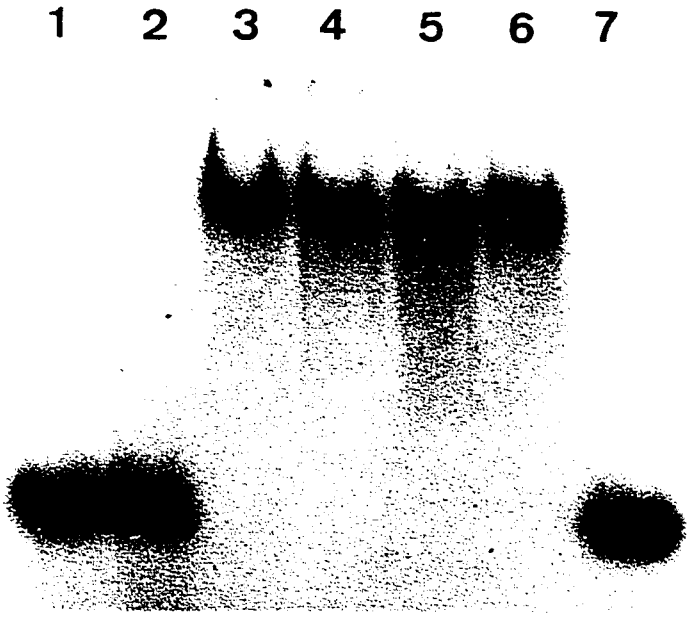
C.18. *BAS1* and *BAS2* are required for *GCV3* regulation.

pLNGCV3-128 expression was reduced 4-fold when *GCN4* was inactivated by mutation (Table 13 and Table 14). In contrast mutating all three Gcn4p binding sites reduced expression at least 60-fold (Table 13 and Table 14). Therefore these GCRE sites, or additional overlapping sites, interact with another transcriptional activator(s). Since Bas1p and Bas2p (Tice-Baldwin et al., 1989) can bind to promoter regions that overlap GCRE elements and both Bas1p and Bas2p are required for the regulation of *HIS4* and several purine genes subject to regulation by Gcn4p (Tice-Baldwin et al., 1989), Bas1p and/or Bas2p may also regulate *GCV3*. To test this possibility, I examined the expression of pLNGCV3-211 in the isogenic yeast strains CY614 (*gcn4*, *bas1*, and *bas2*), CY615 (*bas1* and *bas2*), CY945 (*gcn4* and *bas1*), 5014 (*bas2* and *gcn4*), and CY946 (*bas1* and *bas2*) (Devlin et al., 1991).

pLNGCV3-211 expression in strains CY614 (*gcn4*, *bas1*, and *bas2*), 5014 (*gcn4* and *bas2*), and CY945 (*gcn4*, and *bas1*) did not vary significantly regardless of the growth condition (Fig. 20). In striking contrast, ß-Gal expression by CY946 (*gcn4*) was dramatically

Fig. 19.

Gcn4p does not bind to gcre-128A. The ability of GCRE mutant PCR products gcre-128A, gcre-128E, gcre-128F, and gcre-128G (shown in Fig. 9) to complex with Gcn4p was tested by a gel retardation analysis. All the reaction mixtures were supplemented with 5 ng of appropriate radiolabelled probe. Lane 1 to 3 contains the GCRE-128 labelled probe. Lane 4, lane 5, lane 6 and lane 7 contain the mutant GCREs radiolabelled probe gcre-128E, gcre-128F, gcre-128G and gcre-128A, respectively. Lane 1 was not supplemented with any protein. Lane 2 was supplemented with 3.2 μg of *E.coli* total soluble protein. Lane 3 to 7 were supplemented with 3.2 μg of partially purified Gcn4p.

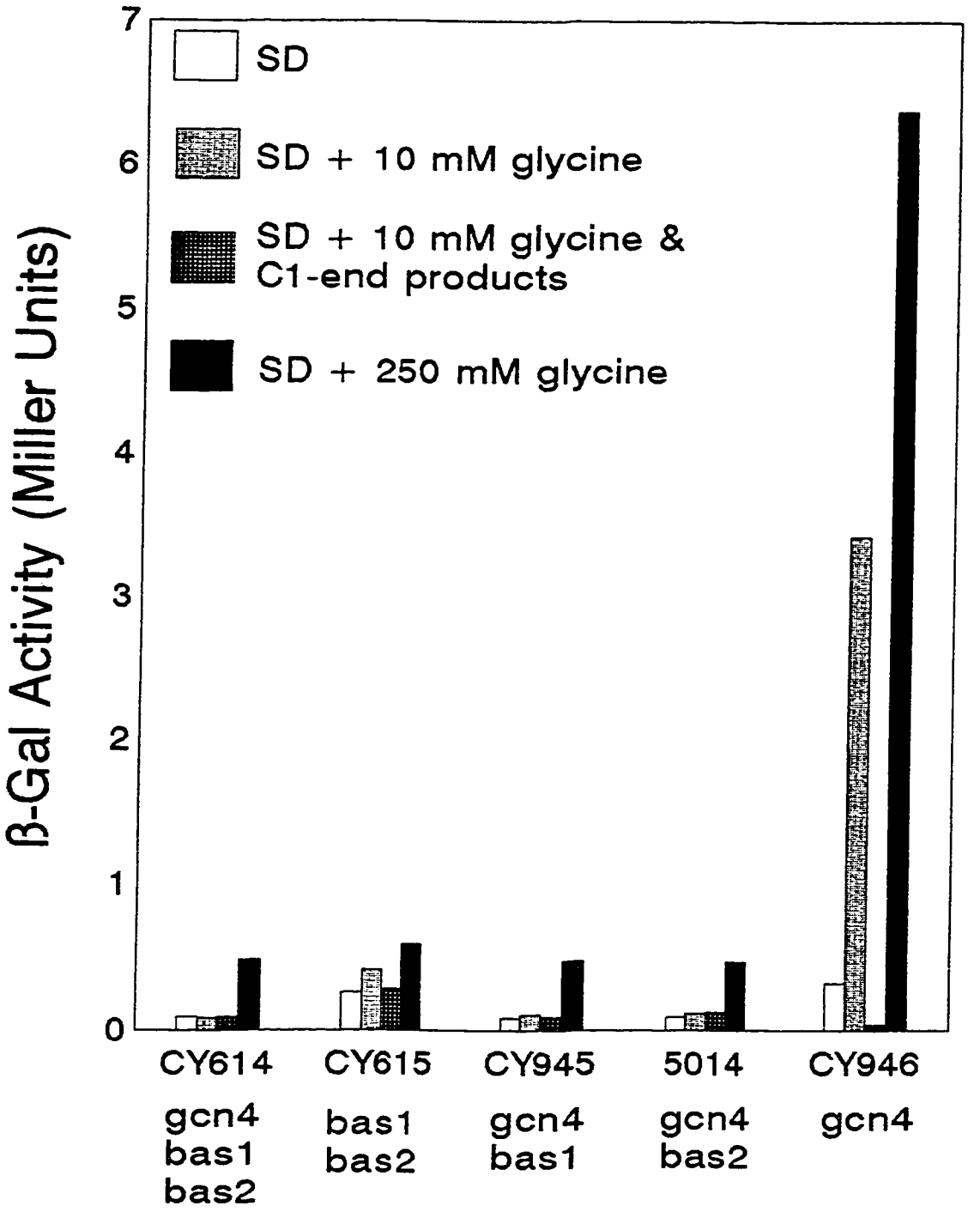

1 2 3 4 5 6 7

induced by glycine (38-fold) and reached levels similar to those obtained with the wild-type strain (Fig. 20). The lack of Bas1p and Bas2p in strain CY615 appears to have little or no effect on expression in the absence of glycine. Supplementing SD + 10 mM glycine with the C1-end products dramatically repressed expression (at least 50-fold) in strain CY946 to levels that were even lower than the low levels expressed by the other three strains in this medium. In SD the CY615 and CY946 transformants expressed 2-fold and 4-fold more ß-gal activity than the CY614, 5014 and CY945 transformants. These results support previous findings (Fig. 16) that both general control and the basal control contribute to *GCV3* expression in SD. That ß-gal expressed by CY615 (*bas1*, *bas2*) in SD was induced 1.5-fold by glycine suggests that expression is modestly induced by a Gcn4p-dependent mechanism in response to 10 mM glycine. The absence of pLNGCV3-211 induction by 10 mM glycine in CY614 (*gcn4*, *bas1*, and *bas2*), 5014 (*bas2* and *gcn4*), and CY945 (*gcn4*, and *bas1*) and the very modest induction (1.5 fold) in CY615 (*bas1*, and *bas2*) shows that the elevated levels of expression that are observed in the presence of glycine are very dependent on Bas1p/Bas2p and to a much lesser extent Gcn4p dependent.

Including the C1-end products with glycine repressed pLNGCV3-211 expression at least 80-fold in CY946 (*gcn4*) relative to the levels expression in the presence of only glycine. In contrast the C1-end products did not significantly repress expression in strains CY615, 5014, and CY945. Therefore, both Bas1p and Bas2p are required for the glycine-dependent induction of *GCV3*. Furthermore, the activation of *GCV3* expression by the Bas1p/Bas2p complex is inhibited when the growth medium is supplemented with the C1-end products.

Fig. 20.

Bas1p and Bas2p are important for *GCV3* regulation. ß-gal activity of the pLNGCV3-211 construct was measured under four different physiological conditions. SD, standard minimal media; SD + 10 mM glycine, SD + 10 mM glycine and C1-end products adenine (40 ug/ml), histidine (20 ug/ml) methionine (20 ug/ml), serine (5 mM) and formate (10 mM)). SD + 250 mM glycine as the sole nitrogen source. CY614, CY945, and CY5014 lack both general control and basal control. CY615 lacks basal control alone. CY946 lacks general control alone. Relevant genotype of these strains are shown in the bottom of the graph.


