MOLECULAR AND CELLULAR BIOLOGY, Mar. 1989, p. 1120–1127
0270-7306/89/031120-08$02.00/0

Vol. 9, No. 3

cys-3, the Positive-Acting Sulfur Regulatory Gene of *Neurospora crassa*, Encodes a Protein with a Putative Leucine Zipper DNA-Binding Element

YING-HUI FU,[1] JOHN V. PAIETTA,[2] DANIEL G. MANNIX,[1] AND GEORGE A. MARZLUF[1]*

Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210,[1] and Department of Biochemistry, Wright State University, Dayton, Ohio 45435[2]

Received 3 October 1988/Accepted 8 December 1988

The sulfur-regulatory circuit of *Neurospora crassa* consists of a set of unlinked structural genes which encode sulfur-catabolic enzymes and two major regulatory genes which govern their expression. The positive-acting *cys-3* regulatory gene is required to turn on the expression of the sulfur-related enzymes, whereas the other regulatory gene, *scon*, acts in a negative fashion to repress the synthesis of the same set of enzymes. Expression of the *cys-3* regulatory gene was found to be controlled by *scon* and by sulfur availability. The nucleotide sequence of the *cys-3* gene was determined and can be translated to yield a protein of molecular weight 25,892 which displays significant homology with the oncogene protein Fos, yeast *GCN4* protein, and sea urchin histone H1. Moreover, the putative *cys-3* protein has a well-defined leucine zipper element plus an adjacent charged region which together may make up a DNA-binding site. A *cys-3* mutant and a *cys-3* temperature-sensitive mutant lead to substitutions of glutamine for basic amino acids within the charged region and thus may alter DNA-binding properties of the *cys-3* protein.

In the filamentous fungus *Neurospora crassa*, a high degree of genetic and metabolic regulation governs the expression of sets of enzymes within various global areas of metabolism such as nitrogen, phosphorus, and sulfur catabolism (7, 18). The sulfur control circuit of *N. crassa* consists of a set of unlinked structural genes which specify enzymes involved in sulfur metabolism. Synthesis of this entire family of sulfur-related enzymes, which includes aryl sulfatase, choline sulfatase, sulfate permease, a high-affinity methionine permease, and an extracellular protease occurs only when cellular levels of sulfur become limited (10, 16, 19, 22). The expression of these sulfur-catabolic enzymes is controlled by two distinct regulatory genes. One of these, designated *scon* (for sulfur controller), appears to act in a negative fashion; *scon* mutants are insensitive to sulfur catabolite repression and thus express the sulfur-related enzymes in a constitutive fashion (4). The other sulfur-regulatory gene, known as *cys-3*, acts in a positive manner to activate the expression of the various sulfur-related genes (17, 21).

Two different sulfate permease species, specified by distinct and unlinked structural genes, are both members of the sulfur circuit (16). The structural gene for sulfate permease II, *cys-14*, has been cloned and shown to encode an mRNA of approximately 3 kilobases (kb) whose content is highly regulated by *cys-3*, by *scon*, and by the sulfur status of the cells (11). Thus, it appears that both of the regulatory genes as well as sulfur repression act at the level of transcription or at a closely related step such as mRNA processing or stability. The *cys-3* major sulfur control gene has been postulated to encode a regulatory protein which is needed to turn on the expression of the entire set of sulfur-related activities, presumably by binding at target DNA sequences adjacent to each structural gene (11). Mutants containing mutations of *cys-3* lack all of the sulfur-related enzymes and cannot use various secondary sulfur sources which are readily utilized by *cys-3*$^+$ strains; revertants of *cys-3* regain these various activities in a single step. Moreover, temperature-sensitive *cys-3* mutants have been studied which have a *cys-3*$^+$ phenotype at 25°C but which behave as tight *cys-3* mutants at 37°C, consistent with the possibility that *cys-3* specifies a regulatory protein (17). The *cys-3*$^+$ regulatory gene was recently cloned via complementation by using the sib selection procedure, and its identity was confirmed by restriction fragment length polymorphism analyses (21). Expression of the *cys-3* regulatory gene itself was found to be subject to a high degree of regulation, such that it is expressed at a considerably higher rate upon sulfur derepression, yielding two transcripts, of approximately 1.3 and 1.6 kb.

We report here the entire nucleotide sequence of the *cys-3* regulatory gene and demonstrate that transcription begins at several closely spaced sites just upstream of a putative AUG initiation codon. We show that the *cys-3* gene can be translated to yield a protein of molecular weight 25,892, composed of 236 amino acids. We also describe the use of the polymerase chain reaction to clone a *cys-3* mutant gene and a *cys-3* temperature-sensitive gene and present the alteration which occurs in each of these mutations. These two mutations affect neighboring basic amino acid residues which occur in a highly charged region of the protein that lies immediately adjacent to a putative leucine zipper structure. Finally, experiments are presented which demonstrate that the *scon* gene in fact controls the expression of *cys-3*.

MATERIALS AND METHODS

Strains. The *N. crassa* wild-type strain 74OR231A and the *cys-3* mutant (allele P22) were obtained from the Fungal Genetics Stock Center, University of Kansas Medical Center, Kansas City. A *cys-3* temperature-sensitive revertant (allele 65t) was described previously (17), and two *scon*^{c} mutants (alleles 36-21 and 36-28) were obtained from R. L. Metzenberg. Cultures were grown in Vogel liquid medium

* Corresponding author.

