Coordinate expression of ribosomal protein genes in *Neurospora crassa* and identification of conserved upstream sequences

Yuguang Shi* and Brett M. Tyler

Department of Plant Pathology, University of California-Davis, Davis, CA 95616, USA

Received August 13, 1991; Revised and Accepted October 30, 1991

GenBank accession no. M63879

ABSTRACT

The relative levels of rRNAs and ribosomal proteins are coordinately regulated by growth rate and carbon nutrition in *Neurospora crassa*. However, little is known about the mechanisms involved. To investigate the transcriptional regulation of ribosomal protein genes in *N. crassa*, we cloned and sequenced a ribosomal protein gene (*crp-3*). The inferred *crp-3* protein sequence shares 89% and 83% homology at its N-terminus with the yeast rp51 and the human S17 ribosomal proteins respectively. The *crp-3* gene contains two introns, neither of which are conserved in position with the *RP51* or the *S17* genes. The *crp-3* gene is present in a single copy and was mapped by RFLP analysis to the right arm of linkage group IV, near the *cot-1* locus. Sequence comparisons of the upstream regions of the three sequenced *crp* genes revealed several common features. These include a 'Taq box' (consensus: ARTTYGACTT) at –39, a CG repeat (consensus: CCCRCRRR) at –65, and a major transcription initiation site embedded in a purine rich region flanked by an upstream pyrimidine rich sequence. Using four *N. crassa* ribosomal protein genes as probes, we demonstrated that the levels of the four ribosomal protein mRNAs were closely coordinated during a nutritional downshift from sucrose to quinic acid.

INTRODUCTION

Ribosomes consist of 60–80 ribosomal proteins and 3–4 rRNA molecules. Their synthesis comprises a major metabolic investment by cells. Under favorable growth conditions, ribosomes contain about 85% and 15% of the total cellular RNA and protein respectively (1–5). Therefore regulation of the levels of ribosome synthesis and coordinate expression of individual ribosomal component genes are essential. In *E. coli* and yeast (*Saccharomyces cerevisiae*), the ribosome content is proportional to growth rate (1, 6) and the rates of ribosomal synthesis are promptly adjusted to changes in cell growth and environmental conditions (4, 5). In *E. coli*, the rate of ribosome synthesis is

controlled by the rate of rRNA synthesis; ribosomal protein synthesis is coordinated with rRNA synthesis by autogenous control of the translation of the ribosomal protein mRNAs (4). In eukaryotes, coordinate control of the rates of rRNA and r-protein synthesis (1, 2, 7, 8) is more complex, involving genes transcribed by the three nuclear RNA polymerases. The large rRNA precursors are synthesized by RNA polymerase I, r-protein mRNAs by RNA polymerase II, and 5S rRNAs by RNA polymerase III. Coordination appears to be achieved via two basic mechanisms: first, there may be concerted de novo synthesis of the different rRNAs and r-protein mRNAs (1). Second, a variety of feedback mechanisms controlling RNA splicing and stability, and protein turnover may act to precisely balance the supply of free ribosomal components (9).

In yeast, transcription of r-protein genes is coordinated during nutritional shifts of carbon or nitrogen source or amino acid starvation (1, 7, 8, 10, 11). Comparison of sequenced yeast r-protein genes identified a 15 bp conserved sequence, UAS_{rpg}, in the 5' flanking region of most but not all r-protein genes (2, 5, 12, 13). Promoter deletion and gene fusion studies of several genes confirmed that the UAS_{rpg} was responsible for coordinate activation of r-protein genes (14–16). The UAS_{rpg} also is found 5' to the gene encoding a common subunit of RNA polymerase I and polymerase III (17), suggesting a possible mechanism for coordinating the synthesis of r-proteins with that of the rRNAs. However, this has not yet been tested, and the UAS_{rpg} is found also in many other genes not associated with protein synthesis (5). A regulatory protein known as TUF, RAP or GRF which binds to the UAS_{rpg} has been identified and characterized (18–22). Autogenous control of splicing of the pre-mRNA for yeast ribosomal protein L32 has been described (9, 23) but this mechanism does not appear to be general.

In mammalian systems, the promoter regions essential for efficient transcription of r-protein genes are often very short (24), and share some common features. These include the lack of a canonical TATA box, a cap site embedded in a pyrimidine tract flanked by GC rich sequences, and one or two intragenic elements in the first exon or intron (24–28). These common features suggest that expression of these genes may be coordinated via one or several of the common promoter elements. But in general,

* Present address: Department of Microbiology & Immunology, Hormone Research Institute, University of California-San Francisco, CA, USA

co-ordinate regulation of r-protein genes in mammalian systems is not as well understood as in yeast. Translational control of r-protein synthesis has been reported in higher eukaryotes, such as *Drosophila* (29, 30), *Xenopus* (31, 32), and mouse (30, 33), but this control is not autogenous.