β-Gal Activity (Miller Units)
0
1
2
3
4
5
6
7
SD
SD + 10 mM glycine
SD + 10 mM glycine & C1-end products
SD + 250 mM glycine
CY614
gcn4 bas1 bas2
CY615
bas1 bas2
CY945
gcn4 bas1
5014
gcn4 bas2
CY946
gcn4

In SD + 250 mM glycine where glycine serves as the sole nitrogen source, an additional mechanism, can enhance expression in the absence of Gcn4p and Bas1p/Bas2p (Fig 20). That is, both in the absence of general and basal transcriptional control mechanism (strains CY614, 5014, and CY945) expression was about five fold higher when growth was in the presence of 250 mM glycine as the nitrogen source than it was under the other growth conditions.

C.19. Bas2p binds to the *GCV3* promoter.

To find out whether Bas2p interacted directly with the *GCV3* promoter, interactions between Bas2p and the *GCV3* promoter were analyzed by following DNA-protein binding *in vitro*. The PCR products depicted in Figure 9 were employed in this experiment. The binding conditions were identical to those used for Gcn4p binding (Fig. 19) except 3 μg of partially purified extract of *E.coli*-produced Bas2p replaced Gcn4p. The results show that Bas2p binds to the *GCV3* promoter region (Fig. 21, lane 7). They also show that the GCRE mutations that abolished Gcn4p binding do not alter Bas2p binding. Apparently, either the Bas2p binding site does not overlap the GCRE site (Fig. 21) or if it does the GCRE mutations included in these probes do not prevent Bas2p binding.

Since the binding sites for Bas2p and Gcn4p can overlap (Devlin et al., 1991; Tice-Baldwin et al., 1989), Bas2p may compete with Gcn4p for binding sites on the *GCV3* promoter. To test for this the PCR product GCRE-128 (Fig. 22) was employed. These studies showed that Bas2p does not prevent Gcn4p binding but rather both Bas2p and Gcn4p can bind the *GCV3* promoter simultaneously (Fig. 23).

Fig. 21.

GCRE mutations did not affect Bas2p binding. All the lanes contained 0.5 μg of γ-^{32}P-labelled probe. Lane 1 to 3 contains GCRE-128 labelled probe. Lane 4, lane 5, lane 6, and lane 7 contain GCRE mutant radiolabelled probes GCRE-128E, gcre-128F, gcre-128G and gcre-128A respectively (shown in Fig. 9). Lane 1 contains labelled probe alone. Lane 2 contained 3 μg of *E.coli* protein. Lane 3 to lane 7 contain 3 μg of Bas2p for each lane.

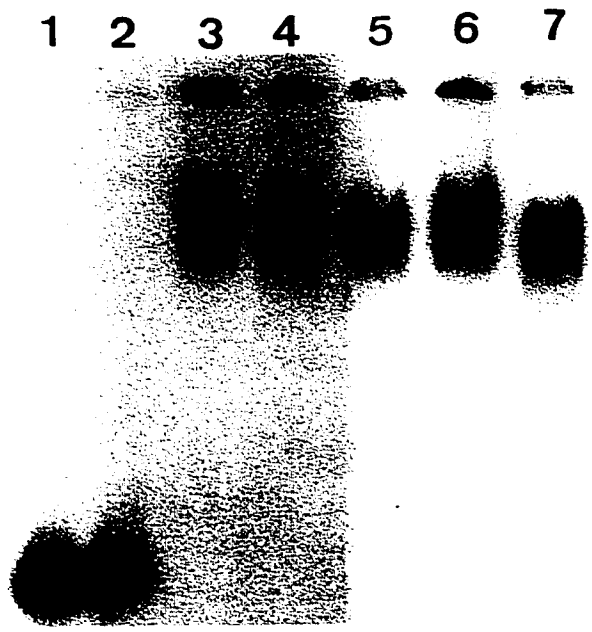

1 2 3 4 5 6 7

Fig. 22.

Bas2p does not titrate out the prebound Gcn4p-GCRE-128 complex. All the lanes contained 5 μg of γ-^{32}P-radiolabelled GCRE-128 probe. Lane 1 contains GCRE-128 labelled probe alone. Lane 2 contains 3.2 μg of Gcn4p alone. Lane 3 to 9 contained prebound Gcn4p-GCRE-128 complex. Varying amounts of Bas2p was supplemented starting from lane 3, 2 μg; lane 4, 4 μg; lane 5, 6 μg; lane 6, 8 μg; lane 7, 10 μg; lane 8, 12 μg; and lane 9, 14 μg. Lane 10 contains 3 μg of Bas2p alone. Different amounts of Bovine serum albumin (BSA) were added to keep the total protein equal.

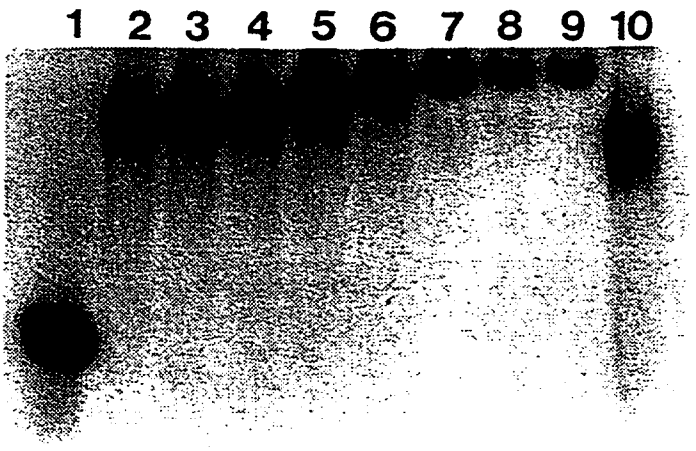

1 2 3 4 5 6 7 8 9 10

Fig. 23.

Gcn4p and Bas2p can bind to GCRE-128 simultaneously. All the lanes contain 5 μg of radiolabelled GCRE-128 probe. Lane 1 contains radiolabelled probe alone. Lane 2 contains 3.2 μg of *E.coli* protein. Lane 3, contains 3.5 μg of Gcn4p alone. Lane 4 contains 3 μg of Bas2p alone. Lane 5 contains both Gcn4p (3.2 μg) and Bas2p (3 μg).


1
2
3
4
5

C.20. Glycine is toxic to basal control mutants.

Although strains missing both the general and basal control systems are histidine auxotrophs (Tice-Baldwin et al., 1989; Devlin et al., 1991), each transcriptional control system can independently support *HIS4* expression at levels sufficient for normal growth. The expression studies presented above show that Bas1p/Bas2p and Gcn4p also regulate *GCV3* expression. To test whether Bas1p/Bas2p and/or Gcn4p were also important for glycine cleavage activity, I performed the following experiment. Yeast strains CY614 (*gcn4 bas1 bas2*), CY615 (*bas1 bas2*), CY945 (*bas1 gcn4*) and CY946 (*gcn4*) were grown on minimal medium plates supplemented with histidine, histidine and glycine, or histidine, glycine and formate. Histidine was included because strains defective in basal and general control are auxotrophic for histidine. The rational behind this experiment was to determine whether glycine is toxic for basal and/or general control mutants in a fashion similar to its toxicity for GCS mutants. I also wanted to test whether formate supplement could rescue the toxic effects of glycine (if it was toxic). All four strains grew with doubling time of 6 hr or less in minimal medium supplemented with histidine and with doubling times of less than 3.5 hr in medium with histidine, glycine and formate (Fig. 24 and Table 16). However, in the presence of histidine and glycine, the *gcn4* strain grew well (doubling time 4.5 hr) whereas the three other strains grew extremely poorly with doubling times of at least 18 hr (Fig. 24, and Table 16). Therefore, glycine severely inhibits the growth of the basal control mutants. These results also show that the inhibiting action that glycine exerts on basal control mutants can be reversed by formate supplement. These results show that basal control system is required to express GCS activity at levels that are sufficient to prevent

Fig. 24.

Glycine is toxic to mutants lacking basal transcriptional control. Both Bas1p and Bas2p are required for the normal growth in the presence of glycine. *Plate sector* A strain CY614, *Plate sector* B strain CY615, *Plate sector* C strain CY945, and *Plate sector* D strain CY946. All the plates contains histidine except Plate 4. Plate 1, SD; Plate 2, SD + 10 mM glycine; Plate 3, SD + 10 mM glycine and 10 mM formate; Plate 4, SD without histidine.

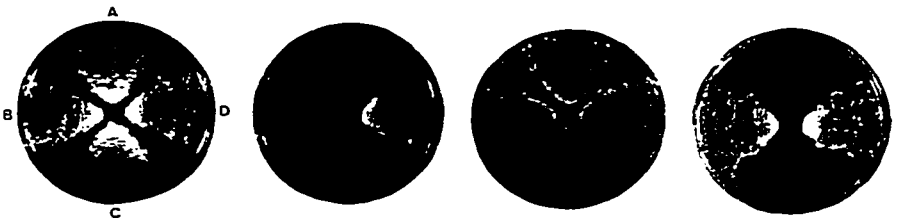

A
B
D
C

Table 16.

Growth rate of basal control and general control deficient mutants in SD with various supplements.

