Cell, Vol. 43, 177–188, November 1985, Copyright © 1985 by MIT 0092-8674/85/110177-12 $02.00/0

GCN4 Protein, Synthesized In Vitro, Binds *HIS3* Regulatory Sequences: Implications for General Control of Amino Acid Biosynthetic Genes in Yeast

Ian A. Hope and Kevin Struhl
Department of Biological Chemistry
Harvard Medical School
Boston, Massachusetts 02115

Summary

The yeast GCN4 gene product is necessary for the transcriptional induction of many amino acid biosynthetic genes in response to conditions of amino acid starvation. We synthesized radioactively pure GCN4 protein by in vitro translation of mRNA produced by in vitro transcription with SP6 RNA polymerase. GCN4 protein binds specifically to the 20 bp region of the *HIS3* gene that is critical for transcriptional regulation in vivo and contains the TGACTC sequence common to coregulated genes. A synthetic GCN4 mutant protein lacking the 40 C-terminal amino acids fails to bind DNA; this correlates with a *gcn4* mutant gene that is nonfunctional in vivo. Finally, GCN4 protein binds to the promoter regions of coordinately regulated genes, but not to analogous regions of other genes. We suggest that GCN4 protein is a specific transcription factor, and we describe a molecular model for the general control of amino acid biosynthetic genes.

Introduction

The general control system of the yeast Saccharomyces cerevisiae coordinately regulates the expression of genes encoding amino acid biosynthetic enzymes from different pathways (Schurch et al., 1974; Delforge et al., 1975; Wolfner et al., 1975; reviewed by Jones and Fink, 1982). Under normal growth conditions, these genes are constitutively expressed at a basal level. However, when cells are starved for any one of several amino acids, expression of all of these genes is induced 2- to 10-fold. For all cases that have been tested, this induction is due to increased transcription (Struhl and Davis, 1981; Zalkin and Yanofsky, 1982; Donahue et al., 1983; Messenguy et al., 1983; Aebi et al., 1984).

The *HIS3* gene, which encodes a histidine biosynthetic enzyme, is regulated in response to amino acid starvation and hence subject to general control. Extensive deletion analyses have defined the sequence elements upstream of the transcriptional initiation site that are critical for proper *HIS3* expression. Two regions are required for constitutive expression, a TATA element 35–55 nucleotides before the site of transcriptional initiation (nucleotides −35 to −55) and an upstream element located between nucleotides −115 and −129 (Struhl, 1981; 1982a; submitted). A separate region between −86 and −99 is critical for positive regulation of *HIS3* expression in response to amino acid starvation (Struhl, 1982b; Hill and Struhl, unpublished). This regulatory region contains the sequence TGACTC, which with some variation is repeated 6 times in the *HIS3* promoter region. The same sequence is repeated upstream of other coregulated genes (Hinnebusch and Fink, 1983a), and deletion analysis of *HIS4* indicates that it is critical for transcriptional control (Donahue et al., 1983; Hinnebusch et al., 1985). Thus, this sequence is implicated as the *cis*-acting site required by regulated genes for recognition by the general control system.

General control has also been investigated by isolating unlinked mutations that fail to regulate the biosynthetic genes in a proper fashion. Six genes for *trans*-acting regulatory factors have been identified, *GCN1-5* and *GCD1* (Schurch et al., 1974; Wolfner et al., 1975; Penn et al., 1983; Hinnebusch and Fink, 1983b). In *gcn* mutants, transcription of the biosynthetic genes cannot be stimulated, whereas in *gcd1* mutants, transcription is always at the induced levels. The epistatic relationships of the mutations in these six genes suggest that the *GCN1,2,3,5* gene products act indirectly through the *GCD1* gene product, and they implicate the *GCN4* gene product as having the most direct role in the transcriptional stimulation of regulated genes (Hinnebusch and Fink, 1983b). However, a direct link between the genetic analysis of proteins involved in the general control system and the molecular studies of regulated genes remains to be established. The simplest hypothesis is that the *GCN4* protein binds to the TGACTC regulatory sequences upstream of the coregulated genes and coordinately activates their transcription.

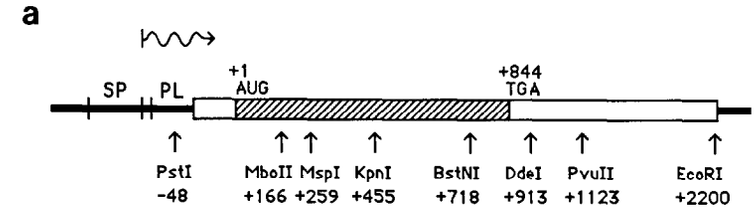
To show that the *GCN4* gene product recognizes genes regulated by the general control system, GCN4 protein was first synthesized by in vitro translation of RNA that was obtained by in vitro transcription of the cloned gene with SP6 RNA polymerase. It seemed reasonable to expect that if the GCN4 protein, as synthesized in vivo, binds to the upstream region of regulated genes, then this property might be shown by the GCN4 primary translation product. First, it was estimated that reasonable quantities of ^{35}S-labeled GCN4 protein could be produced at a high specific activity. Second, the level of GCN4 protein increases during conditions of amino acid starvation (Thireos et al., 1984; Hinnebusch, 1984), suggesting that posttranslational events are not essential for the activity of the protein.

Here we demonstrate that the GCN4 in vitro translation product binds specifically to DNA fragments containing the *HIS3* regulatory site; the DNA sequences necessary and sufficient for protein binding in vitro and transcriptional regulation in vivo appear indistinguishable; the C-terminal 40 amino acids are necessary for DNA binding and for activity in vivo; and the GCN4 protein binds to the upstream regions of coordinately regulated genes, but not to analogous regions of other genes.

Results

Synthesis of GCN4 Protein

GCN4 protein was generated by in vitro transcription and



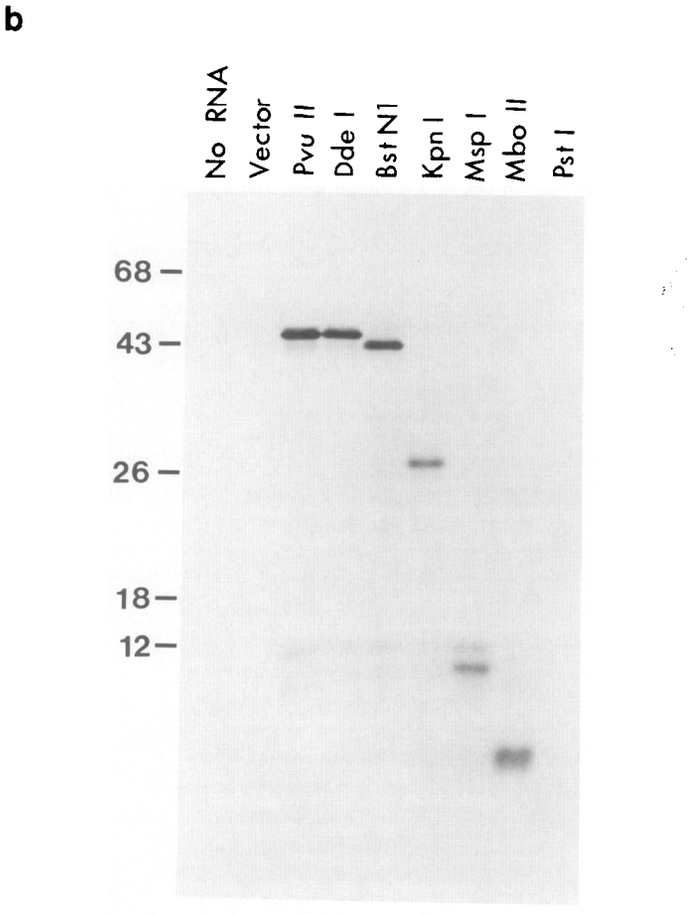


Figure 1. Synthesis of GCN4 Protein

(a) Relevant structural details of pSP64-*GCN4*, the construction of which is described in Experimental Procedures. The bold line represents pSP64 (vector) DNA, the open box is yeast DNA, and the hatched area is the GCN4 coding region, SP is the SP6 promoter, PL is the polylinker in the vector, and the wavy arrow indicates the direction of transcription. For each restriction enzyme, the cleavage sites indicated are the first encountered after the SP6 promoter, and the numbers refer to their positions (in bp) from the AUG initiation codon. (b) Autoradiograph of in vitro translation products (1 μl/lane) generated after in vitro transcription of pSP64-*GCN4* DNA cleaved with various restriction enzymes (Pvu II, Dde I, Bst NI, Kpn I, Msp I, Mbo II, or Pst I), of pSP64 DNA (vector), cleaved with Pvu II, or generated in the absence of exogenous RNA (no RNA). Samples were analyzed by SDS-PAGE and were compared with proteins of known molecular weight (units are kilodaltons).