In *Neurospora crassa*, the rates of both rRNA and r-protein synthesis increase markedly within 30 min following a nutritional upshift from acetate to glucose. During a downshift from glucose to glycerol, both rRNA and r-protein synthesis are severely inhibited for at least two hours after the shift before a slower rate of synthesis recommences (34–37). Studies of 5S and 40S transcription *in vitro* have led to the identification of a common *cis* element which is essential for *in vitro* transcription of the 5S and 40 rRNA genes (38, 39), but not tRNA genes (40). This element, the Ribo box, also is present at three locations in the promoter region of a r-protein gene (*crp-2*), including the transcription initiation site (41). Therefore, the Ribo box may play a role in coordination of transcription of ribosomal genes in *N. crassa*. However, there have been no studies of how r-protein genes are regulated *in vivo*. Here we report the cloning and characterization of the r-protein gene *crp-3*, homologous to the yeast *rp51* and mammalian *S17* genes, and show that the mRNAs of four *N. crassa* r-protein genes are coordinately regulated during a nutritional downshift. Sequence comparisons of the 5' flanking regions of the *crp-1*, *crp-2* and *crp-3* genes reveal two conserved upstream elements.

MATERIALS AND METHODS

Amplification of a *crp-3* genomic fragment by the polymerase chain reaction

Two regions highly conserved between yeast *rp51* and human *S17* r-protein sequences were used to design degenerate oligonucleotide primers for PCR. The strongly biased codon usage of actively expressed genes in *N. crassa* was used to eliminate many possible codons. Total genomic DNA (1 µg) from *N. crassa* strain 105C was used as template DNA for PCR amplification of *crp-3* genomic fragments. Conditions used for DNA amplification reactions were essentially as described (42). The PCR reaction was carried out in a programmable DNA thermal cycler (Perkin Elmer Cetus). The thermal program used for PCR was 94°C for 2 min, followed by 25 cycles of 1 min at 94°C, 2 min at 50°C and 3 min at 72°C, and finally 72°C for 10 min. The amplified fragments were then gel purified and used as probes for screening both cDNA and genomic libraries from *N. crassa*.

Cloning and sequencing of *crp-3* cDNA and genomic DNA clones

Fragments of the yeast *RP51* gene (provided by M. Rosbash, Brandeis University) and *N. crassa crp-3* genomic fragments produced by PCR were used as probes to screen cDNA and genomic libraries. The two cDNA libraries (provided by M. Sachs, Stanford University) were prepared with mRNA from either mycelia or germinating conidia, using λ-ZAP (Stratagene) as the vector. The genomic library (provided by J. Baum and R. Geever, University of Georgia) consisted of six independent pools, each derived from 10⁴ phage plaques with λ-EMBL3 as vector. Hybridizations to the cDNA and genomic libraries were carried out at 42–55°C when using the yeast *RP51* gene as a probe, and at 62°C when using the *crp-3* PCR fragment as a probe. The hybridization buffer consisted of 2×SSC (0.3 M

NaCl, 0.03 M Na₃ citrate), 0.5% SDS, 0.1% bovine serum albumin, 0.1% polyvinylpyrrolidone, 0.1% Ficoll, 100 µg/ml heparin, and 1 mM EDTA. Following screening of the cDNA libraries, the positive plaque which carried the longest cDNA insert was identified by re-screening 10 strongly hybridizing plaques with PCR, using the *crp-3* 3' primer paired with either of the *universal* or *reverse* primers which flank the cDNA insertion sites of vector. This clone was sub-cloned by M13-assisted plasmid release (pop-out), avoiding plaque purification and phage propagation (43). Following screening of the genomic libraries, λ phage DNA was prepared (44) from a single positive clone from each of four sub-libraries. Restriction mapping and further subcloning of cDNA and genomic clones were carried out essentially as described (44). Exonuclease III was used to generate deletion clones for sequencing (45). Dideoxynucleotide sequencing with Sequenase (USB) was carried out using either alkali-denatured double stranded plasmid DNA or single stranded DNA. Single stranded plasmid DNA was prepared from *E. coli* cells containing Bluescript (Stratagene) or pTZ18R plasmids carrying the f1 ori site, by superinfection with M13K07 helper phages. Chemical sequencing was conducted according to Maxam and Gilbert (46) with a modification (47).

5' End-mapping of *crp-3* mRNA

Transcription initiation sites were determined by S₁ nuclease digestion of DNA-mRNA hybrids, and by primer extension. The DNA probe used for S₁ mapping was a 175 bp *Apa*LI-*S*tyI fragment spanning the 5' upstream and first exon regions. The primer was a 30 bp *Bst*UI-*S*tyI fragment within the first exon. S₁ mapping and primer extension assays were carried out essentially as described (41) except that hybridization was carried out at 50°C instead of 56°C for S₁ mapping.

Genetic mapping of *crp-3*

The genomic location of the *crp-3* gene was mapped using restriction fragment length polymorphisms (RFLPs). The probe used for mapping was prepared from a 20 kb genomic clone by random primer labeling. One fragment was found to be polymorphic after *S*tyI digestion of genomic DNAs of Oak Ridge (OR, 3.5 kb fragment) and Mauriceville (M, 3.8 kb fragment) strains of *N. crassa*. The location of *crp-3* was determined by comparing the segregation of M and OR alleles of *crp-3* among 38 progeny with the known segregation of 55 other RFLP loci (48).