	Generation time (hr)		
Stain[a]	**+Histidine**	**+Histidine +Glycine**	**+Histidine +Glycine +Formate**
CY614 *bas1 bas2 gcn4*	6.0	21.5	3.5
CY615 *bas1 bas2*	5.0	18	3.5
CY945 *bas1 gcn4*	6.0	20.5	3.5
5014 *bas2 gcn4*	6.0	20.0	3.5
CY946 *gcn4*	3.5	4.5	3.0

[a] All the strains are basal control deficient except CY946, which is general control deficient.

All the strains grown with formate and histidine grew at 2.5 hr doubling time. Growth without formate but with other C1-end products (serine, histidine, adenine, and methionine) did not rescue glycine toxicity. Formate supplement alone did not compensate histidine auxotrophy of general control and basal control deficient strains.

glycine from inhibiting growth of wild-type yeast. A possible explanation for the poor growth of strains expressing low levels of GCS activity in the presence of glycine is that glycine inhibits synthesis of the C1-donor N^5 N^{10} MTHF.

C.21. Glycine uptake and degradation.

The above results suggest that glycine is toxic to *gcv3* strains (Table 5) and strains that lack basal transcriptional control (Table 16) because these strains express GCS activity at reduced levels. To directly test this possibility, glycine transport and breakdown assays were done on yeast strains 3634 (wt), 4404 (*gcv3*), CY614 (*bas1*, *bas2*, and *gcn4*), CY615 (*bas1*, and *bas2*), CY945 (*bas1* and *gcn4*), 5014 (*bas2* and *gnc4*) and CY946 (*gcn4*) (Table 17). The methods used for measuring transport and degradation are described in the materials and methods section.

The rates of glycine transport by all the strains used in this study were linear for at least 90 min (data not shown) and all the strains accumulated glycine at a rate of approximately 22 nmoles per $1x10^6$ cells in 30 min. Therefore the *gcv3*, *gcn4*, *bas1* and *bas2* mutations do not affect glycine transport.

To follow the fate of the accumulated glycine, degradation assays were performed. These results (Table 17) show that intracellular glycine decreased 6-fold when the wild-type and *gcn4* strains were incubated for 8 hr. In contrast, intracellular glycine levels in strains 4404, CY614, CY615, 5014 and CY945 hardly decreased during the 8 hr incubation.

Table 17.

Determination of intracellular glycine levels.

Strains	**OD_{600}**	**nmoles glycine/ $1x10^6$ cells after 30 min**	**nmoles of glycine/ $1X10^6$ cells after 8 hr**
3634 Wt	0.213	22.98	3.16
4404 *gcv3*	0.210	23.61	21.51
CY614 *gcn4 bas1 bas2*	0.208	21.24	18.92
CY615 *bas1 bas2*	0.217	21.81	15.34
CY945 *bas1 gcn4*	0.196	19.97	17.30
5014 *bas2 gcn4*	0.212	22.56	18.74
CY946 *gcn4*	0.206	22.24	3.93

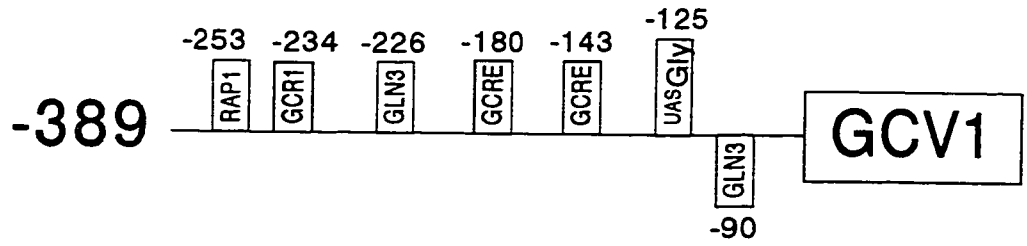
Therefore, the ability to degrade glycine by both the GCS defective (*gcv3* strain) and basal transcriptional control mutant strains were severely impaired. These results suggest that during growth in the presence of glycine, intracellular glycine accumulates to levels that can inhibit growth in the absence of a functional glycine cleavage system.

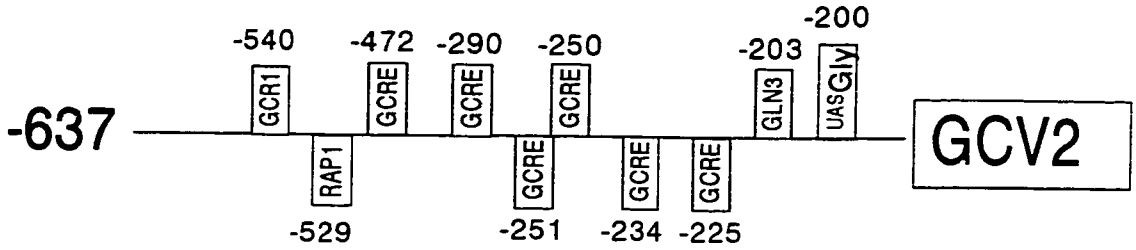
C.22. Potential upstream elements of *GCV1, GCV2* and *LPD1*.

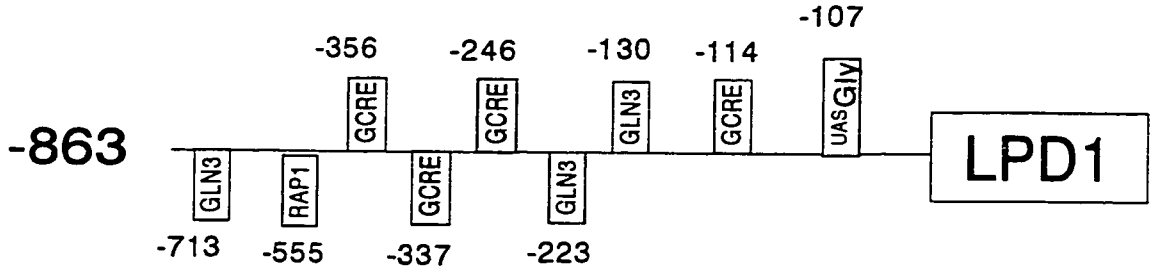
Genes coding for enzymes involved in the same biochemical pathway are often coordinately regulated. The results presented above show that *GCV3* is induced by glycine, repressed by the C1- end products, and regulated by the general amino acid control, basal control, nitrogen regulation, and glycolytic regulation. To assess whether the other glycine cleavage system genes (*GCV1, GCV2,* and *LPD1*) were regulated in a fashion similar to the regulation of *GCV3*, the promoter regions of *GCV1*, *GCV2*, and *LPD1* were searched for the presence of potential Gcn4p (GCRE), Gln3p, Gcr1p, and Rap1p binding elements. In addition the promoter regions of all four genes were searched for the presence of a common element that could be responsible for their induction by glycine (UAS_{Gly}). Potential Gcn4p, Gln3p, Gcr1p, and Rap1p binding sites were found in the upstream regions of all four genes (Fig. 25). A potential UAS_{Gly} was also identified in the upstream regions of all four genes. The presence of binding sites for these four transcriptional regulators under potential UAS_{Gly} upstream of all four GCS genes suggests that the *GCV1*, *GCV2,* and *LPD1* genes are regulated in a similar fashion to *GCV3*.

Fig. 25.

Potential upstream elements of *GCV1*, *GCV2*, and *LPD1* promoters. Boxes represents potential cis-acting elements found in the upstream regions of the genes. The boxes above the line represent the potential cis-acting elements of coding strand. Similarly the boxes below the line represent the potential cis-acting elements of non-coding strand. The numbers near the boxes indicate the positions of these cis-acting elements relative to the first base of the respective start codons. The promoter size (intergenic region) of these genes is indicated on the left. GLN3 stands for nitrogen dependent transcriptional control; GCR1 and RAP1 stand for glycolysis dependent regulation; UAS_{Gly} stands for substrate (glycine) dependent regulation; GCRE (general control recognition element), stands for general control and basal control.


-389
-253
-234
-226
-180
-143
-125
RAP1
GCR1
GLN3
GCRE
GCRE
UASGlv
GLN3
-90
GCV1


-637
-540
-472
-290
-250
-203
-200
GCR1
GCRE
GCRE
GCRE
GLN3
UASGlv
RAP1
GCRE
GCRE
GCRE
-529
-251
-234
-225
GCV2


-863
-356
-246
-130
-114
-107
GCRE
GCRE
GLN3
GCRE
UASGlv
GLN3
RAP1
GCRE
GLN3
-713
-555
-337
-223
LPD1

DISCUSSION

D.1. Molecular characterization of *GCV3* and single carbon donor pathways of *S.cerevisiae.*

There are three ways to synthesize the C1-donor 5,10-MTHF in *S.cerevisiae* (Nagarajan and Storms, 1997). These routes, which obtain C1-units from serine, glycine and formate, are catalysed by glycine hydroxymethyltransferase, the glycine cleavage system and C1-tetrahydrofolate synthase, respectively. I conclude, however, that on minimal medium serine derived from 3-phosphoglycerate provides the vast majority and possibly all C1-units. This conclusion is based on the following results: Glycine supplement rescues the very poor growth phenotype of a *ser1* mutant. Therefore, *ser1* strains growing on minimal medium cannot synthesize enough glycine to meet cellular requirements for glycine, serine and 5,10-MTHF. Furthermore, since supplement with both glycine and formate is required to rescue a *ser1 gcv3::URA3* strain, effectively all cellular formate is derived from 5,10-MTHF via C1-tetrahydrofolate synthase.