translation of a hybrid DNA molecule consisting of the GCN4 coding region cloned into the pSP64 vector. The *GCN4* gene used in these experiments was isolated by complementation of a *gcn4* mutation; it is structurally indistinguishable from the *GCN4* gene the DNA sequence of which was determined previously (Thireos et al., 1984; Hinnebusch, 1984). The GCN4 coding region extending from 27 bp upstream of the *GCN4* translational initiation codon to 1400 bp beyond the termination codon was cloned into pSP64 (Melton et al., 1984) at a position just downstream of the promoter for SP6 RNA polymerase (Figure 1a; see Experimental Procedures for details). After cleavage with Pvu II, this pSP64-GCN4 molecule was transcribed with SP6 RNA polymerase in the presence of diguanosine triphosphate (G-5′ppp5′-G) to produce a 5′-capped "run-off" transcript. To generate GCN4 protein, this artificial *GCN4* mRNA was translated in vitro using an extract from wheat germ.

A high level of specific synthesis of GCN4 protein was obtained. Typically, starting with 0.5 μg of pSP64-GCN4 DNA, 1.5 × 10^6 cpm of ^{35}S-methionine was incorporated into GCN4 protein in a translation volume of 30 μl. From the specific activity of the ^{35}S-methionine and the predicted amino acid sequence of the GCN4 protein (Thireos et al., 1984; Hinnebusch, 1984), a minimum of approximately 5 ng of GCN4 protein was synthesized per reac-

Table 1. Predicted and Apparent Molecular Weights of Truncated GCN4 Proteins

Endonuclease Used to Digest Transcription Template	Predicted Protein MW*	Observed Protein MW*
Pvu II		44,000
Dde I	44,000	44,000
Bst NI	37,500	41,000
Kpn I	23,700	25,700
Msp I	13,500	11,600
Mbo II	8,000	8,400

* Expressed in daltons.

tion, which corresponds to a concentration of approximately 8 nM. The autoradiogram in Figure 1b indicates that in terms of ^{35}S-labeled material, the GCN4 protein is essentially pure (roughly 95%). Minor bands of lower molecular weight, probably arising from premature termination, are also observed.

The apparent molecular weight of the GCN4 translation product as determined from its relative electrophoretic mobility is 44,000 daltons. This is considerably larger than 30,000 daltons, the molecular weight predicted from the amino acid sequence (Hinnebusch, 1984; Thireos et al., 1984). To confirm that the observed translation product was the correct protein, the pSP64-GCN4 template DNA was cleaved with restriction enzymes at various positions within the *GCN4* coding region before being transcribed and translated (Figure 1). The apparent molecular weights of the truncated translation products are in reasonable agreement with those predicted by multiplying 44,000 daltons (the apparent molecular weight of the full-size product) by the proportion of the coding region remaining after digestion of the template with each restriction enzyme (Table 1). In particular, cleavage 70 bp beyond the termination codon does not affect the size of the GCN4 protein, whereas cleavage 120 bp before the termination codon shortens the protein by an expected amount. Thus this method accurately generates the primary GCN4 translation product that migrates anomalously under the conditions described here. Such anomalous behavior has been observed with other translation products (Kimelman et al., 1984).

Specific DNA Binding Activity of the GCN4 Protein

The electrophoretic mobilities of DNA–protein complexes are different from those of free DNA or free protein (Garner and Revzin, 1981; Fried and Crothers, 1981). Thus, GCN4 binding activity was detected as an altered electrophoretic mobility of the ^{35}S-labeled translation product following incubation with specific DNA fragments. Similar assays for DNA binding activity carried out previously (Garner and Revzin, 1981; Fried and Crothers, 1981) have used excess protein to alter the mobility of unlabeled or ^{32}P-labeled DNA fragments and have served as the basis for the purification of proteins from crude mixtures (Strauss and Varshavsky, 1984). In contrast, the GCN4 protein represents a small percentage of the total protein in the in vitro translation mixture, but is essentially pure in terms of ^{35}S-labeled species. So, in this study, DNA to be bound was used in excess, and an altered electrophoretic mobility of labeled protein in the presence of unlabeled DNA was sought.

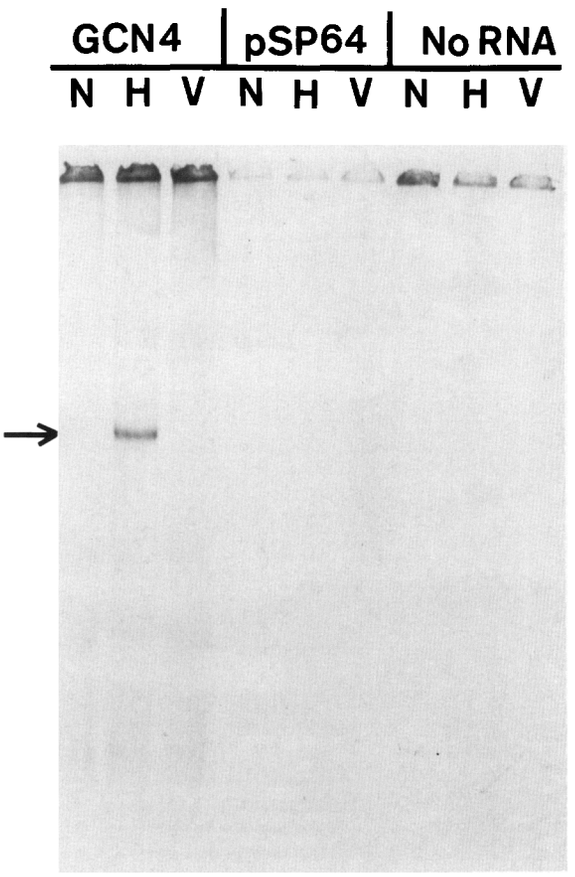


Figure 2. Specific DNA Binding Activity of GCN4 Protein

^{35}S-translation products generated by transcription and translation of pSP64-*GCN4* or pSP64 DNA, or in the absence of exogenous RNA were incubated with Taq I-digested pUC8-*HIS3* (H), pUC9 (V) DNAs, or with no DNA (N). Products were analyzed by native PAGE and autoradiography. The arrow indicates the mobility of the complex between GCN4 protein and *HIS3* DNA.

The initial DNA substrate, pUC8-Sc2676, contains a 1.8 kb *HIS3* DNA fragment cloned into pUC8. This DNA was cleaved with Taq I to produce eight fragments, one of which contains the entire *HIS3* regulatory region on a 294 bp fragment (nucleotides −24 to −310). The DNA was incubated with the in vitro translation mixture containing GCN4 protein and was then subject to electrophoresis in a 5% native polyacrylamide gel. The gel was then fluorographed to reveal the position of labeled protein.

In the presence of Taq I-cleaved pUC8-*HIS3* DNA, a single strong band of GCN4 protein was observed (Figure 2). This band was not observed if DNA was omitted or if pUC9 DNA was substituted for pUC8-*HIS3* DNA. Furthermore, this band was not detected if the translation products used in the binding were generated in the absence of exogenous RNA or in the presence of the RNA transcribed from the pSP64 vector DNA. Thus, GCN4 is a specific DNA

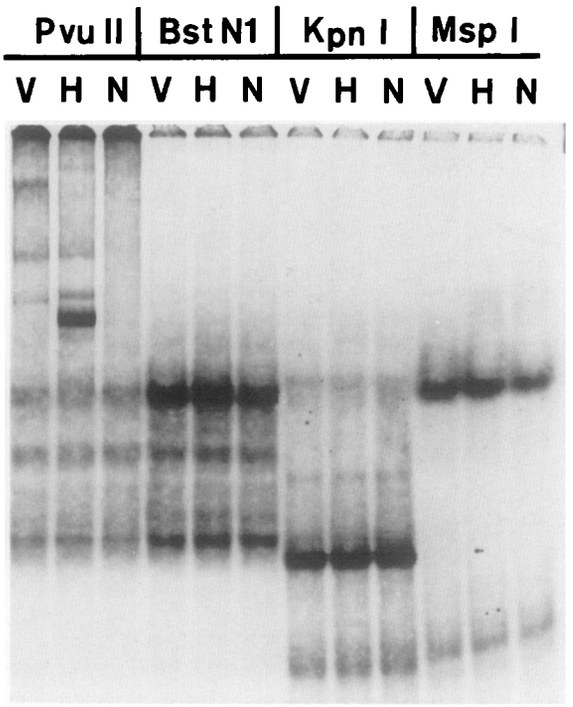


Figure 3. Analysis of DNA Binding Activity of Mutant GCN4 Proteins Lacking C-Terminal Amino Acids

Full-size and truncated GCN4 proteins, produced by transcription and translation of pSP64-*GCN4* DNA cleaved with Pvu II, Bst NI, Kpn I, or Msp I (described in Figure 1) were incubated with Taq I-digested pUC8-*HIS3* (H) or pUC9 (V) DNAs or in the absence of DNA (N), and then were analyzed by native PAGE and autoradiography.

binding protein. It will be shown below that the appearance of this band is due to the formation of a complex between GCN4 protein and a single Taq I fragment of *HIS3* DNA.