Isolation and measurement of level of mRNA of *crp* genes

Conidia from each strain were shaken at 300 rev./min at 30°C overnight without light in 800 ml of minimal Vogel's medium plus 1.5% sucrose to allow germination and outgrowth. After 13.5 hours, each culture was harvested by filtration onto Miracloth (Calbiochem), rinsed with ice-cold sterile distilled water, and then divided into three equal portions. Portion 1 (S3) was cultured in 1.5% sucrose medium for 3 hours. Portion 2 (Q3) and portion 3 (Q10) were cultured in minimal Vogel's medium plus 0.3% quinic acid (49) for 3 and 10 hours respectively under the original conditions. Finally, the mycelia were harvested by filtration onto miracloth, rinsed with ice-cold sterile distilled water, frozen immediately in liquid N₂ and stored at –80°C. mRNA was prepared from the frozen mycelia using the methods similar to those of Patel *et al.* (50).

S1 nuclease mapping was used to measure the mRNA levels of four different r-protein genes and the β-tubulin gene. The method used for S1 mapping was the same as that for mapping

Fig. 1. PCR amplification of *N. crassa* *crp-3* sequences. (A). Highly conserved regions of the yeast rp51 and human S17 proteins used to design PCR primers for cloning *crp-3*. The degenerate primer sequences were derived according to the biased codon usage of highly expressed *N. crassa* genes (41). (B) *Crp-3* DNA fragment (370 bp) amplified from *N. crassa* genomic DNA and resolved on a 1.5% agarose gel. Lane a: PCR product from 10 μ M of each primer; lane b: PCR product from 1 μ M of each primer; lane M: DNA size marker (lambda DNA digested with HindIII).

the 5' end of *crp-3* mRNA, except that the hybridizations were carried out at 58°C. The fragments used for the S1 analysis were 580 bp *Bst*BI-*Eco*NI (5' end labelled) for *crp-1* (51), 213 bp *Eco*RI-*Bgl*II (5' end labelled) for *crp-2* (41), 373 bp *Bam*HI-*Ava*I (5' end labelled) for *crp-3*, 365 bp *Eco*O109I fragment (3' end labelled) for *crp-4* (a homolog of yeast YL3, de la Serna, Shi and Tyler, unpublished), and a 353 bp *Pvu*II-*Bam*HI (3' end labelled) for the β -tubulin gene (52). The mRNA level for each gene was measured by scanning the resulting radioactive band in the dried gels using an AMBIS radioanalytic imaging system (San Diego, CA).

RESULTS

Cloning and sequencing of *crp-3*

The yeast rp51 protein is strongly homologous to the human S17 protein, primarily at the N-terminus (53). Primers for PCR were synthesized based on two highly conserved regions of the yeast rp51 and human S17 r-protein sequences (Fig. 1A). The degenerate primers were based on the most common codon usage in several *N. crassa* genes (Fig. 1A). PCR amplification with these primers resulted in a specific DNA fragment of about 370 base pairs (Fig. 1B).

The amplified genomic fragment was used as a probe for screening the cDNA libraries. The clone with the longest insert (750 bp) was identified by PCR screening of 10 strongly hybridizing plaques. The insert was subcloned by M13-assisted plasmid release, then sequenced. The 5' end of the cDNA sequence (the first 276 bp of the coding region) showed 74% homology with yeast *RP51* and 75% with human *S17*. To clone the *crp-3* gene, the genomic library was screened with the PCR amplified genomic fragment. Restriction digestion and Southern analysis of four positive clones indicated that all contained a 1.3 kb *Sal*I fragment which hybridized to the cDNA clone (not shown).

Sequences of the cDNA and genomic clones are shown in Fig. 2. The introns were identified by comparing the cDNA sequence with the genomic sequence. Unlike the yeast *RP51* gene, which contains only one intron, the *N. crassa* gene possesses two introns and three exons. Both intron sequences start with GT and end with AG. An 8 bp consensus sequence common to *N. crassa* introns (54) was also found in each intron (underlined). The inferred protein sequence of the *crp-3* gene shares 89% homology

Fig. 2. Nucleotide sequence of the *crp-3* gene. The triangle indicates the major transcription start point and the diamonds indicate the minor starts (see Fig. 3). Slashes show the intron boundaries. Underlined sequences in the introns are conserved heptamers found in *N. crassa* introns (54). Asterisks indicate identities between the predicted *crp-3* protein sequence and the yeast rp51 and human S17 protein sequences. The reversed double arrow at the 3' end marks the polyadenylation site in the cDNA clone. The full *crp-3* 5' flanking sequence (817 bp) is available in Genbank (accession number M63879). Boxed sequences are conserved among *crp-1*, *crp-2* and *crp-3*. Hatched boxes indicate pyrimidine-rich sequences flanking the transcription startpoint. The pair of arrows indicates a perfect 11 bp tandem repeat in the 5' untranslated region.

with the yeast *rp51* protein and 83% homology with the human *S17* protein at the N-terminus, previously shown to be highly conserved (53). This suggests that the *N. crassa* gene is a homolog of both genes. Since this is the third *N. crassa* r-protein gene characterized, it was named *crp-3* (cytoplasmic ribosomal protein). The polyadenylation site of the *crp-3* cDNA clone is located in a purine rich region (marked by the reverse arrow head in Fig. 2).

The transcription initiation sites of the *crp-3* gene were mapped by *S*₁ nuclease digestion of RNA-DNA hybrids and by primer extension. The DNA probe used for S1 mapping was a 175 bp *Apa*LI-*Sly*I fragment spanning the 5' upstream and first exon regions. A 30 bp *Bst*UI-*Sly*I fragment within the first exon was used for primer extension. The results are shown in Fig. 3. The major initiation site was mapped at -57 relative to the first AUG codon by both S₁ and primer extension assays (indicated by the arrow in Fig. 3). A total of 10 initiation sites were revealed in a region between 107 to 43 bp upstream from the first AUG codon (Fig. 2). It is common in fungi that the 5' ends of mRNAs are heterogeneous (41, 55, 56).