The relatively slow growth of a *gcv3::URA3* strain on minimal medium (Table 5) was unexpected, since it should be possible to synthesize serine, glycine and the C1-donor 5,10-MTHF without a functional glycine cleavage system (Fig. 2). One explanation for the reduced growth rate of *gcv3::URA3* strains is that the amount of 5,10-MTHF synthesized is rate limiting for growth. Rate limiting amounts of 5,10-MTHF might arise in a *gcv3::URA3* strain, because without a functional glycine cleavage system only one molecule of 5,10-MTHF can be produced for each serine consumed and/or because the accumulation of any

excess glycine generated by the action of glycine hydroxymethyltransferase inhibits the production of 5,10-MTHF. Two observations support the notion that the synthesis of 5,10-MTHF is rate limiting in a *gcv3* mutant. First, formate supplement enables the *gcv3::URA3* strain to grow at the normal rate (Table 5). Second, extracellular glycine severely inhibits growth of the *gcv3* mutant but has little effect on the wild-type (Table 5).

Glycine cleavage assays have shown that cells grown with glycine as the sole nitrogen source had significantly enhanced levels of glycine cleavage system activity (Sinclair et al., 1996). I found that this increased activity is controlled, at least in large part, by a transcriptional control mechanism, because a *GCV3-LacZ* reporter gene was strongly induced in response to extracellular glycine.

In light of the observed regulation of *GCV3* and glycine cleavage system activity in response to extracellular glycine, I examined the *LPD1*, *GCV1*, *GCV2* and *GCV3* genes for consensus elements that might coordinate their induction by glycine. One potential element with the consensus 5'-RCCTCGM-3' is present in the upstream regions of the four glycine cleavage system subunit encoding genes.

That *GCV3* expression is not induced by glycine when the C1-end products, serine, methionine, histidine and adenine are included in minimal medium suggests that the glycine cleavage system is regulated to meet cellular demand for C1-units. Transcriptional control of *GCV3* in response to demand for C1-units could occur via a mechanism that responds to intracellular concentrations of 5,10-MTHF, one or more of the other C1-donors or even the C1-end products. A second possibility is that the intracellular concentration of glycine regulates *GCV3* transcription. For example, in the absence of extracellular C1-end products

essentially all C1-units must be derived from serine. Two C1-units can be derived from each molecule of serine. One C1-unit is added to THF during the conversion of serine into glycine by the enzyme glycine hydroxymethyltransferase. Glycine generated in excess of that needed for cellular metabolism, for example protein, purine and heme biosynthesis, can also be used as a source of C1-units in the series of reactions catalysed by the glycine cleavage system. Therefore, *GCV3* transcription might be regulated by the intracellular glycine concentration to ensure the efficient utilization of serine and/or to prevent the potentially toxic effects of glycine.

My results also show that *GCV3* is subject to transcriptional regulation by the general amino acid control system. It is, however, less obvious as to why *GCV3* would be subject to general control. Perhaps, since C1-units derived via the glycine cleavage system can be used in the synthesis of amino acids like serine, methionine, and formyl methionine, *GCV3* is subject to general control for the same reason that so many other amino acid biosynthetic genes are subject to general amino acid control (Hinnebusch, 1992). Consistent with the glycine cleavage system being subject to general control, the upstream regions adjacent *LPD1*, *GCV1*, and *GCV2* also contain copies of the GCRE element that is utilized by Gcn4p.

D.2. Multiple mechanisms regulate *GCV3* transcription.

The results presented in this study also demonstrate that *GCV3* is regulated by the general amino acid control system, the nitrogen regulation system, basal control system, a glycine specific response element (UAS_{Gly}) and the Gcr1p-dependent system that regulates glycolytic gene expression (Fig. 25). The complexity of *GCV3* expression may seem

unnecessary; however, each of the products of glycine cleavage (NH_3, 5,10-MTHF and NADH) plays a major role in distinctly different aspects of cellular metabolism. For example, the NH_3 generated by glycine cleavage can be used directly for the biosynthesis of glutamate and then glutamine. These two amino acids are subsequently used for the biosynthesis of other nitrogen containing compounds. 5,10-MTHF ultimately contributes to the synthesis of the other one-carbon carrier derivatives of THF, S-adenosylmethionine and formate. These C1-unit donors are required for the biosynthesis of thymidylate, purines, methionine, S-adenosylmethionine, and histidine. Formate is required for the shuttling of C1-units between cellular compartments (Pasternack et al., 1994). The glycine cleavage system also provides the reduced electron carrier NADH for oxidative phosphorylation. In light of the distinct metabolic functions performed by these products it is not unexpected that glycine cleavage regulation would respond to its substrate, the availability of alternative nitrogen sources and general amino acid control.

The expression studies presented show that general amino acid control of *GCV3* is dependent upon one or more of three GCRE elements and Gcn4p. General control of *GCV3* regulation may occur because 5,10-MTHF is required for the biosynthesis of a number of amino acids including histidine, an amino acid whose biosynthesis is subject to general amino acid control.

It is also not surprising that glycine induced *GCV3-lacZ* expression, since many examples exist where the genes for the catabolism of a nutrient are induced either directly or indirectly by that nutrient. The upstream region that responds to glycine supplement overlaps the 17-bp region immediately downstream of base pair -128. The element 5'-

RCCTCGM-3', which is found in this region and is also present in the upstream regions of the other glycine cleavage genes, is a candidate for the cis-acting sequence that is used to induce gene expression in response to glycine. Induction of expression by glycine has only a modest effect, since deleting the region harbouring this element reduced expression about 4-fold.

Expression is also regulated in response to the nitrogen source that is present in the medium. Deleting the region with a UAS_{GATAA} sequence affects expression when the nitrogen source is glycine or ammonia (Table 11). Interestingly, deleting this region caused a 4-fold reduction in expression when glycine is the nitrogen source but has the opposite effect (a 5-fold increase) when ammonia is used. Therefore, deleting this region greatly reduced the magnitude of the difference in expression levels observed between growth with ammonia and glycine (54-fold (Table 6 and 9) versus 4-fold (Table 9) when this region is deleted). These effects can be explained because both a transcriptional activator (Gln3p) and a transcriptional repressor (Dal80p) bind to UAS_{GATAA} sites.

Gcr1p apparently functions in conjunction with Rap1p to enhance glycolytic gene expression. Perhaps the involvement of Gcr1p relates to the close metabolic association between glycine cleavage and glycolysis, where phosphoenolpyruvate can be used for glycine biosynthesis via serine as an intermediate. Indeed, under optimal growth conditions with glucose as the carbon source, excess glycine would be reported to be produced during the SHMT-dependent production of C1-units from serine. The presence of an active glycine cleavage system would ensure that glycine did not accumulate.

D.3. *In vivo* and *in vitro* assessment of GCRE and basal control.

In vivo and *in vitro* analysis demonstrated that the three GCREs (Fig. 3, and Fig. 6) located at -107 bp, -87 bp and -71 bp are very important for the transcriptional regulation of *GCV3*. Mutating all three GCREs abolished expression from pLNGCV3-128 (construct pLNGCV3-128A, Fig. 17, Table 13, 14, and 15). The abolished expression of pLNGCV3-128A could be due to the inability of Gcn4p to bind to the promoter of this construct (Fig. 19). This result might mean that one or more of these GCREs are absolutely required to maintain Gcn4p-mediated *GCV3* regulation. Interestingly, the pLNGCV3-128 construct was expressed at significant levels in a *gcn4* strain (3646). The abolished expression of pLNGCV3-128A in the wild-type strain (3634) suggests that expression of the pLNGCV3-128 construct is totally dependent upon the GCREs. Therefore the mutations introduced to the GCREs affect at least two transcriptional control mechanisms: one is Gcn4p-dependent general control, the other is general control independent.

I speculated that *GCV3* regulation might be controlled by one or more of several proteins with binding sites that can overlap general control elements. Positive regulators with binding sites that can overlap general control elements are Bas1p, Bas2p and Yap1p (Mösch et al., 1990; Devlin et al., 1991). My results show that Bas1p and Bas2p cooperatively mediate the basal control of *GCV3*. *GCV3-lacZ* expression in isogenic wild-type and *yap1* strains was similar indicating that under the condition tested Yap1p did not play a significant regulatory role (data not shown).

The *in vitro* Bas2p binding studies demonstrated that Bas2p binding is not affected by mutating the GCRE elements (Fig. 21). Knowing that the Bas1p is also required for

basal control (Fig. 20), I speculated that the mutations introduced into the GCREs might be preventing Bas1p binding. Hovring et al., (1994) determined the consensus sequence required for Bas1p binding to be 5'-AGTGACTCCGGTA-3'. This sequence contains a GCRE sequence (5'-TGACTC-3') demonstrated to be important for Bas1p binding. Therefore, it is not unexpected that the mutations introduced into the GCREs would affect Bas1p binding and the basal control of pLNGCV3-128A expression (Table 13, 14, and 15). Since basal control requires both Bas1p and Bas2p for transcriptional activation. Initially the mutations introduced in the GCREs were designed to eradicate general control (Gcn4p binding). My results demonstrate that pLNGCV3-128 expression is GCRE-dependent and that mutations introduced into the GCREs affect both general and basal control.

The consensus sequence (5'-GGTAAATTAAGTTAATTAATTG-3') reported for Bas2p binding (Svetlov and Cooper, 1995) is not apparent on either strand of the *GCV3* promoter, however the Bas2p binding site in the *GCV3* promoter remains unknown. Therefore, the role of GCREs on basal control in *GCV3* deserves further research. The Bas1p and Bas2p binding sites in the *HIS4* promoter have been shown to overlap the GCREs (Tice-Baldwin, et al 1989; Devlin et al., 1991). In the *TRP4* gene, one of the GCREs shown to be important for general control also serves as a target for Bas2p binding. During simultaneous amino acid and phosphate starvation these two proteins compete for this site and Bas2p appears to efficiently bind and suppress general control (Braus et al., 1989).