When GCN4 protein is incubated with either pUC8-*HIS3* or pUC9 DNA, several faint bands are observed in addition to the single strong band (Figure 2). Since these bands are not observed in the absence of DNA, they presumably represent nonspecific binding to the other Taq I fragments in the incubation mixture. Nonspecific DNA binding is a characteristic of essentially all DNA binding proteins.

Mutant GCN4 Proteins Lacking the C-Terminus Do Not Bind DNA

Because the GCN4 protein is synthesized in vitro from cloned DNA segments, it is possible to generate mutant proteins simply by altering the template. Therefore, any desired GCN4 derivative can be tested for its ability to bind DNA. As described in Figure 1, cleavage of the pSP64-*GCN4* template with appropriate restriction endonucleases permits the synthesis of truncated GCN4 proteins that lack the C-terminus. When any of these mutant proteins was assayed for DNA binding activity, no new bands corresponding to protein–DNA complexes were observed (Figure 3). For each mutant protein, identical ^{35}S

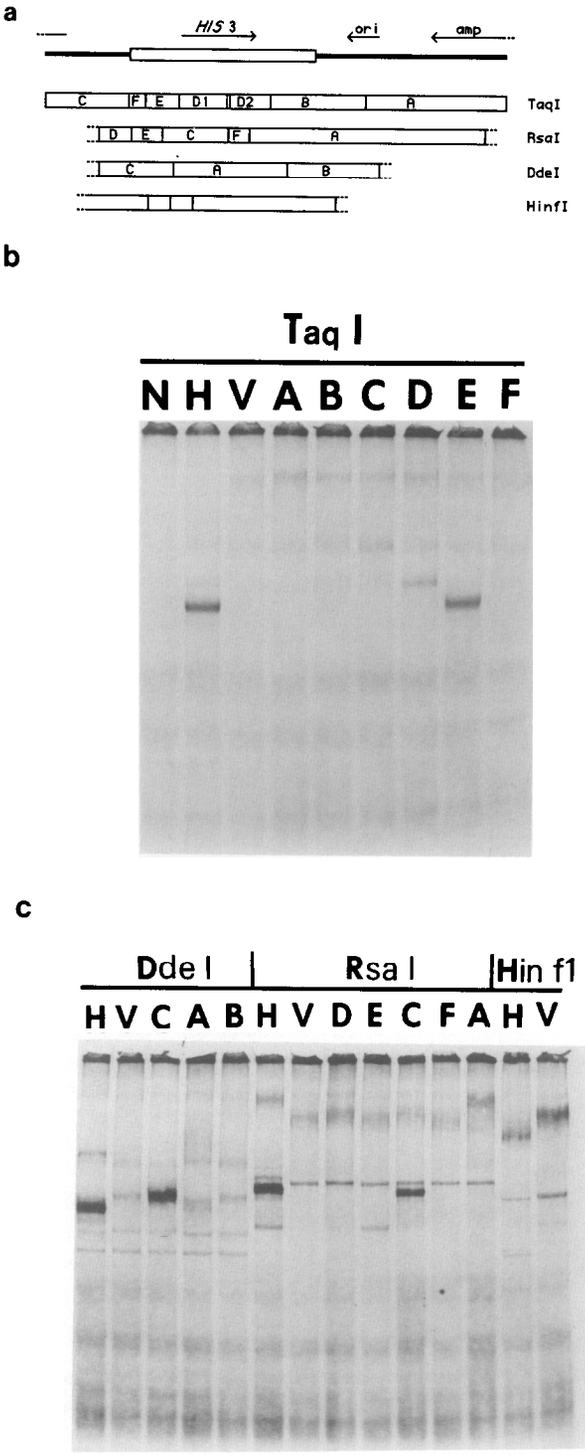


Figure 4. Localization of the GCN4 Protein Binding Site Using Restriction Fragments

(a) Taq I, Rsa I, Dde I, and Hinf I restriction maps of pUC8-*HIS3* DNA. The structure of pUC8-*HIS3* is depicted uppermost with the bold line representing pUC8 vector DNA and the open box representing the 1.8 kb Bam HI yeast DNA fragment containing the *HIS3* gene (Sc2676) (Struhl and Davis, 1980). The positions of the *HIS3* gene, the origin of plasmid DNA replication, and the ampicillin resistance gene are indicated with arrows. For each restriction enzyme, the fragments are labeled alphabetically according to size (fragment A is the largest). The D1 and D2 fragments produced by Taq I digestion were not resolved by agarose gel electrophoresis and were assayed together. (b) Analy-

profiles were observed in the presence and absence of pUC8-*HIS3* DNA. The free proteins migrate idiosyncratically because they differ in charge and in molecular weight. These results suggest that the truncated GCN4 proteins do not possess any specific or nonspecific DNA binding activity. Of particular interest is the "Bst N1 generated" protein that lacks the 40 C-terminal amino acids of GCN4 protein and is defective in DNA binding activity. Deletion of the region of the *gcn4* gene encoding this segment of the protein abolishes the regulatory function in vivo (Hinnebusch, 1984). Therefore, these results indicate a relationship between DNA binding in vitro and transcriptional activation in vivo, and they implicate the C-terminus of the GCN4 protein as a critical component for both.

Localization of the GCN4 Protein Binding Site on *HIS3* DNA Fragments

The assay described above establishes that GCN4 protein has DNA binding activity that correlates with its function as inferred from previous genetic analysis. To complete the link between the genetic studies of genes involved in general control and the molecular studies of elements required in *cis* by the regulated genes, the sites with which the GCN4 protein interacts must be identified. The recognition site was first localized by using isolated restriction enzyme DNA fragments in the binding assay (Figure 4a).

The DNA fragments generated by Taq I digestion of pUC8-*HIS3* DNA were purified by agarose gel electrophoresis. Individual DNA fragments were combined in equimolar amounts with Taq I-cleaved pUC9 DNA and were then tested by the gel mobility assay (Figure 4b). The 294 bp Taq I fragment was the only one to generate the strong DNA–protein complex band observed with the total Taq I digestion products of pUC8-*HIS3* DNA. Thus, it is concluded that the site recognized by GCN4 protein is present on this Taq I fragment that contains the nucleotides from −317 to −23 relative to *HIS3* transcription initiation. This region contains all the sequences necessary for proper *HIS3* regulation (Struhl, 1982b).

From extensive analysis of the *HIS3* and *HIS4* genes, TGACTC has been proposed as the consensus sequence of the element required in *cis* for general control regulation (Struhl, 1982b; Donahue et al., 1983). The 294 bp Taq I fragment contains two copies of the TGACTC sequence at position −99 to −94 and −258 to −263 as well as 4 other sites with 5 out of 6 matches to this sequence at −142 to −137, −181 to −176, −216 to −221, and −225 to −230 (Figure 5a). So, the site on the Taq I fragment recognized by GCN4 protein was further localized by assaying individual DNA fragments produced with two other restriction enzymes, Rsa I and Dde I (Figure 4c). The major GCN4 binding site was found on a 600 bp Rsa I fragment containing the *HIS3* nucleotides −160 to +440 and on a 620 bp Dde I fragment extending from −80 through −447 into pUC8 DNA. Therefore, the nucleotide sequence necessary and sufficient for GCN4 binding is contained between −80 and −160 with respect to the *HIS3* transcriptional initiation site. This region contains one perfect TGACTC sequence (from −99 to −94) and one imperfect sequence (from −141 to −136).