Fig. 3. Mapping of the transcription initiation site of *crp-3* by primer extension and S1 nuclease mapping. The major transcription site is indicated by the arrow as the +1 site. The lanes are A+G: A+G chemical sequencing of S₁ probe as marker; P: primer extension; S₁: S₁ nuclease mapping; N: *N. crassa* mRNA, and E: *E. coli* control RNAs. The four strong bands appearing above the primer in both the primer extension lanes are restriction fragments contaminating the primer. The greater heterogeneity of the bands in the S₁ analysis is an artifact of the enzyme digestion.

Genetic mapping of the *crp-3* gene

Gel blot hybridization analysis using the *crp-3* cDNA as a probe showed a single band hybridizing to *N. crassa* genomic DNA digested with three different restriction enzymes (*Bam*HI, *Sal*I, and *Pst*I), indicating that the *crp-3* gene is present in a single copy (data not shown). This differs from yeast where the *RP51* gene is present in two copies (13). To genetically map the *crp-3* gene, a 20 kb genomic fragment containing the *crp-3* gene was used as a hybridization probe to screen for restriction fragment length polymorphisms (RFLPs) in genomic DNAs of two *N. crassa* strains, Oak Ridge (OR) and Mauriceville (M). A *Sty*I fragment (3.8 kb in M and 3.5 kb in OR) was found to be polymorphic (lane OR and M, Fig. 4A). The 20 kb probe was

Fig. 4. Genetic mapping of *crp-3* using RFLPs. (A) Segregation of an RFLP adjacent to the *crp-3* gene. DNAs from 38 progeny of an Oak Ridge (OR) × Mauriceville (M) cross (57) were digested with *Sty*I and hybridized with a 20 kb genomic fragment carrying the *crp-3* gene. The polymorphic fragments from the digestion are indicated by the arrows. Higher molecular weight products in some lanes are partial digestion products. OR and M indicate DNA from the Oak Ridge and Mauriceville parents respectively. A1, A4, B6 etc indicate progeny DNAs. Progeny with same letter code (e. g. A1 and A4) originate from the same ordered tetrad (57). (B) The RFLP map containing *crp-3*. The circle indicates the centromere. The Greek letters represent 5S rRNA genes. *Crp-3* was aligned at the same location as the *cot-1* locus, but due to uncertainties in the mapping, they may be several cM apart.

then hybridized to *Sty*I digested DNAs from 38 progeny of a cross between the Oak Ridge and Mauriceville strains (lane A1 to R4, Fig. 4A). This progeny set has previously been used to construct RFLP genetic maps of *N. crassa* (57). The *crp-3* gene was mapped to the right arm of linkage group IV at the same location as the *cot-1* locus (Fig. 4B), but *crp-3* and *cot-1* have not been directly mapped with respect to each other. Sequence comparison showed that *crp-3* and *cot-1* are distinct genes (C. Yanofsky, personal communication).

Coordinate changes in r-protein mRNA levels during a nutritional downshift

To begin to examine the regulation of *N. crassa* r-protein gene expression, and in particular to determine the level of coordinate control, we measured the mRNA levels of four *N. crassa* r-protein genes at several points during a nutritional downshift. Three strains derived from 246-89601-2A by transformation (for another purpose) were grown to mid log phase in sucrose medium, then the mycelia were harvested and divided into three portions. The portions were reinoculated into 1.5% sucrose for 3 hr (S3), into 0.3% quinic acid for 3 hr (Q3) or into 0.3% quinic acid for 10 hr (Q10). Quinic acid is a poor carbon source for *N. crassa* and adaptation to stable growth on quinic acid takes 4–6 hr (49). In each culture, the mRNA levels were measured for r-protein genes *crp-1* (51), *crp-2* (41), *crp-3* (this paper) and *crp-4* (de la Serna, Shi and Tyler, unpublished). *Crp-4* was cloned

Fig. 5. Coordinate regulation of r-protein mRNAs during a carbon nutritional downshift (A) Quantitation of mRNAs by S1 nuclease digestion of RNA-DNA hybrids. The mRNA was isolated from three *N. crassa* strains grown in sucrose for three hours (S3), or in quinic acid for 3 hours (Q3) and 10 hours (Q10) respectively. Ec indicates control RNA from *E. coli*. The DNA probes used were fragments from the *crp-1*, *crp-2*, *crp-3*, *crp-4*, and β -tubulin genes (see Material and Methods). The respective fragments protected from S1 nuclease are marked by the arrows. The multiple bands (marked by a vertical line) from *crp-4* reflect multiple termination sites of transcription. (B) Correlation of relative mRNA levels among the *crp* genes during the downshift. All data were derived from quantitation of the radioactive bands shown in the autoradiograph and were standardized with the mRNA level of the β -tubulin gene. The relative mRNA levels were expressed as a percentage of the level in sucrose cultures. The levels of the three other *crp* genes were plotted against *crp-2* (on the left) or *crp-3* (on the right). The diagonal lines on each graph indicate the points at which the three mRNA levels (y-axis) are identical to the fourth (x-axis).