To test the nitrogen regulation and glycolytic regulation of *GCV3* in the absence of both the general and basal control systems, the expression of pLNGCV3-gcre-156 and pLNGCV3-gcre-176 were examined. Inclusion of the nitrogen regulation upstream

information (construct pLNGCV3-gcre-156) did not support ß-gal expression in SD where the rich nitrogen source ammonia is present. However, the same construct maintained a significant level of expression when glycine served as the sole nitrogen source (Fig. 8). The longer promoter region, which includes the glycolytic regulatory information, expressed modest ß-gal activity in SD and resulted in slightly increased expression in SD + 250 mM glycine (Fig. 8). These results demonstrate that the upstream nitrogen and glycolytic regulatory information can support expression in the absence of general and basal control.

The expression of the pLNGCV3-128 construct resembles the expression of *HIS4*, because its expression is solely dependent upon general and basal transcriptional control for expression. The absence of these two transcriptional controls abolishes the expression of both *HIS3* and the *GCV3*-128 fusion gene. The regulation of the full *GCV3* promoter construct (pLNGCV3-211) is different than *HIS4* because nitrogen regulation, glycolytic regulation and a glycine response element also contribute to its regulation (Fig. 8).

It is interesting to note that the mutation introduced to the GCRE located at -107 bp resulted in increased expression in all the strains tested regardless of the growth conditions. This suggests that this upstream GCRE behaves like an URS (construct pLNGCV3-128B) (Table 13, 14, and 15). A possible candidate for the repressor protein that binds to this site is Acr1p. Acr1p has been shown to negatively regulate *HIS3* by binding to a GCRE like element (Vincent and Struhl, 1992). The role of Acr1p on *GCV3* regulation has not yet been tested. The GCRE located at -107 bp also appears to repress the general control response (Table 15) during growth with 250 mM glycine as the sole nitrogen source. That is, when this element is intact *GCV3-lacZ* expression is not subject to general amino acid control

under these growth conditions. However, when this site is mutated expression is subject to general amino acid control. Therefore, the GCRE located at -107 bp can also apparently repress the general control response. This possibility is consistent with the absence of a detectable general control response when pLNGCV3-128C, and pLNGCV3-128D constructs (where this site is intact) are expressed in the presence of glycine as the sole nitrogen source (Table 15).

In vitro studies indicated that Gcn4p binds to the GCRE located at -107 bp (Fig. 19, lane 4). However, mutating this GCRE did not reduce the effect of the general control system on *GCV3* expression. To date, there is no evidence in the literature that the Gcn4p could bind and repress gene expression. Understanding the the regulatory role of this site requires further analysis. DNA binding studies using Acr1p, Gcn4p and the various GCRE mutants should provide important insights.

The mutations introduced to the GCRE located at -87 bp (pLNGCV3-128C construct) demonstrate that this GCRE is a target of more than one protein, because mutating this site affected both general and basal control particularly in the presence of glycine (Table 13, 14, and 15). The mutations introduced to the GCRE located at -71 bp (construct pLNGCV3-128D) did not affect the general control response in SD (Table 13), therefore Gcn4p is not targeted to this site under this physiological condition. However, in the presence of glycine (both SD + 10 mM glycine and SD + 250 mM glycine) these mutations not only affected the general control response but also severely reduced expression in the presence of glycine. Therefore, this GCRE may also be included within the binding site for more than regulatory protein.

The evidence presented in section C.16 implies that both GCREs located at -87 bp and -71 bp are critically important for both general and basal control especially in the presence of extracellular glycine. It is also possible that these two GCREs may synergistically interact through their trans-acting factors to cause induced expression in the presence of glycine. Synergistic interaction of GCREs through their trans-acting factors have been shown in the *TRP4* (Mösch et al., 1990).

The results presented in Figure 20 demonstrate that the glycine-dependent expression (SD + 10 mM glycine) is largely mediated by the Bas1p- and Bas2p-dependent basal control. Under similar conditions the Gcn4p-dependent general control system appears to very poorly support glycine-dependent *GCV3-lacZ* expression. In addition to glycine-dependent expression, the basal control system appears to be responsible for the repressed expression in response to presence of C1-end products. In contrast, the general control system does not respond to presence of C1-end products. When glycine served as a nitrogen source an additional mechanism (probably nitrogen regulation) joins the glycine-dependent expression. Strains lacking basal and general control systems (CY614, CY615, CY945, and 5014) do not respond to the presence of extracellular glycine or glycine and C1-end product. However, the same strains exhibit glycine-dependent expression only when glycine served as the sole nitrogen source (Fig. 20). Overall, Bas1p and Bas2p are very important for *GCV3* regulation.

The study of *GCV3-lacZ* showed that glycine-dependent expression requires at least three transcriptional mechanisms (Fig. 20, Fig. 18, and Table 11). They are basal control, nitrogen regulation, and UAS_{Gly} dependent regulation. Basal control plays a major role in

the glycine-dependent expression of *GCV3* (Fig. 20). UAS_{Gly} alone can not mediate glycine-dependent expression (Fig. 18). Apparently transcriptional activation by UAS_{Gly} is dependent upon the basal control (Fig. 18) and basal control requires one or more of the upstream GCREs.

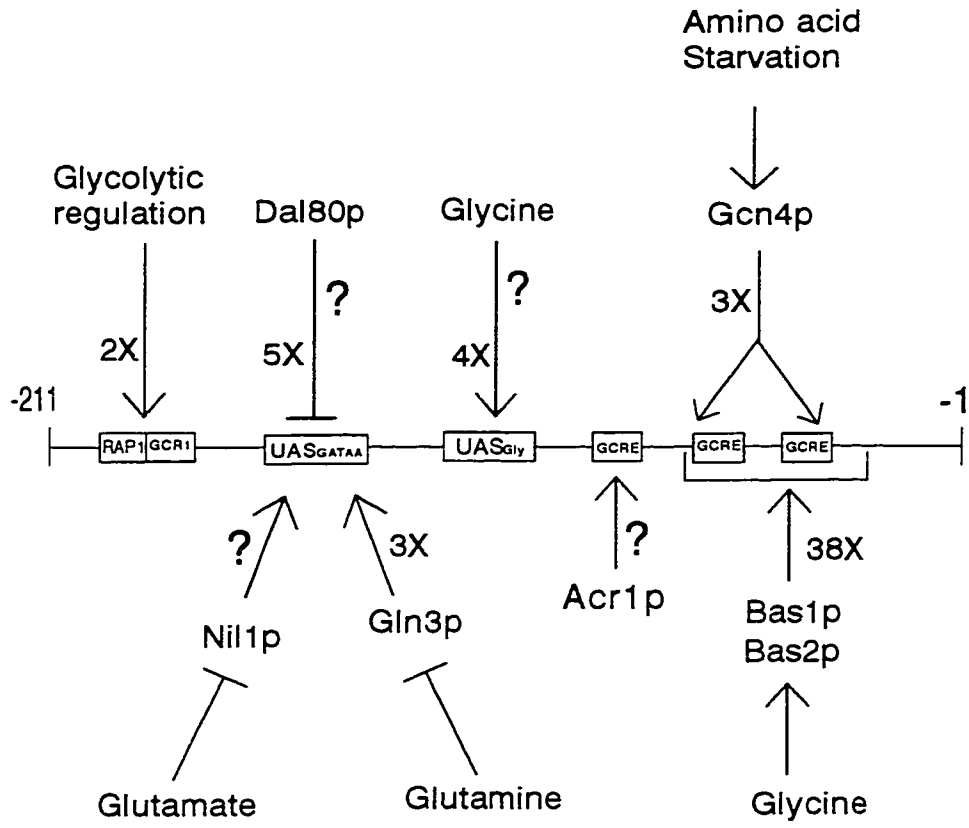
The results presented in Figure 24 and Table 16 demonstrate that extracellular glycine inhibits the growth of strains lacking basal control (CY614, CY945, and 5014). These strains grew much slower (approximately 20 hr versus 5 hr) than strain CY946 which has an intact basal control system. I believe that the inability of these cells to degrade intracellular glycine reduces 5, 10-MTHF synthesis which leads to growth limiting amounts of one or more of the C1-end products. Consistent with this explanation supplement with formate which can be used as a source of C1 units for 5, 10-MTHFsynthesis effectively reverses this poor growth phenotype (Fig. 2).

Overall *GCV3* transcription appears to be regulated by at least five transcriptional activators (Fig. 26). Depending upon the physiological conditions, these activators result in *GCV3* being subject to general control, basal control, nitrogen regulation, glycolytic regulation, and glycine-dependent regulation. In addition, there are at least two negative regulators of *GCV3* transcription. One appears to be playing a negative role regardless of the physiological condition, which negatively regulates expression in the presence of rich nitrogen sources, apparently involves Gln3p.

Determining the importance of the other regulators of nitrogen metabolism like Nil1p and Dal80p, should be the subject of further studies. Similar to *GCV3*, promoters of

Fig. 26.

Model showing all the known and potential transcriptional mechanisms of the *GCV3* gene. Boxes on the horizontal line indicate potential cis-acting factors found between -1 to -211 bp of the *GCV3* promoter. Arrows represents positive effect on expression. Blunted arrows represent negative effect on expression. Magnitude of positive and negative effect on expression are indicated with the numbers near the arrows. A question mark indicates that the effect on expression has not been verified experimentally.