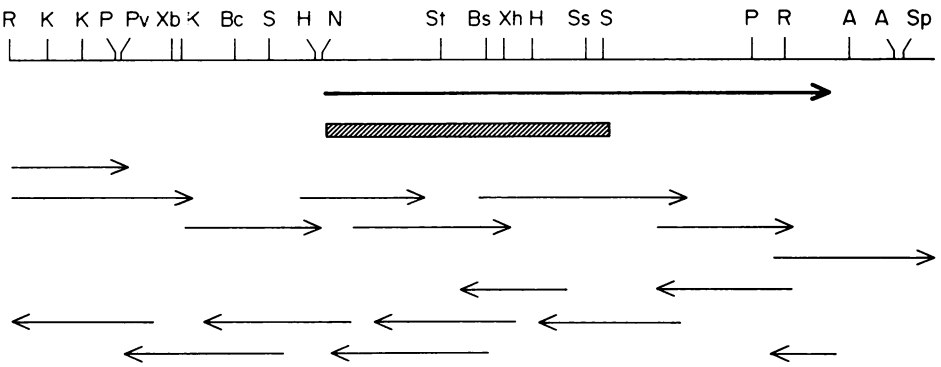


FIG. 1. *cys-3* gene structure and sequencing strategy. The direction of transcription is indicated by the bold arrow, which represents the 1.3-kb transcript. Symbols: ▨, predicted *cys-3* protein composed of 236 amino acid residues; ←, →, sequencing strategy that was achieved with deletion clones and with oligonucleotide primers so that both strands were sequenced. Restriction sites: A, *Aha*II; Bc, *Bcl*I; Bs, *Bst*EII; R, *Eco*RV; H, *Hin*dIII; K, *Kpn*I; N, *Nco*I; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Sp, *Sph*I; Ss, *Sst*I; St, *Stu*I; Xb, *Xba*I; Xh, *Xho*I.

supplemented as indicated for each experiment with shaking at 30°C as described previously (7, 11).

DNA sequencing and S1 nuclease mapping. DNA sequencing was accomplished by the dideoxy-chain termination method (28) with [α-^{32}P]dATP and a modified T7 bacteriophage DNA polymerase, Sequenase (United States Biochemical Corp.). dITP was successfully used in place of dGTP to sequence through compression regions. Plasmid DNA templates were prepared as minipreps (3). Deletion clones of pCys3 for sequencing were constructed by using exonuclease III and mung bean nuclease as described previously (8). Oligonucleotide primers for DNA sequencing were synthesized on an Applied Biosystems model 380B DNA synthesizer at the Ohio State University Biochemical Instrument Center. An *Xba*I-*Xho*I restriction fragment was end labeled at the *Xho*I site with [α-^{32}P]ATP via T4 polynucleotide kinase and isolated from an agarose gel for S1 nuclease mapping (2). Primer extension experiments were performed by mixing a 5′-end-labeled 17-mer oligonucleotide primer that hybridizes at position +56 (see Fig. 2) with poly(A)$^+$ RNA as a template for Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories, Inc.).

Site-directed mutagenesis. Site-directed mutagenesis was performed as described by Kunkel (13). The *Eco*RV fragment of pCYS3 was cloned into vector pTZ18U, and single-stranded plasmid DNA (pCys3-18U), obtained by use of helper phage, was used as the template for mutagenesis. A single base change was engineered into a possible TATA box element, changing it from TATATCA to GATATCA, by using the mutagenic primer CCTGTCTTGGATATCAG (altered base underlined). This change results in the generation of an *Eco*RV restriction site (GATATC). Of 12 potential mutant clones examined, 9 (75%) had gained the expected *Eco*RV site. One of these was sequenced to confirm its identity before use.

Cloning mutant *cys-3* genes. The polymerase chain reaction was used to amplify the *cys-3* gene present in total *Neurospora* genomic DNA of two mutant strains in order to clone them (27). Neurospora DNA was isolated from a *cys-3* null mutant and a *cys-3* temperature-sensitive mutant as described previously (7). A 29-mer oligonucleotide primer was used which hybridized at 25 bases from −68 to −44 at the 5′ end of the gene; the additional 4 bases at the 5′ end of the primer were included to yield an *Eco*RI site. Similarly, a 29-mer that hybridized to the opposite strand at 27 bases from +1116 to +1090 was used at the 3′ end of the gene, with the additional 2 bases providing a *Bgl*II site. Each polymerase chain reaction took place in a 1.5-ml microcentrifuge tube containing 1.3 μg of *Neurospora* DNA, 1 μg of each primer, and 5 U of *Thermus aquaticus* DNA polymerase (Perkin-Elmer Cetus Corp.) in a total volume of 100 μl covered with 100 μl of mineral oil. After 25 cycles, each consisting of denaturation (2 min at 94°C), hybridization (2.5 min at 37°C), and elongation (6 min at 72°C), a DNA band of the expected size, 1.2 kb, was readily visible after agarose gel electrophoresis. The amplified 1.2-kb *Eco*RI-*Bgl*II fragment for each *cys-3* mutant was isolated from the gel and cloned into *Eco*RI-*Bam*HI-digested Bluescript vector.