Within the region so far defined to contain the GCN4 protein recognition site, there is a single Hinf I site and it is located within the TGACTC sequence. When the products of Hinf I cleavage of pUC8-*HIS3* are used in the binding assay, no strong GCN4 protein–DNA complex was observed (Figure 4c). Although alternative explanations are possible, for example, GCN4 protein binding sites located near the ends of DNA fragments may not be bound by GCN4 protein, this result is consistent with Hinf I cleavage of *HIS3* DNA disrupting the sequence recognized by the GCN4 protein.

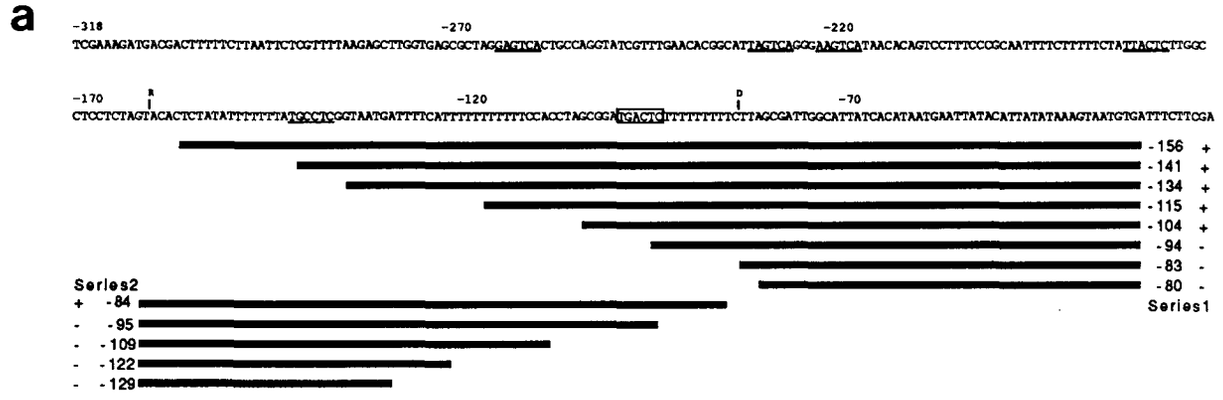
Fine Mapping of the GCN4 Protein Binding Site

To map the GCN4 protein binding site for *HIS3* more precisely, DNA constructs containing deletion endpoints at various positions through the HIS3 upstream region were used in the binding assay (Figure 5a). Two series of DNA constructs were employed. Series 1 consists of *HIS3* DNA from the Bam HI site at +1318 to an Eco RI site introduced at positions from −80 to −156. Series 2 consists of *HIS3* DNA from the Bam HI site at −447 to an Eco RI site introduced at positions from −129 to −83. These constructions were generated previously in the deletion analysis of the *HIS3* upstream region (Struhl, 1982a; 1982b; submitted). For both series, the relevant Bam HI-Eco RI fragments were recloned into pUC9. The resulting DNAs were digested with Pvu II and Rsa I (for series 1) or with Pvu II alone (for series 2) before use in the binding assay; this places the *HIS3* deletion break points close to the center of approximately 600 bp fragments for both series.

The results of the binding assays using these two DNA series is presented in Figure 5b. For series 1, derivatives with *HIS3* endpoints at −104 or further upstream contain the strong GCN4 binding site, whereas derivatives with endpoints at −94 or further downstream do not. For series 2, the strong GCN4 binding site is present on *HIS3* DNA from −447 to −83 but not for constructs with *HIS3* DNA from −447 to −95 and shorter. Therefore, the GCN4 protein recognition sequence is contained entirely within the −83 to −104 region, and the TGACTC sequence within this region is critical for binding. This firmly establishes that GCN4 protein recognizes a gene regulated by general control by binding to an element with TGACTC as part of its identity. TGACTC alone is not sufficient because strong binding to the other TGACTC sequence in the *HIS3* upstream region (−258 to −263) has not been observed and because this sequence is found in several of the pUC9 control DNA fragments.

It should be noted that for series 2 derivatives lacking the strong binding site, GCN4 protein appears to bind above the level for nonspecific interactions (for example,

sis of GCN4 protein binding to individual Taq I DNA fragments (A to F) as well as to Taq I-digested pUC8-*HIS3* (H) and pUC9 (V) DNAs. (c) A similar analysis for individual Rsa I (A,C to F) and Dde I fragments (A to C) and for Hinf I-digested pUC8-*HIS3* and pUC9 DNAs. Strong bands of radioactivity (observed for Taq I fragment E, Rsa I fragment C, and Dde I fragment C) indicate specific protein–DNA complexes; weaker bands indicate nonspecific binding.



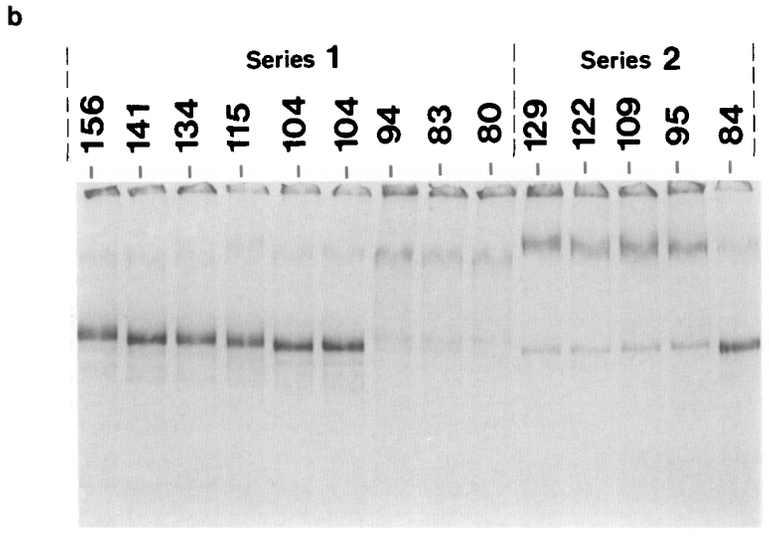


Figure 5. Fine Mapping of the GCN4 Protein Binding Site Using *his3* Deletion DNA Fragments

(a) The nucleotide sequence of the 294 bp Taq I-E fragment, which is numbered relative to the *HIS3* transcriptional initiation site (+1). The TGACTC sequence implicated as important for GCN4 protein binding is boxed, the other TGACTC sequence and related sequences are underlined, and the Rsa I and Dde I sites are indicated (R and D). Each bold line under the DNA sequence represents the region of *HIS3* DNA that is present in that deletion fragment. The *HIS3* DNA fragments extend from the nucleotide indicated by the number adjacent to each segment to +1318 (for series 1) or to −447 (for series 2). The DNA binding properties of these deleted derivatives (as determined from the autoradiograph in b) are indicated (+ or −). (b) Assay of *HIS3* deletion DNA fragments for GCN4 protein binding sites. GCN4 protein was incubated with restriction fragments of pUC9 containing the *HIS3* deletion DNA fragments (extending to the nucleotide indicated above each lane), before native PAGE and autoradiography.

compare the band intensities to that of the −80 derivative of series 1). Similarly, the Rsa I fragment E (nucleotides −161 to −415) also demonstrates binding above the nonspecific level (Figure 4c). These observations may be due to weak binding interactions to one or several copies of the TGACTC-related sequences located upstream of the strong binding site.

DNAase I Protection of *HIS3* DNA by GCN4 Protein

To augment the conclusions drawn above, the footprinting technique of Galas and Schmitz (1978) was employed. Initially, the DNA to be bound was the 255 bp Taq I-Alu I fragment from the upstream region of *HIS3* (nucleotides −24 to −279), which was 3′ end-labeled at the Taq I site. DNAase I treatment was carried out for 2 min after 20 min equilibration of DNA and translation mix, plus or minus GCN4 protein. The digestion products were examined on a sequencing gel with autoradiography.

GCN4 protein partially protects a 10 bp region from DNAase I digestion (Figure 6). The protected region, from −102 to −92 with respect to the *HIS3* transcriptional initiation site, includes the TGACTC implicated as part of the recognition site in the binding assays above. No other protection from DNAase I action was detected between −40 and −200. Interestingly, the bound GCN4 protein causes a marked increase in DNAase I sensitivity just downstream from the protected region, between nucleotides −86 and −84.