by homology to the yeast 5S rRNA-binding r-protein YL3 using PCR (Shi and Tyler, unpublished). The preliminary sequence of a *crp-4* cDNA clone encodes an amino acid sequence with 68% identity to that of yeast YL3 (de la Serna, Shi and Tyler, unpublished; J. Woolford, personal communication). β -tubulin mRNA levels were measured as a standard. S1 nuclease mapping was used to quantitate simultaneously all five mRNAs in the same hybridization reaction. The results are shown in Fig. 5A. In each strain, after 3 hours growth in quinic acid all four r-protein

A CG REPEAT			
<i>crp-1</i>	-66	CCCACCAAAA	-57 >
<i>crp-2</i>	-74	CCCACCAAGT	-65 >
<i>crp-3</i>	-65	CCCACCGGAT	-56 >
<i>crp-2</i>	-192	CCCAGCAAAAT	-183 >
<i>crp-2</i>	-26	CCCACCGAAT	-17 >
consensus		CCC ^G CAAAAT ^G	>
B TAQ BOX			
<i>crp-1</i>	-40	AATTTGACTT	-31 >
<i>crp-2</i>	-30	AGTTCGACTT	-39 <
<i>crp-3</i>	-40	AGTTCGACTT	-31 >
consensus		A ^G TT ^C GACTT ^A	>
C RIBO BOX			
α 5S rRNA	+18	CTGGAAACTCG-GGATC	>
β 5S rRNA	+18	GTAGAAACTCG-GGATC	>
40S rRNA	-65	GTTGAAAGTCGCGGCTC	>
<i>crp-2</i>	-9	CTGGAAACTCG-AGATC	>
consensus		C ^T TGAAAC ^T C ^G -G ^A ATC ^A	>
<i>crp-2</i>	-271	CTGGcAtcGTcGcAGATC	<
<i>crp-2</i>	-89	GgGGAAAGTGGcAGaAa	<
<i>crp-1</i>	-457	CTGGAAgACaGG-GGAaC	>
<i>crp-1</i>	-384	GcTgcAAgaTCG-GGCTC	<
<i>crp-1</i>	-270	GaAGgcAAGaGG-GGCTC	>
<i>crp-3</i>	-291	CTTGAARtCTtGgcGATC	>

Fig. 6. Consensus sequences for the CG repeat (A), Taq box (B) and Ribo box (C). > and < indicate relative orientations. Lower case letters represent mismatches with the consensus. *Crp-1* and *crp-2* data are from refs (41) and (51) and the *crp-3* data are from this paper

mRNAs had dropped sharply to 3–5% of the sucrose grown levels (compare lane S3 with Q3 in Fig. 5A), while the β -tubulin mRNA levels remained about the same. After 10 hrs in quinic acid (Q10 in Fig. 5A), each of the r-protein mRNA levels had partially recovered. The extent of recovery varied among the three strains (20% to 80% of the S3 levels). However, the relative levels of the four r-protein mRNAs remained closely coordinated. Fig 5B shows the relative levels of the *crp-1*, *crp-3* and *crp-4* mRNAs plotted against *crp-2* for all three strains after 3 and 10 hr in quinic acid (on the left). The levels of *crp-1*, *crp-2* and *crp-4* mRNAs are also plotted against *crp-3* in a similar manner (on the right). In each case the mRNA levels are expressed as a percentage of the level in the sucrose cultures, after normalization to the β -tubulin mRNA level. In all cases there is a close coordination of the four r-protein mRNA levels irrespective of the absolute levels.

DISCUSSION

In *Neurospora crassa*, the syntheses of rRNAs and r-proteins are coordinately regulated during nutritional shifts (3). However, little is known about the mechanism involved in coordination, either at the transcriptional or translational level. To facilitate our studies on mechanisms involved in co-regulation of ribosomal genes in *N. crassa*, we cloned and sequenced the *crp-3* gene of *N. crassa*, the third r-protein gene identified in this fungus. The N-terminus of the inferred *crp-3* protein sequence shows 89% homology with yeast rp51 and 83% homology with human S17, as is the case between the rp51 and the S17 proteins (53). *Crp-3* is a single copy gene, like *crp-1* and *crp-2*, in contrast to the

RP51 gene and many other yeast r-protein genes which are duplicated in the genome (5). *Crp-3* contains two introns, neither of which are conserved in position with the yeast *rp51* gene which carries only one intron (58), nor with the human *S17* gene which possesses four introns (59). The first intron of human *S17* is conserved in position with that of yeast *RP51*. In the *N. crassa* *crp-2* gene, two intron positions are conserved in the human *S14* gene but no intron positions are conserved between the yeast *CRY1* and human *S14* genes (41). In *crp-2* and *crp-3*, no intron positions are conserved in the *N. crassa* and yeast genes, though *N. crassa* and yeast are evolutionarily closer to each other than to humans.