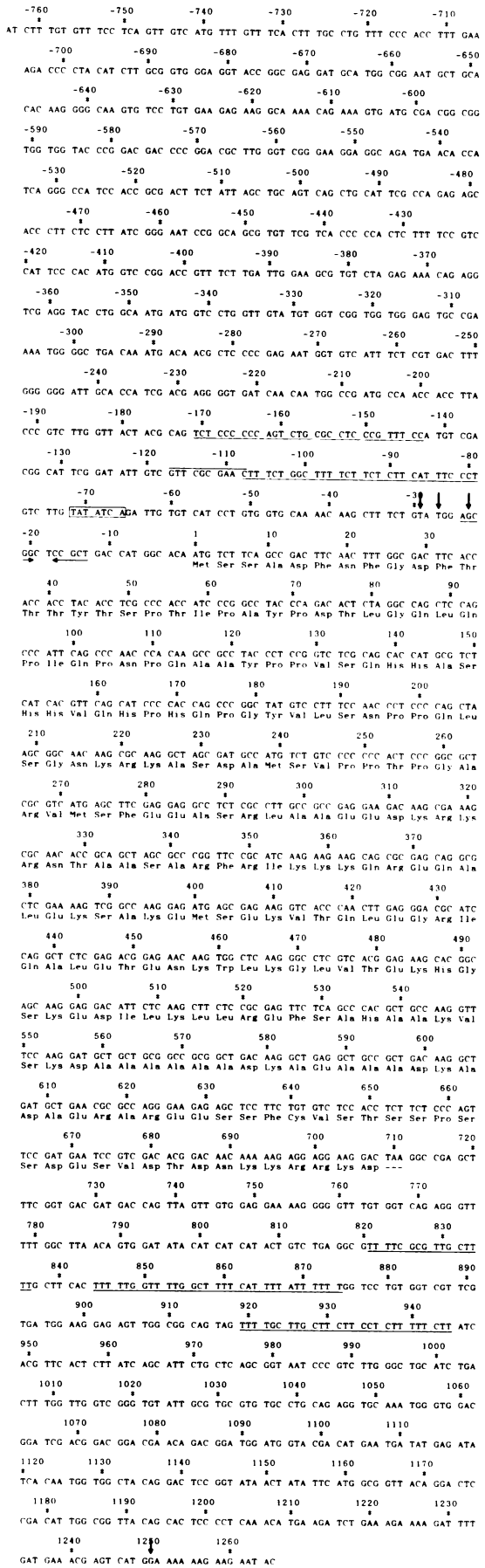
Poly(A)$^+$ RNA isolation and Northern blot analyses. *N. crassa* RNA was isolated by the method of Reinert et al. (24) from cells grown with either high sulfur (5 mM methionine) or low sulfur (0.25 mM methionine). The poly(A)$^+$ RNA fraction was isolated with oligo(dT)-cellulose (1). Formaldehyde-agarose gel electrophoresis and Northern (RNA) blots were carried out as described by Maniatis et al. (15), by following the prehybridization and hybridization protocols described before (7). Plasmid DNAs for use as probes were labeled with [^{32}P]dCTP nick translation (25).

Isolation and analysis of cDNA clones. *N. crassa* poly(A)$^+$ RNA was prepared as described above, except that it was passed through oligo(dT)-cellulose twice prior to use for cDNA synthesis. A cDNA library was constructed in lambda gt10 by using *Eco*RI adaptors to eliminate the need for methylation of the double-stranded cDNA and *Eco*RI digestion prior to ligation into the vector arms, followed by packaging with Packagene (Promega Biotec). Four rounds of plaque hybridization with pCys3 as a probe were required to isolate several *cys-3* cDNA clones (15). Lambda DNA was purified, digested with *Eco*RI, and the insert cDNA was subcloned into the Bluescript plasmid vector and sequenced as described above.

Computer methods. The handling of sequences, their analysis, translation, and hydropathy and codon bias analyses were accomplished with Pustell software (International Biotechnology, Inc.). Protein homology searches were conducted with GenBank, the Protein Identification Resource (National Biomedical Research Foundation), which contains 3,800 different protein sequences.

RESULTS

***cys-3*$^+$ nucleotide sequence.** A restriction map of the *cys-3* gene and the strategy used to sequence it are shown in Fig.



1. The nucleotide sequence for both strands was determined by using the dideoxy sequencing technique for the entire *cys-3* gene, with the use of overlapping clones to confirm the sequence of each segment. The nucleotide sequence of the *cys-3* gene and its flanking regions is presented in Fig. 2.

Initiation of *cys-3* transcription. The 1.88-kb *Eco*RV fragment shown in Fig. 1 transforms the *cys-3* mutant strain at a high frequency. The direction of transcription as shown in Fig. 1 was determined by using single-stranded probes in Northern blot experiments (results not shown). To help to define the 5′ border of the gene, the *Eco*RV fragment was cut with various restriction enzymes and then assayed for its transforming capacity. When cut with *Pvu*II, *Kpn*I, or *Bcl*I, the cloned DNA was still *cys-3*$^+$, but restriction with *Nco*I resulted in loss of its transforming ability, suggesting that the 5′ border of the gene was located between the *Bcl*I and *Nco*I sites. Various restriction fragments used as probes in Northern blot experiments indicated that the 5′ end of the gene occurred near the *Hin*dIII site at position −41. An S1 nuclease mapping experiment showed that the 5′ end of the *cys-3* transcript was located at approximately position −30 (Fig. 3). This result was confirmed by a primer extension assay which revealed that transcription of *cys-3* is initiated at three closely spaced sites, at −29, −27, and −23 (Fig. 2 and 3). The complete oligonucleotide sequence of a 1.3-kb *cys-3* cDNA was determined and demonstrated that the 1.3-kb *cys-3* transcript contained no introns and identified the approximate location of its 5′ and 3′ ends (Fig. 2).

Analysis of a potential TATA box. A possible TATA box, TATATCA, is located approximately 40 bases upstream of the 5′ start sites (Fig. 2). To determine whether this sequence was essential for *cys-3* expression, site-directed mutagenesis was used as described in Materials and Methods to change the initial T into a G, yielding an easily assayed *Eco*RV restriction site (GATATC). Of 12 potential mutant clones, 9 had gained the expected *Eco*RV site, giving a mutagenic yield of 75%. One of these was sequenced to confirm its identity, and a transformation assay was used to determine whether the altered gene was still functional. The potential *cys-3* TATA box mutant gene transformed the *cys-3* mutant strain at a high frequency, identical to that observed with the positive control, a completely normal *cys-3*$^+$ gene. In *Neurospora* spp., most transformants result from the integration of the transforming DNA at nonhomologous sites in the genome, and thus their functional expression depends upon the insertion of an intact gene (5, 6). Therefore, these results imply that the TATATCA sequence is not required for the expression of the *cys-3* gene and thus is not an essential TATA box element.