The DNAase I protection by GCN4 protein was examined on the other *HIS3* strand by using a 602 bp Hin PI-Hind III fragment (nucleotides −270 to +332) that was 3′ end-labeled at the Hin PI site. The results suggest that protection by GCN4 protein extends over a similar region (nucleotides −100 to −95), but may be less pronounced. Interestingly, enhanced DNAase I cleavage is also observed for this strand, but at the opposite end of the protected region, between nucleotides −101 and −103.

These observations are further evidence for specific binding of GCN4 protein to the segment of DNA upstream of *HIS3*, which includes the most proximal copy of the TGACTC sequence. Weaker binding to other DNA segments upstream of *HIS3* may not have been detected in this experiment. The enhancement of DNAase I cleavage adjacent to the region of protection suggests that DNA

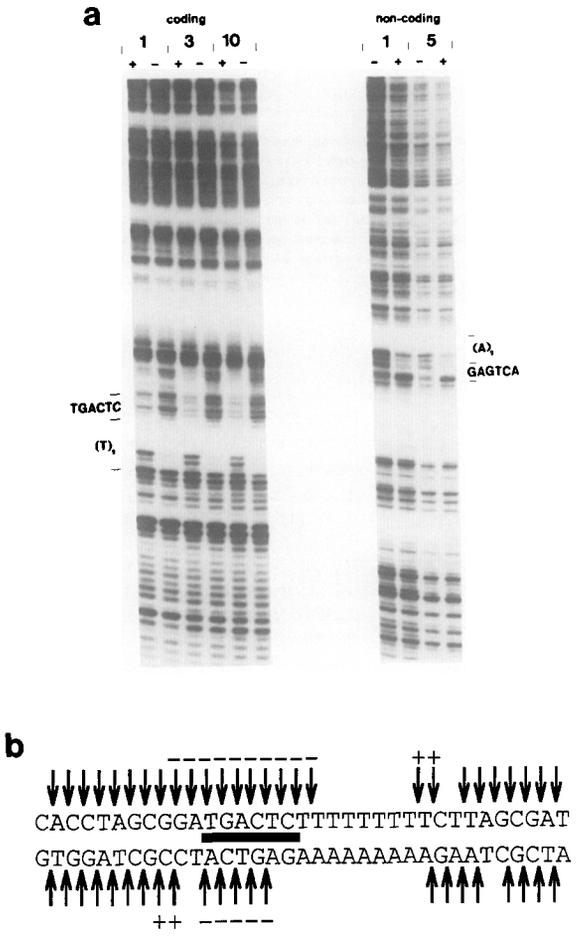


Figure 6. Protection of *HIS3* DNA from DNAase I Cleavage by GCN4 Protein

(a) *HIS3* DNA fragments, 3′ end-labeled on either the coding or non-coding strands, were incubated with various amounts of translation products (the number of μl used per 20 μl reaction is indicated above each lane) that either contain (+) or do not contain (−) GCN4 protein. After controlled digestion with DNAase I, the products were examined on a sequencing gel along with the same DNA fragments digested with selected restriction enzymes to serve as size markers. Regions of enhanced and reduced DNAase I action as a result of GCN4 protein are indicated with the nucleotide sequences. (It is noted that there appears to be some DNAase I activity in the translation products such that the addition of larger amounts increases degradation). (b) Results of the footprinting analysis at nucleotide resolution. Sites of DNAase I action on both strands of DNA between nucleotides −110 and −86 are indicated with arrows. DNAase I cleavages enhanced or suppressed by GCN4 protein are indicated (+ or −, respectively). DNAase I cleavage between certain nucleotides, and in particular at the dA:dT region, is poor (absence of arrows), and differences between the presence and absence of GCN4 protein were not detected. The black bar highlights the TGACTC sequence.

binding by GCN4 protein induces a change in DNA conformation.

Binding of GCN4 Protein to Other Yeast Promoter Regions

Gcn4 mutant strains fail to induce the transcription of all genes so far examined that are normally subject to general control (Penn et al., 1983; Hinnebusch and Fink, 1983b). If the role of the *GCN4* gene product in general control is to recognize the promoters of coregulated genes, then the *GCN4* protein should bind only to the promoter regions of these genes.

The promoter regions of eight yeast genes were isolated on approximately 700 bp DNA fragments (Table 2) and then were examined for their ability to bind GCN4 protein (Figure 7). The genes include four that are regulated by general control (*HIS4*, *ARG4*, *TRP5*, and *HIS3*), and four that are not regulated (*URA3*, *TRP1*, *GAL1,10*, and *DED1*). As expected, GCN4 protein binds specifically to the upstream regions of the genes subject to general control and not to the DNA fragments of the other four genes tested. Therefore, it is likely that the transcriptional stimulation of all genes that are coregulated by the general control system depends on the binding of GCN4 protein to the promoter regions.

Discussion

The Use of In Vitro Synthesized Proteins for DNA Binding Studies

This paper describes new methods for determining whether a cloned gene encodes a specific DNA binding protein and for analyzing DNA–protein interactions. The basis of the method is to clone the protein coding sequences into a vector containing a promoter for SP6 RNA polymerase, to produce messenger RNA by transcribing the DNA template with this enzyme, and to synthesize the desired protein as a ^{35}S-labeled species by translation of this mRNA in vitro. To detect DNA binding activity, the labeled protein is incubated with specific DNA fragments, and protein–DNA complexes are separated from free protein by electrophoresis in native acrylamide gels.

This method has a number of useful features. First, the protein is synthesized essentially as a radioactively pure species. Because the assay method follows the fate of the radioactive label in the presence of DNA, the observed effects can be ascribed directly to the DNA binding activity of the protein. Second, it is possible to synthesize relatively large amounts of protein in a short amount of time. In our standard experiment, it takes a few hours to synthesize 5 ng of protein in a volume of 30 μl; this amount is enough for standard DNAase I footprinting experiments. Moreover, because the substrates (cloned DNA, SP6 RNA polymerase, wheat germ extract, and cofactors) are available in essentially unlimited quantities, we estimate that it should be possible to synthesize at least 1 μg of protein. Third, by incubating the protein with a variety of DNA fragments, it is possible to examine its specific and nonspecific DNA binding properties and to localize precisely the DNA binding sequences. Fourth, any desired mutant protein can be created simply by altering the DNA template and can then be tested for its specific and nonspecific DNA binding properties. In yeast, the function in vivo of such mutant proteins can be assessed by reintroducing the appropriate DNAs into cells. This approach is considerably easier than the isolation and purification of mutant proteins from whole organisms. Fifth, the ^{35}S-protein itself can be used as a substrate to assay for cellu-

Table 2. DNA Fragments Used to Assay Eight Yeast Genes for GCN4 Protein Binding Sites

Gene	DNA Fragment			Region of Gene Contained, Relative to Initiation of	References
	Size (bp)	Enzymes Used	Source		
HIS4	663	Dde I	YIp302	−494 to +169, Transcription	Farabaugh and Fink, (1980); Donahue et al. (1983)
HIS3	599	Rsa I	pUC8-*HIS3*	−161 to +438, Transcription	This Work
TRP5	581	Bam HI, Bst EII	pYE(*trp5*)1	−453 to +128, Transcription	Walz et al. (1978); Zalkin and Yanofsky (1982)
ARG4	734	Rsa I, Eco RV	pYE(*arg4*)511	−473 to +261, Translation	Clarke and Carbon, (1978); Beacham et al. (1984)
DED1	745	Bgl II	pUC8-*HIS3*	−483 to +262, Transcription	This Work
URA3	735	Hind III, Hinc II	YIp55	−228 to +507, Translation	Bach et al. (1979); This Work
TRP1	619	Eco RI, Hind III	pUC8-Sc4101	−102 to +517, Translation	Struhl et al. (1979); This Work
GAL1,10	390	Eco RI, Hind III	pUC8-Sc3296	*	Struhl (1984)

* *GAL1* and *GAL10* are transcribed divergently, and the DNA fragment used in the assay extends from −97 relative to *GAL1* to −130 relative to *GAL10*, transcriptional initiation sites. pYE(*trp5*)1 and pYE(*arg4*)511 were provided by Monica Penn. YIp302 was provided by Fred Winston.

lar factors that modify the protein, alter the activity of the protein, or directly interact with the protein.