We measured the mRNA levels of four r-protein genes (*crp-1*, *crp-2*, *crp-3* and *crp-4*) during a nutritional downshift from growth on sucrose to growth on quinic acid. Quinic acid is a poor carbon source for *N. crassa* and adaptation to stable growth on quinic acid takes 4–6 hr (49). The levels of all four mRNAs dropped sharply to about 3–5% of the sucrose levels after 3 hrs culture in quinic acid. After a total of 10 hr in quinic acid, the levels of the four r-protein mRNAs had recovered to about 20%, 80%, or 50% of sucrose levels in strain 1, 2, and 3 respectively. These results show that r-protein mRNA levels in *N. crassa* are co-ordinately regulated during a carbon downshift. Presumably the mRNA levels are responding to a change in the growth rate (3), or to signals from the carbon catabolite regulatory system (60). An *N. crassa* r-protein mRNA homologous to rat *S26* also was reported recently to be regulated by carbon nutrition (61). Our results also showed that there was close coordination of the relative levels of the four r-protein mRNAs during the downshift. This was particularly apparent when the r-protein mRNA levels in the three strains were compared after 10 hr in quinic acid; although the absolute mRNA levels in the three strains had recovered to different levels, the relative levels of the four r-protein mRNAs were very similar within each strain.

The rates of rRNA synthesis in *N. crassa* have been measured during a carbon downshift from glucose to glycerol (34). The rate of rRNA synthesis dropped to 10–15% of glucose grown levels during the downshift, and remained at that level for 2 hr (34). After 3 hr, the rate of rRNA synthesis recovered to about 50% of the glucose grown level and remained at that level until the end of the experiment at 6 hr (34). These changes in the rate of rRNA synthesis closely parallel the changes in r-protein mRNA levels we observed. Therefore, in *N. crassa*, the levels of r-protein mRNAs appear to be down-regulated in concert with the rate of rRNA synthesis during a carbon downshift. However, we cannot tell presently whether the r-protein mRNA levels are regulated by changes in the stability of the mRNAs or by changes in the transcription rate. The rate of *N. crassa* rRNA synthesis also drops to 3–4% of the control during phosphorus starvation (62). It will be of interest to measure r-protein mRNA levels during phosphorus starvation, as well as other types of nutritional shifts.

In *Saccharomyces cerevisiae*, the transcription rates of most r-protein genes are coordinately regulated by growth rate and carbon nutrition by the action of a common upstream transcriptional element, the UAS_{rpg} . The UAS_{rpg} is a 15 bp conserved sequence often found in duplicated copies or as tandem repeats 250–450 upstream of the first AUG codon of r-protein genes (5). Deletion of the UAS_{rpg} usually results in a reduction of transcription of a gene carrying the sequence (15, 63–65). The UAS_{rpg} is also found in the 5' flanking region of non-ribosomal protein genes (18, 19), while some r-protein genes, like *RPL3* and *RPS33*, are not co-regulated via the UAS_{rpg} (2,

5). Therefore, the UAS_{rpg} is not confined to coordinating r-protein genes. In addition, little is known about how transcription of yeast r-protein genes is coordinated with rRNA synthesis. In the *crp-1*, *crp-2* and *crp-3* genes the best homologies to the UAS_{rpg} consensus have only 11/15 matches. This is not statistically significant, though protein binding experiments have not been carried out. It is not known if there is an *N. crassa* homologue of the UAS_{rpg} binding protein of yeast.

To search for a conserved sequence element which might be responsible for the coordinated transcription of r-protein genes in *N. crassa*, computer comparisons were carried out among the 5' flanking sequences of the *crp-1*, *crp-2*, and *crp-3* genes. The analysis revealed two highly conserved sequences located at about –39 and –65 in each of the three r-protein genes. The sequence at around –65 was previously described as common to *crp-1* and *crp-2* and was named the CG repeat (41). It is present at –65 in *crp-3* (Fig. 2), at –66 in *crp-1*, and three times in *crp-2*, at –192, –74 and –26 (Fig. 6A). The CG repeat resembles the human GC box (consensus: RYYCCGCC^{G/T}) which is known to be binding site for the transcription activator SP1 (66, 67). The CG repeat in *crp-2* binds human SP1 factor *in vitro* (Y. Shi and B. M. Tyler, unpublished), but it is unknown whether there is an SP1-like factor in *N. crassa*. Of 31 other *N. crassa* genes examined, none had perfect matches to the CG repeat consensus within 300 bp of the transcription startpoint; sequences with two mismatches occurred occasionally, but not at statistically significant frequency. The conserved sequence at –39 is a partially palindromic sequence, designated the Taq box, which is present at –40 in *crp-1* and *crp-3*, and at –39 in *crp-2* but in opposite orientation (Figs. 2 and 6B). The Taq box is relatively A-T rich and resembles in location the A-T rich sequence in mammalian r-protein genes, which is important for transcriptional efficiency *in vivo* (27). We previously identified a sequence, the Ribo box, which is required for transcription of the 5S and 40S rRNA genes and which was present in the promoter region of *crp-2* (38–41). The Ribo box is not highly conserved in the upstream regions of *crp-1* and *crp-3*; the best homologies have 4 mismatches to the consensus (Fig 6C). Therefore it remains unclear whether the Ribo box could coordinate all rRNA and r-protein genes in *N. crassa*. This should be resolved by tests of Ribo box function *in vivo*.