Translation of the *cys-3* sequence. The initiation of transcription occurs just upstream of an ATG codon (designated +1 in Fig. 2), which begins a long open reading frame. The other five reading frames are all constantly interrupted by termination codons. The ATG codon at +1 has a favorable context as an initiation codon, particularly the presence of C and A at −4 and −3, respectively. A more 5′ ATG codon occurs at −8, but it has a very poor context and does not

FIG. 2. Nucleotide sequence of *cys-3* and its flanking regions. The ATG initiation codon is numbered +1. The translated amino acid sequence is shown beneath the DNA sequence. Vertical arrows indicate 5′ and 3′ termini. A 10-base symmetrical sequence at −112 is overlined, a potential TATA box is boxed, pyrimidine-rich tracks are underlined, and a potential hairpin loop structure is indicated by horizontal arrows.

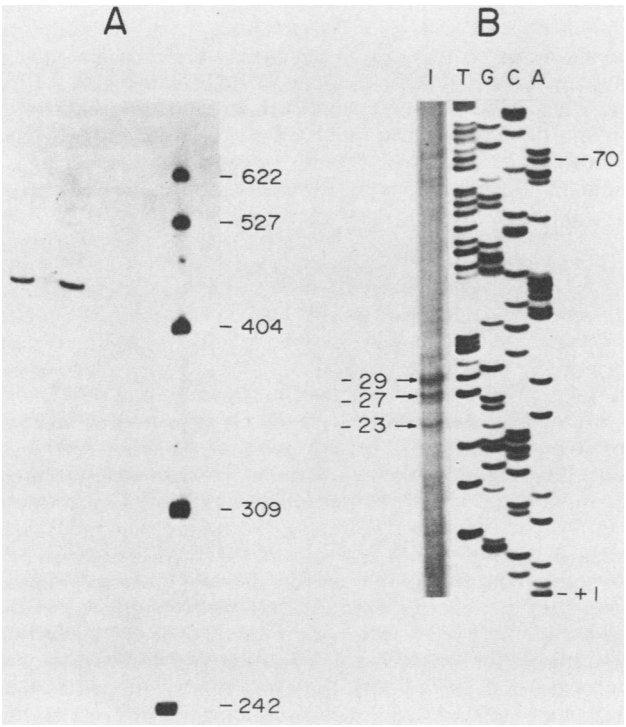


FIG. 3. S1 nuclease and primer extension mapping. (A) The approximate location of the 5′ end of *cys-3* transcripts was determined by S1 nuclease mapping. An 810-base *Xho*I-*Xba*I DNA restriction fragment, ^{32}P end labeled at the *Xho*I site (located at +445), was used for S1 nuclease mapping as described in Materials and Methods. Following the S1 nuclease treatment, electrophoresis revealed a protected fragment (in the two left lanes) estimated to be 475 bases (by comparison with the DNA molecular weight markers in the right lane), indicating that the *cys-3* mRNA 5′ terminus occurs at approximately −30. (B) Primer extension mapping of the 5′ end of *cys-3* transcripts was accomplished as described in Materials and Methods with a 17-mer oligonucleotide that hybridizes at +56. Next to the extension products in lane 1 are shown dideoxy sequencing ladders (lanes T, G, C, and A) of *cys-3* DNA primed with the same oligonucleotide used in the primer extension (lane 1). The 5′ start sites occur at −29, −27, and −23.

begin an open reading frame, indicating that it is not an initiation codon, although it might have some role in translational control of the *cys-3* mRNA. These features strongly imply that the ATG at +1 is the initiation codon for a *cys-3* encoded protein. This coding region, which lacks any intervening sequences, translates to give a protein composed of 236 amino acids with a calculated molecular weight of 25,892. Codon usage for the *cys-3* encoded protein shows a significant bias, as has been found with other *Neurospora* proteins (20, 29), e.g., a predominant use of codons ending in either C or T, especially C, and an almost complete exclusion of codons ending in A. A protein-coding region locator program, which incorporates the coding bias typical of *Neurospora* proteins, demonstrated that the open reading frame which specifies the *cys-3* protein is highly favored as a coding region and that no other reading frame or other segment of the entire sequenced region shows any significant bias. The putative *cys-3* protein has a number of interesting features, including the fact that alanine and serine make up 25% of its total residues, many of which occur in a large cluster near the carboxyl end of the protein. The amino terminus of the protein is deficient in charged amino acids

```
WT          -GAA-GAC-AAG-CGA-AAG-
                     LYS-ARG

CYS-3⁻      -GAA-GAC-CAG-CAA-AAG-
                     GLN GLN

CYS-3 ts    -GAA-GAC-AAG-CAA-AAG-
  REV                LYS GLN
```

FIG. 4. Nucleotide and amino acid substitutions in *cys-3* mutants. A *cys-3* null mutant gene and a temperature-sensitive (ts) revertant (REV) *cys-3* mutant gene were each cloned, and their nucleotide sequences were determined (see Materials and Methods). The two mutational changes which occur at nucleotides 313 and 317 are underlined, and the amino acid substitutions in the translated *cys-3* proteins are displayed. The wild-type (WT) lysine and arginine residues occupy amino acids 105 and 106, respectively, in the translated *cys-3*$^{+}$ protein.

and is followed by six histidines which occur in closely spaced pairs, possibly constituting a ligand-binding site. Nearly all (15 of 16) of the proline residues occur in the amino-terminal one-third of the protein. The carboxy terminal two-thirds of the molecule contains many charged residues which occur in clusters throughout the length of the protein. A hydropathy plot (data not shown) revealed that the putative *cys-3* protein is essentially hydrophilic throughout its entire length.