Amino acid
Starvation
Glycolytic
regulation
Dal80p
Glycine
Gcn4p
2X
5X
?
4X
?
3X
-211
-1
RAP1
GCR1
UASGATAA
UASGly
GCRE
GCRE
GCRE
?
3X
?
38X
Nil1p
Gln3p
Acr1p
Bas1p
Bas2p
Glutamate
Glutamine
Glycine

the other members of the glycine cleavage system genes *(GCV1, GCV2, and LPD1)* also harboured consensus sequences required for nitrogen regulation, general control, basal control, and glycolytic regulation (Fig. 25). Therefore, it appears that all four glycine cleavage system genes are coordinately regulated in a fashion similar to *GCV3*. The sets of isogenic strains constructed as part of my *GCV3* studies can now be used characterize *GCV1*, *GCV2*, and *LPD1* expression.

REFERENCES

-Agre, P., Johnson, P.F. and McKnight, S.L. (1989). Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. Science 246:922-926.

-Alani, E., Cao, L. and Klekner, N. (1987). A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. Genetics 116:541-545.

-Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. 215:403-410.

-Appling, D.R. and Rabinowitz, J.C (1985). Regulation of expression of the *ADE3* gene for yeast C1- tetrahydrofolate synthase, a trifunctional enzyme involved in one-carbon metabolism. J. Biol. Chem. 260:1248-1256.

-Arndt, K. and Fink, G.R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5'TGACTC 3'sequences. Proc. Natl. Acad. Sci. USA. 83:8516-8520.

-Arndt, K.T., Styles, C. and Fink, G.R. (1987). Multiple global regulators control *HIS4* transcription in yeast. Science 237:874-880.

-Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G.,Smith, J.A. and Struhl, K. (1989). Current Protocols in Molecular Biology (New York: Wiley Interscience and Green Publishing Associates).

-Baker, H.V. (1986). Glycolytic gene expression in *Saccharomyces cerevisiae*:nucleotide sequence of *GCR1*, null mutants, and evidence for expression. Mol. Cell. Biol. 11:3774-3784.

-Baker, H.V. (1991). *GCR1* of *Saccharomyces cerevisiae* encodes a DNA binding protein whose binding is abolished by mutations in the CTTCC sequence motif. Proc. Natl. Acad. Sci. USA. 88:9443-9447.

-Barlowe, C.K. and Appling, D.R. (1990). Molecular genetic analysis of *Saccharomyces cerevisiae* C1-tetrahydrofolate synthase mutants reveals a noncatalytic function of the *ADE3* gene product and an additional folate-dependent enzyme. Mol. Cell. Biol. 10:5679-5687.

-Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acid Research 21:3329-30.

-Bitter, G.A., Chang, K.K. and Egan, K.M. (1991). A multi-component upstream activation sequence of the *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase gene promoter. Mol. Gen .Genet. 231:22-32.

-Blinder, D. and Magasanik, B. (1995). Recognition of nitrogen-responsive upstream activation sequences of *Saccharomyces cerevisiae* by the product of the *GLN3* gene. J. Bacteriol. 177:4190-4193.

-Blinder, D., Coschigano, P.W. and Magasanik, B. (1996). Interaction of the GATA factor Gln3p with the nitrogen regulator Ure2p in *Saccharomyces cerevisiae*. J. Bacteriol. 178:4734-4736.

-Boeke, J., Lacroute, F. and Fink, G. Z . (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. 197:345-346.

-Bourguignon, J., Neuburger, M. and Douce, R. (1988). Resolution and characterization of the glycine-cleavage reaction in pea leaf mitochondria. Properties of the forward reaction catalysed by glycine decarboxylase and serine hydroxymethyltransferase. Biochem. J. 255:169-178.

-Bradford, M.M. (1976). A rapid sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

-Braus, G., Mosch, H.U., Vogel, K., Hinnen, A. and Hutter, R. (1989). Interpathway regulation of the *TRP4* gene of yeast. EMBO J. 8:939-945.

-Brazas, R.M. and Stillman, D.J. (1993). The Swi5 zinc-finger and Grf10 homeodomain proteins bind DNA cooperatively at the yeast *HO* promoter. Proc. Natl. Acad .Sci. USA. 90:11237-11241.

-Bussey, H., Kaback, D.B., Zhong, W., Vo, D.T., Clark, M.W., Fortin, N., Hall, J., Ouellette, B.F.F., Keng, T., Barton, A.B., Su, Y., Davies, C.J. and Storms, R.K. (1995). The nucleotide sequence of chromosome I from *Saccharomyces cerevisiae*. Proc. Natl. Acad.Sci. USA. 92:3809-3813.

-Bysanti, N., Daugherty, J.R. and Cooper, T.G. (1991). Saturation mutagenesis of the UAS_{NTR} (GATAA)responsible for nitrogen catabolite repression-sensitive transcriptional activation of the allantoin pathway genes in *Saccharomyces cerevisiae*. J. Bacteriol. 173:4977-4982.

-Casadaban, M.J., Martinez-Arias, A., Shapira, S.K. and Chou, J. (1983). ß-galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. 100:293-308

-Chambers, A., Tsang, J.S., Stanway, C., Kingsman, A.J. and Kingsman, S.M. (1989). Transcriptional control of the *Saccharomyces cerevisiae PGK* gene by *RAP1*. Mol. Cell. Biol. 9:5516-5524

-Chambers, A., Packham, E.A. and Graham, I.R. (1995). Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*). Curr. Genet. 29:1-9.

-Clifton, D. and Fraenkel, D.G. (1981). The *gcr* (glycolysis regulation) mutation of *Saccharomyces cerevisiae*. J. Biol. Chem. 256:13074-13078.

-Coffman, J.A., el Berry, H.M. and Cooper, T.G. (1994). The URE2 protein regulates nitrogen catabolic gene expression through the GATAA-containing UAS_{NTR} element in *Saccharomyces cerevisiae*. J. Bacteriol. 176:7476-7483.

-Coffman, J.A., Rai, R., Cunningham, T., Svetlov, V. and Cooper, T.G. (1996). Gat1p, a GATA family protein whose production is sensitive to nitrogen catabolite repression, participates in transcriptional activation of nitrogen-catabolic genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 16:847-858

-Courchesne, W.E. and Magasanik, B., (1988) Regulation of nitrogen assimilation in *Saccharomyces cerevisiae*: roles of the *URE2* and *GLN3* genes. J. Bacteriol. 170: 708-713.

-Cunningham, T.S. and Cooper, T.G. (1991). Expression of the *DAL80* gene, whose product is homologous to the GATA factors and is a negative regulator of multiple nitrogen catabolic genes in *Saccharomyces cerevisiae*, is sensitive to nitrogen catabolite repression. Mol. Cell. Biol. 11:6205-6215.

-Cunningham, T.S. and Cooper, T.G. (1992). Expression of the *DAL80* gene, whose product is homologous to the GATA factors and is a negative regulator of multiple nitrogen catabolic genes in *Saccharomyces cerevisiae*, is sensitive to nitrogen catabolite repression. Mol. Cell. Biol. 12:2454.

-Cunningham, T.S. and Cooper, T.G. (1993). The *Saccharomyces cerevisiae* DAL80 repressor protein binds to multiple copies of GATAA-containing sequences (URS_{GATA}). J. Bacteriol. 175:5851-5861.

-Cunningham, T.S., Dorrington, R.A. and Cooper, T.G. (1994). The UGA4 UAS_{NTR} site required for GLN3-dependent transcriptional activation also mediates DAL80-responsive regulation and DAL80 protein binding in *Saccharomyces cerevisiae*. J. Bacteriol. 176:4718-4725.

-Cunningham, T.S., Svetlov, V.V., Rai. R., Smart, W. and Cooper, T.G. (1996). G1n3p is capable of binding to UAS_{NTR} elements and activating transcription in *Saccharomyces cerevisiae*. J. Bacteriol. 178:3470-3479.

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

-Devlin, C., Tice-Baldwin, K., Shore, D. and Arndt, K.T. (1991). RAP1 is required for BAS1/BAS2- and GCN4-dependent transcription of the yeast *HIS4* gene. Mol. Cell. Biol. 11:3642-3651.

-Drazinic, C.M., Smerage, J.B., Lopez, M.C. and Baker, H.V. (1996). Activation mechanism of the multifunctional transcription factor repressor-activator protein 1 (Rap1p). Mol. Cell. Biol. 16:3187-3196.

-Fujiwara, K., Okamura-Ikeda, K. and Motokawa, Y. (1990). cDNA sequence, *in vitro* synthesis, and intramitochondrial lipoylation of H-protein of the glycine cleavage system. J. Biol. Chem. 265:17463-17467.

-Fujiwara, K., Okamura-Ikeda, K., Hayasaka, K. and Motokawa, Y. (1991). The primary structure of human H-protein of the glycine cleavage system deduced by cDNA cloning. Biochem. Biophys. Res. commun. 176:711-716.

-Geitz, D., Jean, A.St., Woods, R.A. and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acid Research 20:1425.

-Guarente, L. (1987). Regulatory proteins in yeast. Annu. Rev. Genet. 21:425-452.

-Guarente, L. (1995). Transcriptional coactivators in yeast and beyond. Trends. Biochem..Sci. 20:517-521.

-Hamosh, A, Johnston, M.V. and Valle, D. (1995). Non-ketotic hyperglycinemia. In Scriver CR, Beaudet, A.L., Sly, W.S., Valle, D, eds. The Metabolic and Molecular Bases of Inherited Disease, 7th edn. New York; McGraw-Hill, 1337-1348.

-Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557-580.

-Hatefi, Y. (1985). The mitochondrial electron transport and oxidative phosphorylation system. Annu. Rev. Biochem. 54:1015-1069.