Crp-3 lacks an obvious TATA box, like *crp-2* and many other genes in *N. crassa* (*crp-1* has a possible TATA box at –42). *Crp-3* also shares some common features with *crp-1* and *crp-2* around the major transcription initiation site. In each case, the major initiation site is embedded in a relatively purine rich region and is flanked on the 5' side by a pyrimidine rich sequence, partially resembling the features of r-protein genes in mammalian systems. The transcription start site of the human *S16* r-protein gene is determined *in vivo* by a pyrimidine rich tract (27). In filamentous fungi such as *Aspergillus nidulans* and *N. crassa*, many genes contain a pyrimidine rich region upstream of their transcription initiation sites (68–71). In the case of the *trpC* and *gpdA* genes of *A. nidulans*, the pyrimidine rich sequence is important for efficient transcription (68, 69). It is possible that the conserved elements, including the CG repeat and the Taq box, serve as a common core promoter of r-protein genes in *N. crassa*, similar to mammalian r-protein genes (24–26). The *crp-3* gene also has a perfect tandem repeat of 12 bp at +28 and +39, 7 bp upstream from the first AUG codon. Testing the role of any of these conserved sequences will require detailed mutational analyses of the promoters *in vivo*.

ACKNOWLEDGEMENTS

We thank M. Rosbash (Brandeis University) for the yeast *RP51* gene; Jim Baum and Robert Geever (University of Georgia) for the *N. crassa* [lambda] genomic library; M. Sachs (Stanford University) for *N. crassa* cDNA library; Robert Metzberg (University of Wisconsin, Madison) for DNAs from RFLP-mapping progeny; Layne Huiet for helpful discussion; Julie Billington, Tom Cujec, Eva Steinberger, and Ivana de la Serna for criticizing the manuscript; and Jeff Hall for photography. This work was supported by NIH grants GM36964 and GM42178 (to B. M. T.).