***cys-3* mutant genes.** Both *cys-3* null mutants and *cys-3* temperature-sensitive alleles have been described. The polymerase chain reaction was used to amplify the *cys-3* gene in samples of total genomic DNA in order to clone one of each of these mutant types, as described in Materials and Methods. An oligonucleotide primer which hybridized at 25 bases from −68 to −44 at the 5′ end of the *cys-3* gene and a second primer which annealed at the 3′ end to the opposite strand at 27 bases from +1116 to +1090 were used. Following 25 cycles of the polymerase chain reaction to amplify the specific sequence, a DNA band of the expected size, 1.2 kb, was readily visible after agarose gel electrophoresis. In each case the amplified 1.2-kb DNA fragment was isolated and cloned into the Bluescript plasmid vector, thus representing the molecular cloning of both a *cys-3* mutant gene and a *cys-3* temperature-sensitive gene. The entire nucleotide sequence of each of these mutants was determined and found to be identical to the wild-type *cys-3* gene, except for single base changes which affect codons 105 and 106 (Fig. 4). In the *cys-3* null mutant, two nucleotide base changes result in the replacement of lysine-105 and arginine-106 by glutamine residues. The *cys-3* temperature-sensitive revertant, derived from the *cys-3* mutant, also has glutamine in place of arginine at residue 106, but has restored the wild-type amino acid (lysine) at residue 105.

Regulation of *cys-3* gene expression. It was previously demonstrated that the *cys-3* regulatory gene itself is subject to regulation such that the cellular content of *cys-3* mRNA increases substantially upon sulfur limitation, the very condition which leads to expression of the entire set of sulfur-related enzymes (21). The sulfur control circuit contains two completely distinct control genes, the positive-acting *cys-3* gene and the negative-acting *scon* gene. This feature suggested the possibility that *scon* controls *cys-3* expression, which then activates the various structural genes. To examine this possibility, a Northern blot analysis was carried out

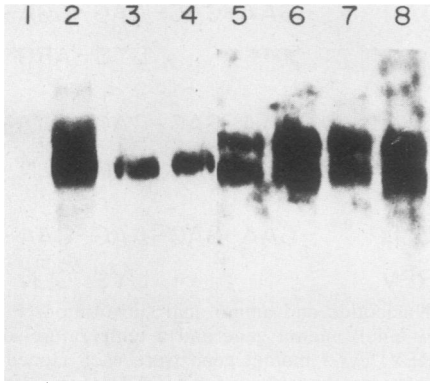


FIG. 5. Northern (RNA) gel blot analysis of *cys-3* gene expression in wild-type and *scon* mutant strains. Poly(A)$^+$ RNA (20 μg) was electrophoresed in each lane of gels, blotted to nitrocellulose, and hybridized to ^{32}P-labeled DNA probes. Lanes: 1 and 3, RNA from S-repressed wild-type cells; 2 and 4, RNA from S-limited wild-type cells; 5, RNA from S-repressed *scon* mutant 36–21; 6, RNA from S-limited *scon* mutant 36-21; 7, RNA from S-repressed *scon* mutant 36–28; 8, RNA from S-limited *scon* mutant 36-28. Lanes 1, 2, and 5 to 8 were probed with labeled pCys3 DNA. Lanes 3 and 4 were probed with a β-tubulin gene probe to demonstrate that equal amounts of RNA were loaded for both S-repressed and S-limited wild-type cells. Exposure times were 10 h for lanes 3 and 4 and 120 h for the others.

with wild-type (*scon*$^+$) and two different *scon*^{c} mutant strains. Unlike the wild type, the *scon*^{c} mutants both express aryl sulfatase and related enzymes constitutively, even during full sulfur repression conditions. The Northern analysis revealed that the wild type contains only a very small amount of *cys-3* mRNA when under sulfur repression conditions, whereas two *cys-3* transcripts of approximately 1.3 and 1.6 kb are readily evident in cells subject to sulfur limitation (Fig. 5). In sharp contrast, two different *scon*^{c} mutants possess a substantial amount of both the 1.3- and 1.6-kb *cys-3* transcripts whether under sulfur-repressing or sulfur-limited conditions; i.e., *cys-3* expression is insensitive to sulfur repression in *scon*^{c} strains (Fig. 5).

DISCUSSION

Nucleotide sequence of the *cys-3* gene. The complete nucleotide sequence of the *cys-3* gene and flanking regions reveals some interesting features. Situated in the 5′ upstream region are two pyrimidine-rich tracts: one is centered at −155, and 24 of 29 bases are T or C; the second tract, centered at −95, has a pyrimidine in 27 of 30 positions. Similarly, the 3′ noncoding region of the *cys-3* transcript is unusually rich in T residues and has three particularly notable T-rich tracts (Fig. 2). The significance of these pyrimidine-rich and T-rich tracts is unknown, but is clear that such monotonous sequences are not found in the protein-coding regions. Fungal genes may contain one or more TATA boxes, which are not located at any fixed position but may lie as far as 100 bases upstream of the transcription start site (9). Moreover, some fungal genes lack any apparent TATA box altogether (23, 29), which appears to be the situation for *cys-3*, since the site-directed mutagenesis study suggested that the only good TATA box candidate was not a completely essential element. However, it is conceivable that disruption of a TATA box could significantly reduce transcription without yielding