-Herschbach, B.M. and Johnson, A.D. (1993). Transcriptional repression in eukaryotes. Annu. Rev. Cell. Biol. 9:479-509.

-Higgins, D.G. and Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237-244.

-Hill, D.E., Hope, I.A., Macke, J.P. and Struhl, K. (1986). Saturation mutagenesis of the yeast his3 regulatory site: requirements for transcriptional induction and for binding by GCN4 activator protein. Science 234:451-457.

-Hinnebusch, A.G. (1988). Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. Microbiological Reviews 52:248-273.

-Hinnebusch, A.G. (1990). Transcriptional and translational regulation of gene expression in the general control of amino-acid biosynthesis in *Saccharomyces cerevisiae*. Progress in Nucleic Acid Research and Molecular Biology. 38:195-240.

-Hinnebusch, A.G. (1992). General and pathway-specific regulatory mechanisms controlling the synthesis of amino acids biosynthetic enzymes in *Saccharomyces cerevisiae*. In The Molecular Biology of the yeast *Saccharomyces cerevisiae*: Metabolism and gene expression. Strathern, J., Jones, E.W., and Broach, J.R (eds). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 319-414.

-Hinnebusch, A.G. and Fink, G.R. (1983). Repeated DNA sequence upsteam from *HIS1* also occur at several other co-regulated genes in *Saccharomyces cerevisiae*. J .Biol. Chem. 258:5238-5247.

-Holland, M.J., Yokoi, T., Holland, J.P., Myambo, K. and Innis. M.A. (1987). The *GCR1* gene encodes a positive transcriptional regulator of the enolase and glyceraldehyde -3-phosphate dehydrogenase gene families in *Saccharomyces cerevisiae*. Mol. Cell .Biol. 7:813-820.

-Hope, I.A. and Struhl, K. (1985). GCN4 protein, synthesized *in vitro*, binds HIS3 regulatory sequences: implications for general control of amino acid biosynthetic genesin yeast. Cell 43:177-188.

-Hope, I.A., Mahadevan, S. and Struhl, K. (1988). Structural and functional characterisation of the short acidic transcriptional activation region of yeast GCN4 protein. Nature 333:635-640.

-Hovring, I., Bostad, A., Ording, E., Myrset, A.H. and Gabrielsen, O.S. (1994). DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. J. Biol. Chem. 269:17663 -17669.

-Huie, M.A., Scott, E.W., Drazinic, C.M., Lopez, M.C., Hornstra, I.K., Yang, T.P. and Baker, H.V. (1992). Characterization of the DNA-binding activity of GCR1: *in vivo* evidence for two GCR1-binding sites in the upstream activating sequence of *TPI* of *Saccharomyces cerevisiae*. Mol. Cell. Biol. 12:2690-2700

-Johnston, M. and Carlson, A.G. (1992). Regulation of carbon and phosphate uitilization. In The Molecular Biology of the yeast *Saccharomyces cerevisiae*: Metabolism and gene expression. Strathern, J., Jones, E.W., and Broach, J.R (eds). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 193-281.

-Kim, Y. and Oliver, D.J. (1990). Molecular cloning, transcriptional characterization and sequencing of cDNA encoding the H-protein of the mitochondrial glycine decarboxylase complex in peas. J. Biol. Chem. 265:848-853.

-Knappe, J. (1987). *Escherichia coli* and *Salmonella typhimurium* in Cellular and Molecular Biology (Neidhardt, C.F., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. eds), vol. 1, pp. 151-155.

-Kunzler, M., Springer, C. and Braus, G.H. (1995). Activation and repression of the yeast *ARO3* gene by global transcription factors. Mol. Microbiol. 1:167-178.

-Kunzler, M., Springer, C. and Braus, G.H. (1996). The transcriptional apparatus required for mRNA encoding genes in the yeast *Saccharomyces cerevisiae* emerges from a jigsaw puzzle of transcription factors. FEMS. Microbiol. Rev. 2:117-136.

-Macherel, D., Lebrun.M., Gagnon, J., Neuburger, M. and Douce, R. (1990). Primary stucture and expression of H-protein, a component of the glycine cleavage system of pea leaf mitochondria. Biochem. J. 268:783-789.

-Magasanik, B. (1992). Regulation of nitrogen utilization. In The Molecular Biology of the yeast *Saccharomyces cerevisiae*: Metabolism and gene expression. Strathern, J., Jones, E.W., and Broach, J.R (eds). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 283-317.

-McKenzie, K.Q. and Jones, E.W. (1977). Mutants in the formyl-tetrahydrofolate interconversion pathway of *Saccharomyces cerevisiae*. Genetics 86:85-102.

-McNeil, J.B., McIntosh, E.M., Taylor, B.V., Zhang, F.R., Tang, S. and Bognar, A.L., (1994). Cloning and molecular characterization of three genes, including two genes encoding serine hydroxymethyltransferases, whose inactivation is required to render yeast auxotrophic for glycine. J. Biol. Chem. 269: 9155-9165.

-McNeil, J.B., Bognar, A.L. and Pearlman, R.E. (1996). *In vivo* analysis of folate coenzymes and their compartmentation in *Saccharomyces cerevisiae*. Genetics 142:371-381.

-McNeil, J.B., Zhang, F, Taylor, B.V., Sinclair, D.A., Pearlman, R.E. and Bognar, A.L. (1997). Cloning, and molecular characterization of the *GCV1* gene encoding the glycine cleavage T-protein from *Saccharomyces cerevisiae*. Gene 186:13-20.

-Meedel, T.H. and Pizer, L.I. (1974). Regulation of One-Carbon Biosynthesis and Utilization in *Escherichia coli*. J. Bacteriol. 118:905-910.

-Melcher,K., Rose, M., Kunzler, M., Braus, G.H. and Entian, K.D. (1995). Molecular analysis of the yeast *SER1* gene encoding 3-phosphoserine aminotransferase: Regulation by general control and serine repression. Curr. Genetics 27:501-508.

-Melcher,K. and Entian, K.D (1992). Genetic analysis of serine biosynthesis and glucose repression in yeast. Curr. Genet. 21:295-300.

-Miller, J.H. (1972). Experiments in molecular genetics, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, pp. 352-355.

-Minehart, P.L. and Magasanik, B. (1991). Sequence and expression of *GLN3*, a positive nitrogen regulatory gene of *Saccharomyces cerevisiae* encoding a protein with a putative zinc finger DNA-binding domain. Mol. Cell. Biol. 11:6216-6228.

-Mortimer, R.K. and Hawthorne, D.C. (1966). Genetic Mapping in Saccharomyces. Genetics 53:165-187.

-Mösch, H.U., Graf, R., Schmidheini, T. and Braus, G. (1990). Three GCN4 responsive elements act synergistically act upstream and as TATA-like elements in the yeast *TRP4* promoter. EMBO J. 9:2951-2957

-Mösch, H.U., Scheier, B., Lahti, R., Mantsala, P. and Braus, G.H. (1991). Transcriptional activation of yeast nucleotide biosynthetic gene *ADE4* by *GCN4*. J. Biol. Chem. 266:20453-20456.

-Moye-Rowley, W.S., Harshman, K.D. and Parker, C.S. (1989). Yeast *YAP1* encodes a novel form of the jun family of transcriptional activator proteins. Genes Dev. 3:283-292.

-Mudd, S.H and Cantoni, G.L (1964). Biological transmethylation, methyl-group neogenesis and other "one-carbon" metabolic reactions dependent upon tetrahydrofolic acid. In Florkin, M and Stotz E.H. (ed.), Comprehensive Biochemistry (Amsterdam: Elsevier Publishing Company), vol.15 pp. 1-47.

-Nagarajan, L. and Storms, R.K. (1997). Molecular characterization of *GCV3*, the *Saccharomyces cerevisiae* Gene coding for the glycine cleavage system hydrogen carrier protein. J. Biol. Chem. 272:4444-4450.

-Nasmyth, K. and Shore, D. (1987). Transcriptional regulation in the yeast life cycle. Science 237:1162-1170.

-O'Connor, J.P and Peebles, C.L. (1992). *PTA1*, an essential gene of *Saccharomyces cerevisiae* affecting pre t-RNA processing. Mol. Cell. Biol. 12:3843-3856.

-O'Dell, B.L. (1947). Synthesis of Tetrahydrofolic acid, J. Am. Chem. Soc. 69:250.

-Okamura-Ikeda, K., Ohmura, Y., Fujiwara, K. and Motokawa,Y. (1993). Cloning and nucleotide sequence of the gcv operon encoding the *Escherichia coli* glycine-cleavage system. Eur. J. Biochem. 216:539-548.

-Olesen, J., Hahn, S. and Guarente, L. (1987). Yeast HAP2 and HAP3 activators both bind to the *CYC1* upstream activation site, UAS2, in an interdependent manner. Cell 51:953-961.

-Oliver, D.J (1994). The glycine decarboxylase complex from plant mitochondria. Annu. Rev. Plant. Physiol. Plant. Mol. Biol. 45:323-337.

-Ord, R.W., McIntosh, E.M., Lee, L., Poon, P.P. and Storms, R.K (1988). Multiple elements regulate expression of the cell cycle-regulated thymidylate synthase gene of *Saccharomyces cerevisiae*. Curr. Genet. 14:363-373.

-Paravicini, G., Mosch, H.U., Schmidheini, T. and Braus, G. (1989). The general control activator protein GCN4 is essential for a basal level of *ARO3* gene expression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9:144-151.

-Pasternack, L.B., Laude, D.A. and Appling D.R. (1992). ^{13}C NMR detection on folate-mediated serine and glycine synthesis *in vivo* in *Saccharomyces cerevisiae*. Biochemistry 31:8713-8719.