Although this approach should be generally applicable, the ability to demonstrate specific DNA binding activity depends on a number of factors. As is the case with other assays, both the absolute binding constant and the relative affinity for specific as compared with nonspecific DNA sequences are critical. However, unlike the standard DNA binding assays that use ^{32}P-DNA, it is difficult to use carrier DNA as a competitor for nonspecific binding. Because the protein is labeled, all DNA–protein complexes are observed, and their electrophoretic mobilities are unpredictable. For the GCN4 protein, our measurement for the binding constant is about 10^{-10} M, which is about average for DNA binding proteins. Our estimate for the relative binding affinity for specific as compared with nonspecific DNA is between 100 and 1000, a typical ratio for DNA binding proteins. However, it should be noted that specific binding will be difficult to detect if the ratio is decreased by a factor of 5. Finally, a general consideration is that the protein synthesized in vitro may not be active either because its structure is different from the native protein or because a critical cofactor is missing.

Correlations between the GCN4 Binding Activity and Its Role In Vivo

Previous genetic studies suggest that the most direct step involved in the transcriptional regulation by the general control system is carried out by the GCN4 gene product (Hinnebusch and Fink, 1983b). Evidence has been presented here demonstrating that the GCN4 protein is a specific DNA binding protein that interacts with *HIS3* regulatory sequences as well as the promoter regions of coregulated genes. Moreover, the C-terminal 40 amino acids of the GCN4 protein appear to be critical for the DNA binding activity in vitro and the transcriptional regulatory ability in vivo.

Analysis of truncated proteins indicate a direct correlation between DNA binding in vitro and transcriptional activation in vivo. In particular, the synthetic protein produced by Bst NI cleavage of the template is almost identical with the predicted protein that should be produced by a *gcn4* mutant strain carrying the *gcn4*-Δ1306 allele of Hinnebusch (1984). This strain is unable to induce the transcription of all coregulated genes in response to starvation conditions. This defect may be due to the inability of the GCN4 mutant protein to bind DNA. By analyzing a series of deletion DNAs, a 20 bp region was shown to be necessary and sufficient for binding by GCN4 protein. This corresponds precisely to the *HIS3* sequences critical for transcriptional regulation in vivo; indeed, the same derivatives analyzed for their DNA binding ability were examined for their phenotypic effects in yeast cells (Struhl, 1982b; Hill and Struhl, unpublished). Independent evidence in support of these results comes from the DNAase I protection experiments. The 10 bp region protected by GCN4 protein lies within the 20 bp segment that is critical for DNA binding. Moreover, the protected nucleotides include the TGACTC sequence, which has been directly implicated as the common recognition signal for genes for which transcription is induced under conditions of amino acid starvation (Struhl, 1982b; Donahue et al., 1983; Hinnebusch and Fink, 1983a).

Although the TGACTC sequence is necessary for GCN4 binding and interacts with the protein, it appears that maximum binding by GCN4 protein depends upon more than this hexanucleotide sequence. No binding was detected to the other perfect TGACTC sequence located further upstream at *HIS3* or to two other TGACTC sequences in the pUC8 vector into which the yeast DNA was cloned. Several considerations suggest that the 9 dT residues immediately downstream from TGACTC are of some importance. First, replacement of 6 of these dT residues with a Sac I linker reduces the transcriptional in-

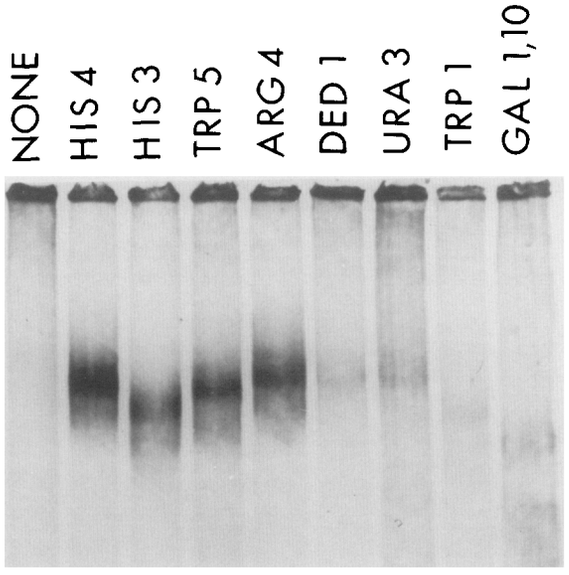


Figure 7. Binding to the Promoter Regions of Other Yeast Genes

GCN4 protein was incubated in the absence of DNA (none) or with isolated DNA fragments (50 ng) containing the upstream regions of the genes indicated (see Table 2 for a full description of these fragments), and the products were subjected to native PAGE.

duction in vivo (Hill and Struhl, unpublished results). Second, other coregulated genes have dT tracts just downstream of the TGACTC sequence, whereas the nonbinding TGACTC sequences lack such tracts. Third, DNAase I cleavage of the downstream-most dT residues are enhanced upon binding by GCN4 protein. It is of interest that poly(dA:dT) sequences have unusual structures; they have a 10.0 bp helix repeat instead of the normal 10.6 bp (Peck and Wang, 1981; Rhodes and Klug, 1981), and they are associated with kinks in DNA (Marini et al., 1982). Aspects of this unusual structure are seen in the experiments here in that the poly(dA:dT) tracts in the *HIS3* upstream region are refractory to cleavage by DNAase I (Figure 6). Thus it is possible that the binding of GCN4 protein induces a change in DNA conformation at the dT residues, and that this is important for transcriptional activation.

At least one property of transcriptional regulation in vivo has not been demonstrated in the in vitro binding experiments. Although the 20 bp region that binds GCN4 protein is necessary for regulation, it is only partially sufficient. Maximal expression in response to starvation conditions also requires the imperfect TGCCTC repeat sequence located between nucleotides −141 to −136 (Hill and Struhl, unpublished). It is worth noting that a 15 out of 17 stretch of dT residues occurs close to and downstream from this imperfect repeat. However, although DNA fragments containing this sequence bind somewhat better than nonspecific DNA, we have yet to demonstrate whether this sequence is responsible. Perhaps the binding of these sequences in vivo depends on cooperativity of GCN4 binding, on cofactors produced during the starvation response, or on a separate protein.

A Molecular Mechanism for the General Control of Amino Acid Biosynthetic Genes

To induce transcription of the co-regulated genes under appropriate conditions, a cell must interpret the physiological state of starvation to produce a molecular signal, and it must transmit this signal to the effector molecule. The experiments described here strongly implicate GCN4 protein as the effector molecule. Specifically, we have shown that the *GCN4* gene, the gene product of which is necessary for the coordinate induction of amino acid biosynthetic genes, encodes a protein that binds specifically to DNA sequences that are critical for this regulation. Thus, the simplest hypothesis is that GCN4 protein is a specific transcription factor for genes that contain the TGACTC binding sites.

How is the stimulus of amino acid limitation transmitted to the GCN4 effector molecule? GCN4 protein levels are dramatically higher in starved cells as compared with normal cells because of increased translation of GCN4 mRNA (Thireos et al., 1984; Hinnebusch, 1984). The structure of *GCN4* mRNA is atypical in that the 5′ untranslated leader is almost 600 bases, and it includes four AUG codons that cannot initiate GCN4 protein synthesis (Thireos et al., 1984; Hinnebusch, 1984). Thus the normal rules of translation preclude the synthesis of GCN4 protein, thereby providing a simple explanation for why the coregulated genes are normally expressed at a basal level. However, under conditions of amino acid starvation, the basic translation rules must be altered in some manner such that GCN4 protein is synthesized from its unusual mRNA. Therefore transcriptional induction of amino acid biosynthetic genes is mediated, at least in part, by the intracellular level of the specific DNA binding protein that recognizes the TGACTC regulatory sequences in the relevant promoter regions.

The translational control mechanism is of interest because the starvation signal is probably produced during the translation process itself. Translation is the only metabolic event for which all 20 amino acids are required, which would explain why starvation for any single amino acid results in the same transcriptional induction. The likelihood that translation is involved in the production of the starvation signal as well as its mode of transmission to the ultimate effector, GCN4 protein, suggests that these aspects of general control may be related mechanistically.