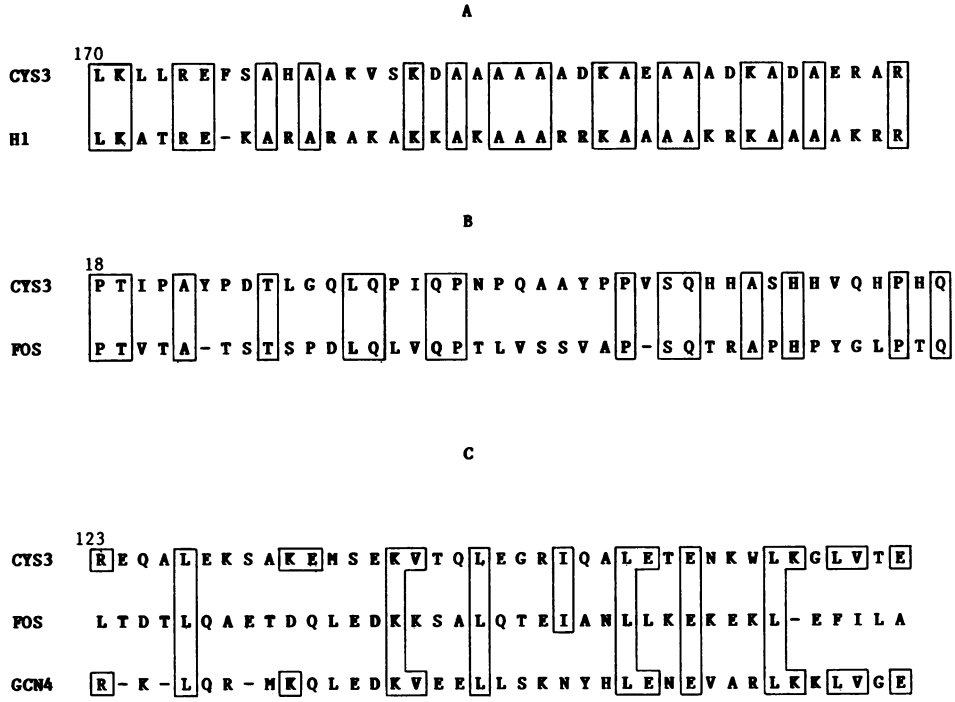


FIG. 6. *cys-3* translated protein homologous domains. Identification and alignment of homologous regions of the translated *cys-3* with other proteins was done by using protein analysis programs (see Materials and Methods). Identical amino acids are boxed. (A) Residues 170 to 208 of *cys-3* protein show 50% homology with the carboxy terminus (residues 129 to 158) of sea urchin histone H1. (B) An amino-terminal segment of *cys-3* (residues 18 to 58) displays 37% amino acid homology with oncoprotein FOS (residues 62 to 100). (C) Homology in the protein segment which includes the hypothetical leucine zipper structure of *cys-3*, FOS, and *GCN4* starts at residues 123, 161, and 245, respectively. In this region, the carboxy terminus of *GCN4* protein shows 35% identical amino acid homology with the indicated *cys-3* region. The *cys-3* protein also has 26% homology with the c-*jun* protein in the same region (not shown).

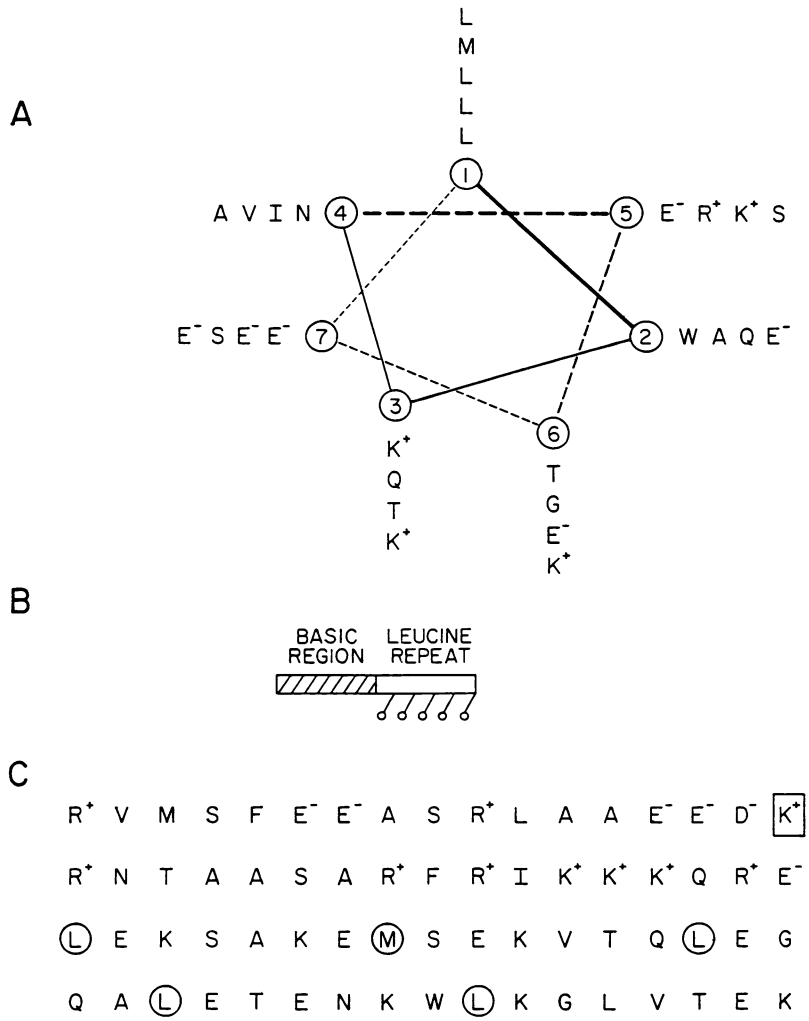


FIG. 7. Leucine zipper DNA-binding element. (A) The hypothetical leucine zipper element in which a leucine or a methionine occurs at exactly every seventh residue, resulting in their being immediately adjacent in the protein alpha-helical structure (viewed from leucine-155, the carboxyl end of the element). (B) Position of the leucine zipper. It lies immediately downstream of a charged (basic) protein segment. (C) Amino acid residues from Arg-89 to Gly-164. The pertinent leucine residues and the single methionine residue which compose the repeating zipper structure are circled. The two basic amino acid residues (Lys-105 and Arg-106) in the upstream charged region at which substitutions occur in *cys-3* mutants are boxed.