-Pasternack, L.B., Laude, D.A. and Appling D.R. (1994) ^{13}C NMR analysis of intercompartmental flow of one-carbon units into choline and purines in *Saccharomyces cerevisiae*. Biochemistry 33:74-82.

-Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature 335:683-689.

-Rai, R, Daugherty, J.R. and Cooper, T.G. (1995). UAS_{NTR} functioning in combination with other UAS elements underlies exceptional patterns of nitrogen regulation in *Saccharomyces cerevisiae*. Yeast 11:247-260

-Rebeille, F., Neuburger, M. and Douce, R. (1994). Interaction between glycine decarboxylase, serine hydroxymethyltransferase and tetrahydrofolate polyglutamates in pea leaf mitochondria. Biochem. J. 302:223-228.

-Rippe, K., von Hippel, P.H. and Langowski, J. (1995). Action at a distance: DNA-looping and initiation of transcription. Trends. Biochem. Sci. 12:500-506.

-Rothstein, R.J. (1983). One-step genedisruption in yeast. Methods Enzymol. 101:202-211.

-Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J. and Higuchi, R. (1988). Primer directed enzymatic amplification with a thermostable DNA polymerase. Science 239:487-491.

-Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: a laboratory manual. 2nd edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

-Santangelo, G.M. and Tornow, J. (1990). Efficient transcription of the glycolytic gene *ADH1* and three translational component genes requires the GCR1 product, which can act through TUF/GRF/RAP binding sites. Mol. Cell. Biol. 10:859-862.

-Scott, E.W. and Baker, H.V. (1993). Concerted action of the transcriptional activators REB1, RAP1, and GCR1 in the high-level expression of the glycolytic gene *TPI*. Mol. Cell. Biol. 13:543-550.

-Sellers, J.W. and Struhl, K. (1989). Changing fos oncoprotein to a jun-independent DNA binding protein with GCN4 dimerization specificity by swapping "leucine zippers". Nature 341:74-76.

-Shannon, K.W and Rabinowitz, J.C. (1988). Isolation and characterisation of the *Saccharomyces cerevisiae MIS1* gene encoding mitochondrial C1-tetrahydrofolate synthase. J. Biol. Chem. 263:7717-7725.

-Shao, D., Creasy, C.L. and Bergman, L.W. (1996). Interaction of *Saccharomyces cerevisiae* Pho2 with Pho4 increases the accessibility of the activation domain of Pho4. Mol. Gen. Genet. 251:358-364.

-Sherman, F. (1991a). Micromanipulation and dissection of Asci. Meth. Enzymol. 194:21-37.

-Sherman, F. (1991b). Getting started with yeast. Meth. Enzymol. 194: 1-21.

-Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988). *lambda ZAP:* A bacteriophage lamda expression vector with *in vivo* excision properties. Nucleic Acids Res. 16:7583.

-Sikorski, R.S. and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics 122:19-27.

-Sinclair, D.A. and Dawes, I.W. (1995). Genetics of the synthesis of serine from glycine and the utilization of glycine as sole nitrogen source by *Saccharomyces cerevisiae*. Genetics 140:1213-1222.

-Sinclair, D.A., Hong, S. and Dawes, I.W. (1996). Specific induction by glycine of the gene for the P-subunit of glycine decarboxylase from *Saccharomyces cerevisiae*. Mol. Microbiol. 19:611-623.

-Srinivasan, R. and Oliver, D.J. (1992). H-protein of the glycine decarboxylase multienzyme complex, complementary DNA encoding the protein from *Arabidopsis thaliana*. J. Plant Physiol. 98:1518.

-Staben, C. and Rabinowitz, J.C. (1986). Nucleotide sequence of the *Saccharomyces cerevisiae ADE3* gene encoding C1-tetrahydrofolate synthase. J. Biol. Chem. 261:4629-4637.

-Stanbrough, M. and Magasanik, B. (1995). Transcriptional and post-translational regulation of the general amino acidx permease of *Saccharomyces cerevisiae*. J. Bacteriol. 177:94-102.

-Stanbrough, M, Rowen, D.W. and Magasanik, B. (1995). Role of the GATA factors Gln3p and Nil1p of *Saccharomyces cerevisiae* in the expression of nitrogen-regulated genes. Proc. Natl. Acad. Sci. U S A. 92:9450-9454

-Stanbrough, M. and Magasanik, B. (1996). Two transcription factors, Gln3p and Nil1p, use the same GATAAG sites to activate the expression of *GAP1* of *Saccharomyces cerevisiae*. J. Bacteriol. 178:2465-2468.

-Stargell, L.A. and Struhl, K. (1996). Mechanisms of transcriptional activation *in vivo*: two steps forward. Trends. Genet. 12:311-315.

-Storms, R.K., Ord, R.W., Greenwood, M.T., Miradammadi, B., Chu, F.K. and Belfort, M. (1984). Cell cycle dependent expression of thymidylate synthase in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 4:2858-2864.

-Stotz, A., Muller, P.P. and Linder, P. (1993). Regulation of the *ADE2* gene from *Saccharomyces cerevisiae*. Curr. Genet. 6:472-480.

-Struhl, K. and Davis, R.W. (1981). Promoter mutants of the yeast *HIS3* gene. J. Mol. Biol. 152:553-568.

-Struhl, K. (1989). Molecular mechanisms of transcriptional regulation in yeast. Annu. Rev. Biochem. 58:1051-1077.

-Struhl, K. (1993). Yeast transcription factors. Curr. Opin. Cell. Biol. 5:513-520.

-Struhl, K. (1995). Yeast transcriptional regulatory mechanisms. Annu. Rev. Genet. 29:651-674.

-Struhl, K. (1996a). Transcriptional enhancement by acidic activators. Biochim. Biophys. Acta. 1288:O15-O17.

-Struhl, K. (1996b). Chromatin structure and RNA polymerase II connection: implications for transcription. Cell 84:179-182.

-Svetlov, V.V. and Cooper, T.G. (1995). Review: Compilation and characteristics of dedicated transcription factors in *Saccharomyces cerevisiae*. Yeast 11: 1439-1484.

-Tada, K. (1992). Non-ketotic hyperglycinaemia: a life-threatening disorder in the neonate. Early Hum. Dev. 29:75-81.

-Tada, K. and Hayasaka, K. (1987). Non-ketotic hyperglycinaemia: clinical and biochemical aspects. Eur. J. Pediatr. 146:221-227.

-Tada, K. (1993). Non-ketotic hyperglycinaemia: molecular lesion, diagnosis and pathophysiology. J. Inherit. Metab. Dis. 16:691-703.

-Tice-Baldwin, K., Fink, G.R. and Arndt, K.T. (1989). BAS1 has a Myb motif and activates *HIS4* transcription only in combination with BAS2. Science 246:931-935.

-Toone, J.R. (1992). Prenatal diagnosis of Non-ketotic Hyperglycinaemia. J. Inher. Metab. Dis. 15:713-719.

-Toone, J.R. (1994). Prenatal diagnosis of Non-ketotic Hyperglycinaemia: experience in 50 at-risk pregnancies. J. Inher. Metab. Dis. 17:342-344.

-Tornow, J., Zeng, X., Gao, W. and Santangelo, G.M. (1993). GCR1, a transcriptional activator in *Saccharomyces cerevisiae*, complexes with RAP1 and can function without its DNA binding domain. EMBO J. 12:2431-2437.

-Uemura, H. and Jigami, Y. (1992). Role of GCR2 in transcriptional activation of yeast glycolytic genes. Mol. Cell. Biol. 12:3834-3842.

-van Vuuren, H.J., Daugherty, J.R., Rai, R. and Cooper, T.G. (1991). Upstream induction sequence, the cis-acting element required for response to the allantoin pathway inducer and

enhancement of operation of the nitrogen-regulated upstream activation sequence in *Saccharomyces cerevisiae*. J. Bacteriol. 173:7186-7195.

-Vincent, A.C. and Struhl, K. (1992). ACR1, a yeast ATF/CREB repressor. Mol. Cell. Biol. 12:5394-5405.

-Wek, R.C., Cannon, J.F., Dever, T.E. and Hinnebusch, A.G (1992). Truncated protein phosphatase GLC7 restores translational activation of *GCN4* expression in yeast mutants defective for the eIF-2 kinase GCN2. Mol. Cell. Biol. 12:5700-5710.

-Wiame, J.M., Grenson, M. and Arst, H.N. (1985). Nitrogen catabolite repression in yeasts and filamentous fungi. Adv. Microb. Physiol. 26:1-88.

-Willett, C.E., Gelfman, C.M. and Holland, M.J. (1993). A complex regulatory element from the yeast gene *ENO2* modulates GCR1-dependent transcriptional activation. Mol. Cell. Biol. 13:2623-2633.

-Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E. and Young, R.A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell. 84:235-244.

-Wu, A., Wemmie, J.A., Edgington, N.P., Goebl, M., Guevara, J.L., and Moye-Rowley, W.S. (1993). Yeast bZip proteins mediate pleiotropic drug and metal resistance. J. Biol. Chem. 268:18850-18858.

-Yamamoto, M., Koyata. H., Matsui, C. and Hiraga, K. (1991). The glycine cleavage system; Occurrence of two types of chicken H-Protein mRNAs presumably formed by the alternative use of the polyadenylation consensus sequences in a single exon. J. Biol. Chem. 266:3317-3322.

-Yoo, H.S. and Cooper, T.G. (1989). The *DAL7* promoter consists of multiple elements that cooperatively mediate regulation of the gene's expression. Mol. Cell. Biol. 9:3231-3243.