It is tempting to speculate that the *GCD1* gene product, which is necessary for cell viability and is required for protein synthesis (Wolfner et al., 1975; Hill and Struhl, unpublished data) is involved in both of these aspects. The *gcd1*-1 mutation results in induced transcriptional levels of all coregulated genes even under normal growth conditions (Wolfner et al., 1975) and in increased levels of GCN4 protein (Hinnebusch, 1984). This effect is not due to artifactual starvation because *gcd1*-1 strains have normal tRNA charging levels for all 20 amino acids (Hill and Struhl, unpublished data). Consequently, it suggests that the *GCD1* gene product is involved in the translational control mechanism. For example, GCD1 protein could sense the translationally stalled ribosomes that undoubtedly occur during starvation because of lack of available charged tRNA spe-

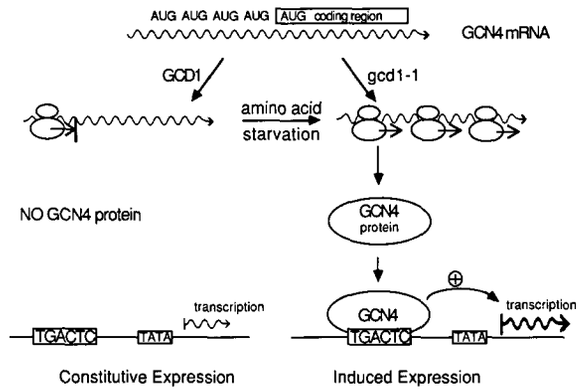


Figure 8. Model for the General Control of Amino Acid Biosynthetic Genes

See text.

cies and respond by releasing the block to translational reinitiation at the critical AUG codon of GCN4 protein.

In Figure 8, these considerations are summarized in a molecular model, aspects of which have been proposed previously (Hinnebusch and Fink, 1983b; Thireos et al., 1984; Hinnebusch, 1984). An appealing feature of the model is that the molecular signal is produced and transmitted to the effector molecule by the basic translation process. The experiments presented here demonstrate that GCN4 protein is the effector molecule and thus provide a link between the genetic and molecular studies of the coordinate transcriptional regulation of the amino acid biosynthetic genes.

Experimental Procedures

Cloning the *GCN4* Protein Coding Sequences into the pSP64 Vector

The *GCN4* gene used in these experiments was isolated by complementation of a *gcn4* mutation, essentially as described previously (Penn et al., 1983; Hinnebusch and Fink, 1983b). Specifically, a collection of DNA molecules, each containing a segment of yeast DNA generated by partial digestion with Sau 3A cloned into the Bam HI site of the *URA3* vector YCp50, was introduced into strain KY484 (a *ura3*-52 *gcn4*-1), and Ura$^+$ transformants that grew in the presence of 10 mM aminotriazole were selected. Restriction endonuclease mapping of plasmid DNAs obtained from four such transformants indicated that they contained the *GCN4* gene. The relevant 2.5 kb Bst EII-Bgl II DNA fragment was treated successively with Bal 31 nuclease, the large fragment of DNA polymerase I (in the presence of all four deoxynucleotide triphosphates), and Eco RI, and the resulting fragments were cloned between the Eco RI and Sma I sites of the mp18 vector (Messing, 1983). Following DNA sequence analysis (Sanger et al., 1980), a derivative that extended from 27 bp upstream of the *GCN4* translational initiation codon to the Eco RI site (roughly 1.4 kb downstream from the termination codon) (Thireos et al., 1984; Hinnebusch, 1984) was recloned into the pSP64 vector (Melton et al., 1984) such that the protein coding sequences were correctly oriented and just downstream from the promoter for SP6 RNA polymerase. This molecule is called pSP64-GCN4.

In Vitro Transcription and Translation

pSP64-GCN4 and pSP64 DNAs, purified by equilibrium centrifugation in CsCl, were cleaved with the appropriate restriction enzyme. Using 0.5 μg of the DNA template, transcription in vitro by SP6 RNA polymerase was carried out in a 25 μl reaction as described by Melton et al. (1984) except that the concentration of GTP was 50 μM, and diguanosine triphosphate (G-5′ppp5′-G)(P-L Biochemicals) was added to 500 μM. This modification, suggested by Doug Melton, generates 5′-capped mRNA in a single reaction; such mRNA is translated more efficiently (Krieg and Melton, 1984). The total RNA generated in such a reaction (roughly 1 μg) was extracted with phenol and was precipitated with ethanol. Translation in vitro was carried out in a reaction volume of 30 μl with 16 μCi of ^{35}S-methionine (1400 Ci/mmol) using a wheat germ extract as directed by the manufacturers (Bethesda Research Laboratories).

Translation products were examined by electrophoresis in a 15% polyacrylamide gel (30 parts acrylamide: 0.8 parts bisacrylamide) containing SDS, as described by Laemmli (1970). A ^{14}C-methylated protein mixture (Bethesda Research Laboratories) was used for molecular weight standards. After electrophoresis, the gel was fixed, treated with en^{3}Hance (New England Nuclear) and autoradiographed. To measure incorporation of ^{35}S-methionine into protein, 1.0 μl of the translation products was incubated first in 50 μl 0.1 M. NaOH for 15 min at 37°C and then in 1 ml 10% trichloroacetic acid for 15 min on ice. The precipitated protein was collected on glass fiber filters, and incorporated ^{35}S-methionine was measured by scintillation counting using aquasol (New England Nuclear).

DNA Binding Assay

The DNAs to be bound were purified either by equilibrium centrifugation in CsCl, or in some instances by a rapid lysate procedure. The initial experiments used pUC8-*HIS3* DNA, which contains the 1.8 kb Bam HI fragment Sc2676 (Struhl and Davis, 1980) cloned into the pUC8 vector (Vieira and Messing, 1982) and used pUC9 DNA as a control. Also used were YIp55, pUC8-Sc3296, and pUC8-Sc4101, which are pUC8 plasmids containing the 1.1 kb Hind III *URA3* fragment (Bach et al., 1979), the 1.4 kb Eco RI-*TRP1* fragment (Struhl et al., 1979), and the 365 bp Dde I-Sau 3A *GAL1,10* fragment (Struhl, 1984), respectively. Individual DNA fragments for binding assays were isolated from 2% low-gelling temperature agarose gels (Wieslander, 1979).

The DNA binding assay was based on the procedures described by Garner and Revzin (1981) and Fried and Crothers (1981). Binding was carried out in 15 μl of 20 mM Tris (pH 7.4), 50 mM KCl, 3 mM $MgCl_2$, 1 mM EDTA, 100 μg/ml gelatin, and 50 μg/ml sonicated salmon sperm DNA (binding buffer). The DNA to be bound (rendered blunt-ended using the large fragment of E coli DNA polymerase I) was included to a final concentration of 9 nM. GCN4 protein or other translation products (0.5 μl), without further purification from the in vitro translation reaction, was then added, and equilibration was achieved by incubating at 25°C for 20 min. To the equilibration products was added 5 μl of loading buffer (binding buffer containing 20% glycerol, 1 mg/ml xylene cyanol FF, and 1 mg/ml bromophenol blue) for immediate electrophoresis. The samples were loaded on to 1.5 mm thick, 5% polyacrylamide gel (30 parts acrylamide: 0.8 parts bisacrylamide) prepared and run in 90 mM Tris-Borate buffer (pH 8.3). Electrophoresis was carried out at 400 volts until the samples had entered the gel and then at 175 volts until the bromophenol blue had migrated the length of the gel (20 cm). The gel was then fixed, treated with en^{3}Hance, and autoradiographed.

DNAase I Footprint Analysis

The DNAase I footprinting technique of Galas and Schmitz (1978) was used with modifications. DNA fragments were 3′ end-labeled using the Klenow fragment of E. coli DNA polymerase I and ^{32}P-labeled dCTP (3000 Ci/mmol). After digestion with a second restriction enzyme, the desired fragments were purified by electrophoresis in a polyacrylamide gel. Approximately 1 ng of end-labeled DNA fragment in binding buffer (without carrier DNA) was combined with various quantities of translation products (without further purification) such that the final volume was always 20 μl. For these experiments, the in vitro translation was carried out with ^{35}S-methionine at only 7 Ci/mmol (7 μCi per 30 μl reaction). This avoided problems of ^{35}S-peptides interfering with autoradiography of the ^{32}P-DNA fragments but still permitted confirmation that GCN4 protein had been synthesized. After equilibration for 20 min at 25°C, 5 μl of DNAase I (20 μg/ml in binding buffer without carrier DNA) was added. DNAase I digestion was terminated after 2 min at 25°C by the addition of 25 μl of 0.5% SDS, 10 mM EDTA, 0.5 mg/ml E. coli tRNA. The DNA was extracted with phenol, precipitated with ethanol, redissolved in 80% formamide, and analyzed as described for DNA sequencing. To relate the observed bands with positions of

DNAase I cleavage, the ^{32}P-labeled DNA fragments were cleaved with appropriate restriction endonucleases and were then treated identically to the other samples.