a phenotype detected by the transformation assay. Moreover, this experiment did not address the possibility of redundant TATA elements in the *cys-3* promoter. Transcription of the *cys-3* gene was found to start at three different but closely spaced sites. The existence of multiple transcription start sites is very common in fungal genes; e.g., a *Neurospora* conidiation-specific gene was recently demonstrated to display six start sites (26), whereas the cross-pathway control gene, *cpc-1* was shown to initiate transcription from seven closely spaced start sites (23). A possible hairpin structure, AGCGG(CT)CCGCT, is situated very near to the 5′ end of the *cys-3* mRNA; such hairpin structures have been shown to modulate mRNA translation (12). One unexpected finding was that although the *Eco*RV fragment (Fig. 1) transforms the *cys-3* mutant at a very high frequency, it lacks approximately 100 bases at the 3′ end of the gene and presumably is also missing any transcriptional stop signals, although it contains the entire protein-coding region.

The *cys-3* gene transcript. The *cys-3* transcript analyzed in this work is 1.3 kb in length and consists of a 5′ leader region of approximately 30 bases, the protein-coding region of 711 bases, and then a 3′ nontranslated region of 539 bases. We do not yet understand the exact nature or significance of the longer 1.6-kb *cys-3* transcript. This longer transcript might reflect an additional complexity in the function of the *cys-3* gene, or it might simply result from the use of a different polyadenylylation site in a fraction of the transcripts that yields a longer 3′ nontranslated region. The 539-base 3′ nontranslated segment of the 1.3-kb *cys-3* transcript is significantly longer than that found for other *Neurospora* mRNAs and could be involved in governing the stability of the *cys-3* mRNA. The *cpc-1* and the *con-10* gene transcripts each possess a 3′ nontranslated region of approximately 240 bases (23, 26). It is interesting that the *cpc-1* mRNA has an extremely long 5′ leader of 720 bases which contains two ATG codons that start short open reading frames which precede the true initiator ATG codon; this unusual 5′ leader may reflect translational control of the *cpc-1* mRNA (23).

Translated *cys-3* protein. The translated *cys-3* protein has a number of interesting features. In a segment of 39 amino acid

residues near the carboxyl end of the *cys-3* protein, the amino acid alanine occurs 16 times in a region that is also rich in basic amino acids. A computer search revealed that the carboxy terminus of histone H1 of the sea urchin is similarly alanine rich; moreover, the relevant regions of the *cys-3* and histone H1 proteins not only have a high content of alanine and basic amino acids but, in fact, also have a 49% amino acid identity (Fig. 6A). Of even greater interest, the predicted *cys-3* protein shows considerable amino acid homology with the oncogene v-*fos* protein and the yeast *GCN4* protein. A stretch of 41 amino acids near the amino terminus of the *cys-3* protein shows approximately 37% homology with the v-*fos* protein, although in this region neither of them has any homology to *GCN4* (Fig. 6B). In a more distal region of 39 amino acids, the *cys-3* protein shows significant homology (35%) with the extreme carboxy terminus of *GCN4* and a lower homology with v-*fos* (Fig. 6C). This is the precise region of *GCN4* which shows homology to v-*jun* and v-*fos* (14, 30). Landschulz et al. (14) have recently suggested that a number of oncogene nuclear proteins and other regulatory proteins including v-*jun*, v-*fos*, and *GCN4* have in common a new type of DNA-binding element, termed the leucine zipper, in which leucine residues occur exactly every seventh amino acid, thus being adjacent to one another in a repeating alpha-helical structure. The adjacent series of leucines are visualized to provide a hydrophobic spine that enables two protein monomers to associate and contribute to a DNA-binding site, which is postulated to be composed of the leucine zipper and an immediately adjacent basic region (14). The *cys-3* protein contains a well-defined leucine zipper composed of four precisely spaced leucines and one methionine, the other permissible amino acid (Fig. 7). It appears to be highly significant that the *cys-3* protein also contains a charged region just upstream of the leucine zipper and that this exact region contains two basic amino acids which are altered in a *cys-3* mutant, one of which is substituted in a temperature-sensitive *cys-3* mutant (Fig. 7). These observations suggest that the leucine zipper and the adjacent charged region may indeed constitute a DNA-binding site for the *cys-3* protein and that mutational changes at this location can result in a nonfunctional regulatory protein or one that is temperature sensitive for DNA binding. It is intriguing that the predicted *cys-3* regulatory protein shows significant homology in segments to other regulatory proteins, to nuclear oncoproteins, and to histone H1, a DNA-binding protein. The *cys-3* protein appears to be a mosaic composed of protein motifs that may represent ancient domains that appear in a variety of nuclear proteins in diverse organisms.

Regulation of *cys-3* expression. It was of considerable interest to determine whether the positive-acting *cys-3* regulatory gene was itself controlled by the negative-acting *scon* regulatory gene. Results presented above demonstrate that the cellular content of *cys-3* transcripts is indeed controlled by the *scon* gene. Only a very limited amount of *cys-3* mRNA is present in *scon*$^+$ during sulfur repression, whereas a substantial amount of a 1.3- and a 1.6-kb *cys-3* transcript is present during sulfur limitation; *scon*^{c} strains possess high levels of *cys-3* mRNA during both sulfur-repressed and sulfur-limited conditions. The most attractive interpretation of these results is that a *scon*$^+$-encoded regulatory protein senses the sulfur status of the cells, perhaps via a cysteine-binding site, and controls in a negative manner the transcription of the *cys-3* gene, preventing *cys-3* expression when sulfur is abundant. Thus, upon sulfur limitation, a dramatic increase would occur in the level of *cys-3* mRNA and of the *cys-3* positive-acting regulatory protein, which, in turn, would lead to expression of the entire set of structural genes of the sulfur circuit. Some evidence has suggested that *cys-3* is autoregulatory such that an increase in its protein product would further enhance transcription of the *cys-3* gene (21). It is important to note that in addition to this transcriptional model, other possible explanations cannot yet be excluded. It is plausible that the *scon* gene product controls the cellular level of *cys-3* mRNA in a different manner, perhaps at a processing step or by regulating the stability of *cys-3* mRNA.