REFERENCES

- Kief, D.R. & Warner, J.R. (1981) *Mol. Cell. Biol.* **1**, 1007–1015.
- Mager, W.H. (1988) *Bioch. Bioph. Acta* **949**, 1–15.
- Alberghina, L. & Sturani, E. (1981) *Microbiol. Rev.* **45**, 99–122.
- Nomura, M., Gourse, R. & Baughman, G. (1984) *Ann. Rev. Biochem.* **53**, 75–117.
- Warner, J.R. (1989) *Microbiol. Rev.* **53**, 256–271.
- Kjeldgaard, N.O. & Gausing, K. (1974)
- Warner, J.R. & Gorenstein, C. (1977) *Cell* **11**, 201–212.
- Warner, J.R. & Gorenstein (1978) *Nature* **275**, 338–339.
- Warner, J.R., Mitra, J.R., Schwindinger, W.F., Studeny, M. & Fried, H.M. (1985) *Mol. Cell. Biol.* **5**, 1512–1521.
- Pearson, N.J. & Haber, J.E. (1980) *Mol. Cell. Biol.* **143**, 1411–1419.
- Kraig, E., Haber, J.E. & Rosbash, M. (1982) *Mol. Cell. Biol.* **2**, 1199–1204.
- Planta, R.J. & Raué, H.A. (1988) *Trends in Genet.* **4**, 64–68.
- Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, H., Planta, R.J., Schultz, L., Friesen, J.D., Fried, H. & Rosbash, M. (1983) *Nucl. Acids Res.* **12**, 8295–8312.
- Donovan, D.M. & Pearson, N.J. (1986) *Mol. Cell. Biol.* **6**, 2429–2435.
- Rotenberg, M.O. & Woolford, J.L.J. (1986) *Mol. Cell. Biol.* **6**, 674–687.
- Herruer, M.H., Mager, W.H., Woudt, L.P., Nieuwint, R.T.M., Wassenaar, G.M., Groeneveld, P. & Planta, R.J. (1987) *Nucl. Acids Res.* **15**, 10133–10144.
- Mann, C., Buhler, J.-M., Treich, I. & Sentenac, A. (1987) *Cell* **48**, 627–637.
- Buchman, A.R., Kimmerly, W.J., Rine, J. & Kornberg, R.D. (1988) *Mol. Cell. Biol.* **8**, 210–225.
- Buchman, A.R., Lue, N.F. & Kornberg, R.D. (1988) *Mol. Cell. Biol.* **8**, 5086–5099.
- Huet, J. & Sentenac, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3648–3652.
- Shore, D., Stillman, D.J., Brand, A.H. & Nasmyth, K.A. (1987) *EMBO J.* **6**, 461–467.
- Shore, D. & Nasmyth, K. (1987) *Cell* **51**, 721–732.
- Wittekind, M., Kolb, J.M., Dodd, J., Yamagishi, M., Memet, S., Buhler, J.M. & Nomura, M. (1990) *Mol. Cell. Biol.* **10**, 2049–2059.
- Dudov, K.P. & Perry, R.P. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8545–8549.
- Hariharan, N., Kelley, D.E. & Perry, R.P. (1989) *Genes and Devel.* **3**, 1789–1800.
- Hariharan, N. & Perry, R.P. (1989) *Nucl. Acids Res.* **17**, 5323–5337.
- Hariharan, N. & Perry, R.P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1526–1530.
- Meyuhas, O. & Klein, A. (1990) *Journal of Biological Chemistry* **265**, 11465–11473.
- Schmidt, T., Chen, P.S. & Pellegrini, M. (1985) *J. Biol. Chem.* **260**, 7645–7650.
- Kay, M.A. & Jacobs-Lorena, M. (1985) *Mol. Cell. Biol.* **5**, 3583–3592.
- Pierandrei-Amaldi, P., Campioni, N., Beccari, E., Bozzoni, I. & Amaldi, F. (1985) *Dev. Biol.* **167**, 281–289.
- Mariottini, P. & Amaldi, F. (1990) *Mol. Cell. Biol.* **10**, 816–822.
- Meyuhas, O., Thompson, E.A. & Perry, R.P. (1987) *Mol. Cell. Biol.* **7**, 2691–2699.
- Sturani, E., Magnani, F. & Alberghina, F.A.M. (1973) *Bioch. Bioph. A.* **319**, 153–164.
- Alberghina, F.A.M., Sturani, E. & Gohlke, J.R. (1975) *J. Biol. Chem.* **250**, 4381–4388.
- Sturani, E., Costantini, M.G., Zippel, R. & Alberghina, F.A.M. (1976) *Exp. Cell Res.* **99**, 245–252.
- Alberghina, L., Sturani, E., Costantini, M.G., Martegani, E. & Zippel, R. (1978) In Burnett, J.H. & Trinci, A.P.J. (e.d.), *Fungal Walls and Cell Growth*. Cambridge University Press, London, pp. 295–318.
- Tyler, B.M. (1987) *J. Mol. Biol.* **196**, 801–811.
- Tyler, B.M. (1990) *Nucl. Acids Res.* **18**, 1805–1811.
- Shi, Y. & Tyler, B.M. (1991) *J. Biol. Chem.* **266**, 8015–8019.
- Tyler, B.M. & Harrison, K. (1990) *Nucl. Acids Res.* **18**, 5759–5765.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988) *Science* **239**, 487–494.
- Short, J.M., Fernandez, J.M., Sorge, J.A. & Huse, W.D. (1988) *Nucl. Acids Res.* **16**, 7583.
- Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Henikoff, S. (1984) *Gene* **28**, 351–359.
- Maxam, A.M. & Gilbert, W. (1980) *Methods. Enzymol.* **65**, 455–560.
- Shi, Y. & Tyler, B.M. (1989) *Nucl. Acids Res.* **17**, 3317.
- Metzenberg, R.L., Stevens, J.N., Selker, E.U. & Morzycka-Wroblewska, E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2067–2071.
- Giles, N.H., Case, M.E., Baum, J., Geever, R., Huiet, L., Patel, V. & Tyler, B. (1985) *Microbiol. Rev.* **49**, 338–358.
- Patel, V.B., Schweizer, M., Dykstra, C.C., Kushner, S.R. & Giles, N.H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5783–5787.
- Kreader, C.A. & Heckman, J.E. (1987) *Nucl. Acids Res.* **15**, 9027–9042.
- Orbach, M.J., Porro, E.B. & Yanofsky, C. (1986) *Mol. Cell. Biol.* **6**, 2452–2461.
- Chen, I.-T., Dixit, A., Rhoads, D.D. & Roufa, D.J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6907–6911.
- Rambosek, J.A. & Leach, J. (1987) *CRC Critical Reviews in Biotechnology* **6**, 357–393.
- Tyler, B.M., Geever, R.F., Case, M.E. & Giles, N.H. (1984) *Cell* **36**, 493–502.
- Struhl, K. (1989) *Annu. Rev. Biochem.* **58**, 1051–1077.
- Metzenberg, R.L. & Groteluschen, J. (1987) *Fungal Genet. Newsl.* **34**, 39–44.
- Teem, J.L. & Rosbash, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4403–4407.
- Chen, I.T. & Roufa, D.J. (1988) *Gene* **70**, 107–116.
- Geever, R.F., Huiet, L., Baum, J.A., Tyler, B.M., Patel, V.B., Rutledge, B.J., Case, M.E. & Giles, N.H. (1989) *J. Mol. Biol.* **207**, 15–34.
- Tarawneh, K.A., Wang, Z. & Free, S.J. (1990) *Nucl. Acids Res.* **18**, 7445–7445.
- Stellwag, E.J. & Metzenberg, R.L. (1984) *Mol. Gen. Genet.* **194**, 105–110.
- Woudt, L.P., Smit, A.B., Mager, W.H. & Planta, R.J. (1986) *EMBO J.* **5**, 1037–1040.
- Larkin, J.C., Thompson, J.R. & Woolford, J.L., Jr. (1987) *Mol. Cell. Biol.* **7**, 1764–1775.
- Schwindinger, W.F. & Warner, J.R. (1987) *J. Biol. Chem.* **262**, 5690–5695.
- Dynan, W.S. & Tjian, R. (1985) *Nature* **316**, 774–778.
- Courey, A.J., Holtzman, D.A., Jackson, S.P. & Tjian, R. (1989) *Cell* **59**, 827–836.
- Hamer, J.E. & Timberlake, W.E. (1987) *Mol. Cell. Biol.* **7**, 2352–2359.
- Punt, P.J., Dingemans, M.A., Kuyvenhoven, A., Soede, R.D.M., Pouwels, P.H. & Hondel, C.A.M.J.J.V.D. (1990) *Gene* **93**, 101–109.
- Ballance, D.J. (1986) *Yeast* **229**–236.
- Gurr, S.J., Unkles, S.E. & Kinghorn, J.R. (1988) *Gene structure in eukaryotic microbes* **23**, 93–139.