Acknowledgments

We thank Alan Hinnebusch and Monica Penn for fruitful conversations over the past several years, particularly about unpublished results, Doug Melton for suggesting the use of diguanosine triphosphate for producing 5′-capped mRNA, members of this laboratory for suggestions concerning the experiments, and Fred Winston and Monica Penn for gifts of DNA. I. A. H. was supported by a fellowship from the Royal Society. This work was supported by grants to K. S. from the National Institute of Health (GM 30186) and the Chicago Community Trust (Searle Scholars Program).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 17, 1985; revised August 26, 1985

References

Aebi, M., Furtur, R., Prantl, F., Niederberger, P., and Hutter, R. (1984). Structure and function of the *TRP3* gene of *Saccharomyces cerevisiae*: analysis of transcription, promoter sequence, and sequence coding for a glutamine amidotransferase. Curr. Genet. *8*, 165–172.

Bach, M. L., Lacroute, F., and Botstein, D. (1979). Evidence for transcriptional regulation of orotidine 5-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene cloned in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *76*, 386–390.

Beacham, J. R., Schweitzer, B. W., Warrick, H. M., and Carbon, J. (1984). The nucleotide sequence of the *ARG4* gene. Gene *29*, 271–279.

Clarke, L., and Carbon, J. (1978). Functional expression of cloned yeast DNA in *Escherichia coli*: specific complementation of arginosuccinate lyase (*argH*) mutations. J. Mol. Biol. *120*, 517–532.

Delforge, J., Messenguy, F., and Wiame, J. (1975). The regulation of arginine biosynthesis in *Saccharomyces cerevisiae*: the specificity of *arg*^{R} mutations and the general control of amino acid biosynthesis. Eur. J. Biochem. *57*, 231–239.

Donahue, T. F., Daves, R. S., Lucchini, G., and Fink, G. R. (1983). A short nucleotide sequence required for regulation of *HIS4* by the general control system of yeast. Cell *32*, 89–98.

Farabaugh, P., and Fink, G. R. (1980). Insertion of the eukaryotic transposable element Ty1 creates a 5-base pair duplication. Nature *286*, 352–356.

Fried, M., and Crothers, D. (1981). Equilibrium and kinetics of *lac* repressor–operator interactions by polyacrylamide gel electrophoresis. Nucl. Acids Res. *9*, 6505–6525.

Galas, D., and Schmitz, A. (1978). DNase footprinting: a simple method for the detection of protein–DNA binding specificity. Nucl. Acids Res. *5*, 3157–3170.

Garner, M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *E. coli* lactose operon regulatory system. Nucl. Acids Res. *9*, 3047–3060.

Hinnebusch, A. G. (1984). Evidence for translational regulation of the activator of general amino acid control in yeast. Proc. Natl. Acad. Sci. USA *81*, 6442–6446.

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

Hinnebusch, A. G., and Fink, G. R. (1983b). Positive regulation in the general control of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA *80*, 5374–5378.

Hinnebusch, A., Lucchini, G., and Fink, G. (1985). A synthetic *HIS4* regulatory element confers general amino acid control on the cytochrome C gene (*CYC1*) of yeast. Proc. Natl. Acad. Sci. USA *82*, 498–502.

Jones, E. W., and Fink, G. R. (1982). Regulation of amino acid biosynthesis and nucleotide biosynthesis in yeast. In The Molecular Biology of the Yeast *Saccharomyces*: Metabolism and Gene Expression. J. N. Strathern, E. W. Jones, and J. R. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 181–299.

Kimelman, D., Lucher, L. A., Brackmann, K. H., Symington, J. S., Ptashne, M., and Green, M. (1984). Synthesis in *Escherichia coli* of human adenovirus type 12 transforming proteins encoded by early region 1A 13S mRNA and 12S mRNA. Proc. Natl. Acad. Sci. USA *81*, 6300–6304.

Krieg, P. A., and Melton, D. A. (1984). Functional messenger RNAs are produced by SP6 in vitro transcription of cloned DNAs. Nucl. Acids Res. *12*, 7057–7070.

Laemmli, U. (1970). Cleavage of the structural proteins during assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Marini, J. C., Levene, S. D., Crothers, D. M., and Englund, P. T. (1982). A bent helix in kinetoplast DNA. Cold Spring Harbor Symp. Quant. Biol. *47*, 279–283.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res. *12*, 7035–7056.

Messing, J. (1983). New M13 vectors for cloning. Meth. Enzymol. *101*, 20–79.

Messenguy, F., Feller, A., Crabeel, M., and Pierard, A. (1983). Control mechanisms acting at the transcriptional and post-transcriptional levels are involved in the synthesis of the arginine pathway carbamoylphosphate synthase of yeast. EMBO J. *2*, 1249–1254.

Peck, L. J., and Wang, J. C. (1981). Sequence dependence of the helical repeat of DNA in solution. Nature *292*, 375–378.

Penn, M. D., Galgoci, B., and Greer, H. (1983). Identification of *AAS* genes and their regulatory role in general control of amino acid biosynthesis in yeast. Proc. Natl. Acad. Sci. USA *80*, 2704–2708.

Rhodes, D., and Klug, A. (1981). Sequence dependent helical periodicity of DNA. Nature *292*, 378–380.

Sanger, F., Coulson, A. R., Barell, B. G., Smith, A. J., and Roe, B. A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. *143*, 161–178.

Schurch, A., Miozzari, G., and Hutter, R. (1974). Regulation of tryptophan biosynthesis in *Saccharomyces cerevisiae*: mode of action of 5-methyl-tryptophan and 5-methyl-tryptophan sensitive mutants. J. Bacteriol. *117*, 1131–1140.

Strauss, F., and Varshavsky, A. (1984). A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell *37*, 889–901.

Struhl, K. (1981). Deletion mapping a eukaryotic promoter. Proc. Natl. Acad. Sci. USA *78*, 4461–4465.

Struhl, K. (1982a). The yeast *his3* promoter contains at least two distinct elements. Proc. Natl. Acad. Sci. USA *79*, 7385–7389.

Struhl, K. (1982b). Regulatory sites for *his3* expression in yeast. Nature *300*, 284–287.

Struhl, K. (1984). Genetic properties and chromatin structure of the yeast *gal* regulatory element: an enhancer-like sequence. Proc. Natl. Acad. Sci. USA *81*, 7865–7869.

Struhl, K., and Davis, R. W. (1980). A physical, genetic, and transcriptional map of the cloned *his3* gene region of *Saccharomyces cerevisiae*. J. Mol. Biol. *136*, 309–332.

Struhl, K., and Davis, R. W. (1981). Transcription of the *his3* gene region in *Saccharomyces cerevisiae*. J. Mol. Biol. *152*, 535–552.

Struhl, K., Stinchcomb, D. T., Scherer, S., and Davis, R. W. (1979). High frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA. *76*, 1035–1039.

Thireos, G., Penn, M. D., and Greer, H. (1984). 5′ untranslated sequences are required for the translational control of a yeast regulatory gene. Proc. Natl. Acad. Sci. USA *81*, 5096–5100.

Vieira, J., and Messing, J. (1982). The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene *19*, 259–268.

Walz, A., Ratzkin, B., and Carbon, J. (1978). Control of expression of a cloned yeast (*Saccharomyces cerevisiae*) gene (*trp5*) by a bacterial insertion element (IS2). Proc. Natl. Acad. Sci. USA *75*, 6172–6176.

Weislander, L. (1979). A simple method to recover intact high molecular weight RNA and DNA after electrophoretic separation in low gelling temperature agarose gels. Anal. Biochem. *98*, 305–309.

Wolfner, M., Yep, D., Messenguy, F., and Fink, G. R. (1975). Integration of amino acid biosynthesis into the cell cycle of *Saccharomyces cerevisiae*. J. Mol. Biol. *96*, 273–290.

Zalkin, H., and Yanofsky, C. (1982). Yeast gene *TRP5*: structure, function, regulation. J. Biol. Chem. *257*, 1491–1500.