ACKNOWLEDGMENTS

We thank Robert Metzenberg for kindly providing the two *scon* mutants used in this study. We acknowledge the valuable service of Richard Swenson and Jane Tolley, Ohio State Biochemical Instrument Center, who prepared the oligonucleotides used in this work.

This research was supported by Public Health Service grant GM-23367 from the National Institutes of Health.

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1971. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Berk A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Burton, E. G., and R. L. Metzenberg.** 1971. Novel mutation causing derepression of several enzymes of sulfur metabolism in *Neurospora crassa*. J. Bacteriol. **109**:140–151.
5. **Case, M. E.** 1983. Transformation in fungi, p. 1–5. *In* P. Lurquin and A. Kleinhofs (ed.), Genetic engineering in eukaryotes. Plenum Publishing Corp., N.Y.
6. **Dhawale, S. S., and G. A. Marzluf.** 1985. Transformation of *Neurospora crassa* with circular and linear DNA and analysis of the fate of the transforming DNA. Curr. Genet. **10**:205–212.
7. **Fu, Y. H., and G. A. Marzluf.** 1987. Molecular cloning and analysis of the regulation of *nit*-3, the structural gene for nitrate reductase in *Neurospora crassa*. Proc. Natl. Acad. Sci. USA **84**:8243–8247.
8. **Fu, Y. H., and G. A. Marzluf.** 1988. Metabolic control and autogenous regulation of *nit-3*, the nitrate reductase structural gene of *Neurospora crassa*. J. Bacteriol. **170**:657–661.
9. **Hahn, S., E. T. Hoar, and L. Guarente.** 1985. Each of three "TATA elements" specifies a subset of the transcription initiation sites at the CYC-1 promoter of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **82**:8562–8566.
10. **Hanson, M. A., and G. A. Marzluf.** 1975. Control of the synthesis of a single enzyme by multiple regulatory circuits in *Neurospora crassa*. Proc. Natl. Acad. Sci. USA **72**:1240–1244.
11. **Ketter, J. S., and G. A. Marzluf.** 1988. Molecular cloning and analysis of the regulation of *cys-14*$^+$, a structural gene of the sulfur regulatory circuit of *Neurospora crassa*. Mol. Cell. Biol. **8**:1504–1508.
12. **Kozak, M.** 1988. Leader length and secondary structure modulate mRNA function under conditions of stress. Mol. Cell. Biol. **8**:2727–2744.
13. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
14. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759–1764.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Marzluf, G. A.** 1970. Genetic and biochemical studies of distinct sulfate permease species in different developmental stages of *Neurospora crassa*. Arch. Biochem. Biophys. **138**:254–263.
17. **Marzluf, G. A., and R. L. Metzenberg.** 1968. Positive control by

the *cys*-3 locus in regulation of sulfur metabolism in *Neurospora*. J. Mol. Biol. **33**:423–437.
18. **Metzenberg, R. L.** 1979. Implications of some genetic control mechanisms in *Neurospora*. Microbiol. Rev. **43**:361–383.
19. **Metzenberg, R. L., and S. K. Ahlgren.** 1970. Mutants of *Neurospora* deficient in aryl sulfatase. Genetics **64**:409–422.
20. **Orbach, M. J., E. B. Porro, and C. Yanofsky.** 1986. Cloning and characterization of the gene for β-tubulin from a benomyl-resistant mutant of *Neurospora crassa* and its use as a dominant selectable marker. Mol. Cell. Biol. **6**:2452–2461.
21. **Paietta, J. V., R. A. Akins, A. M. Lambowitz, and G. A. Marzluf.** 1987. Molecular cloning and characterization of the *cys-3* regulatory gene of *Neurospora crassa*. Mol. Cell. Biol. **7**:2506–2511.
22. **Pall, M. L.** 1971. Amino acid transport in *Neurospora crassa*. IV. Properties and regulation of a methionine transport system. Biochim. Biophys. Acta **233**:201–214.
23. **Paluh, J. L., M. J. Orbach, T. L. Legerton, and C. Yanofsky.** 1988. The cross-pathway control gene of *Neurospora crassa*, *cpc*-1, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene v-*jun*-encoded protein. Proc. Natl. Acad. Sci. USA **85**:3728–3732.
24. **Reinert, W. R., V. B. Patel, and N. H. Giles.** 1981. Genetic regulation of the *qa* gene cluster in *Neurospora crassa*: induction of qa messenger ribonucleic acid and dependency on *qa-1* function. Mol. Cell. Biol. **1**:829–835.
25. **Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
26. **Roberts, A. N., V. Berlin, K. M. Hager, and C. Yanofsky.** 1988. Molecular analysis of a *Neurospora crassa* gene expressed during conidiation. Mol. Cell. Biol. **8**:2411–2418.
27. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**:1350–1354.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
29. **Schechtman, M. G., and C. Yanofsky.** 1983. Structure of the trifunctional *trp*-1 gene from *Neurospora crassa* and its aberrant expression in *Escherichia coli*. J. Mol. Appl. Genet. **2**:83–99.
30. **Struhl, K.** 1987. The DNA-binding domains of the *jun* oncoprotein and the yeast GCN4 transcriptional activator protein are functionally homologous. Cell **50**:841–846.