A *Neurospora crassa* ribosomal protein gene, homologous to yeast *CRY1*, contains sequences potentially coordinating its transcription with rRNA genes

Brett M. Tyler and Karin Harrison*

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

Received June 13, 1990; Revised and Accepted August 31, 1990

EMBL accession no. X53734

ABSTRACT

We have isolated and sequenced a *Neurospora crassa* ribosomal protein gene (designated *crp-2*) strongly homologous to the *rp59* gene (*CRY1*) of yeast and the *S14* ribosomal protein gene of mammals. The inferred sequence of the *crp-2* protein is more homologous (83%) to the mammalian *S14* sequence than to the yeast *rp59* sequence (69%). The gene has three intervening sequences (IVSs) two of which are offset 7 bp from the position of IVSs in the mammalian genes. None correspond to the position of the IVS in the yeast gene. *Crp-2* was mapped by RFLP analysis to the right arm of linkage group III. The 5' region of the gene contains three copies of a sequence, the Ribo box, previously shown to be required for transcription of both 5S and 40S rRNA genes. We speculate that the Ribo box may coordinate ribosomal protein and rRNA gene transcription.

INTRODUCTION

Ribosomes are composed of 70–80 different ribosomal proteins (r-proteins) and four different ribosomal RNAs, present in precise stoichiometric ratios. The genes encoding the ribosomal components are transcribed by RNA polymerase I (large rRNAs), RNA polymerase II (ribosomal proteins) and RNA polymerase III (5S rRNA). When cells are dividing rapidly, rRNA transcription can account for up to 30% of total RNA synthesis and r-proteins can comprise up to 40% of the cell protein (1–4). Consequently, balanced expression of the genes encoding ribosomal components, involving all three RNA polymerases, is critical to the cell. In vertebrates, the transcription rates of some rRNA and r-protein genes are directly co-ordinated during particular stages in development, but also are often distinctly uncoupled (3, 5, 6). In several lower eukaryotes, the regulation of ribosomal RNA and protein syntheses are directly correlated under a variety of growth conditions (4, 7).

In yeast (*Saccharomyces cerevisiae*), rRNA and r-protein synthesis are coordinately modulated according to growth rate, during shifts in carbon and nitrogen nutrition, and upon amino acid starvation. Transcription of the r-protein genes is coordinated

in most cases via one or more copies of a transcriptional activator sequence known as UAS_{rpg} , located in the 5' flanking regions of the genes (8–11, reviewed in 3 & 4). A protein that binds to the UAS_{rpg} DNA sequence, variously called TUF (12), RAP (13, 14) or GRF (15, 16), has been isolated and is presumed to mediate coordinate transcription of the r-protein genes. However, the protein also binds to UAS_{rpg} 's located 5' to several non-ribosomal protein genes, and it does not bind sequences involved in transcription of either the 5S or the large rRNA genes. Therefore its role in linking r-protein gene transcription to rRNA synthesis is unclear. Over-expression of particular r-proteins (17, 18), or under-expression of RNA polymerase I (19), generally cause rapid degradation of proteins not incorporated into ribosomes. For one r-protein, *rpL32*, splicing of the pre-mRNA was inhibited when the protein was overproduced, or rRNA transcription reduced (19, 20).

Vertebrate r-protein genes have complex promoters. Individual genes have different combinations of internal and external elements (21, 22). Many have polypyrimidine tracts and GC-rich sequences at the transcription startpoint (reviewed in 3). However the roles of these sequences, if any, in coordinating the transcription of the genes among themselves or with rRNA genes is unclear. Control of mRNA splicing and translation may play a more important role in the regulation of r-protein synthesis in higher eukaryotes (5, 23, 24, reviewed in 3), but mechanisms, if any, for linking such control to rRNA transcription are unknown.

In the fungus *Neurospora crassa*, as in its distant relative *S. cerevisiae*, the synthesis of rRNA and the levels of ribosomal proteins are regulated by growth rate and by carbon nutrition (1). However, nothing is known about the mechanisms involved. *In vitro* transcription systems have been developed for RNA polymerases I, II and III in *N. crassa* (25–27). These have been used to identify the sequences required for transcription of the 5S rRNA (28) and large (40S) rRNA genes (29). In particular, a common 17–18 bp sequence designated the 'Ribo box' was identified that was required for transcription of both sets of rRNA genes (28, 29). In both cases, small substitutions within the Ribo box reduced transcription 10–20 fold. It was suggested that this sequence might coordinate transcription of the rRNA genes *in vivo* (29).

* Present address: Research School of Biological Sciences, Australian National University, Canberra, ACT 2601, Australia

In order to study the mechanisms by which r-protein genes are regulated in *N. crassa*, and specifically to examine the role, if any, of the Ribo box in r-protein gene transcription, we have isolated and characterized a r-protein gene (*crp-2*) strongly homologous to the S14 gene of mammals and the rp59 gene (*CRY1*) of yeast.

MATERIALS AND METHODS

Cloning and Sequencing of Crp-2 Gene

Screening of the λ phage library with the yeast *CRY1* gene, phage DNA isolation, southern hybridizations, restriction mapping and subcloning were all carried out according to standard procedures similar to those described by Maniatis et al (30). To achieve cross-hybridization between the yeast *CRY1* gene and the *crp-2* gene, hybridizations were carried out at 55°C in a solution containing 3×SSC (0.45 M NaCl, 0.045 M Na₃citrate). Dideoxynucleotide sequencing was carried out using alkali-denatured double stranded plasmid DNA as template (31), and chemical sequencing was carried out according to Maxam and Gilbert (32).

5' End-mapping of RNAs

DNA probes for 5' end-mapping of RNA were labelled by polynucleotide kinase (33). For nuclease mapping, poly(A)⁺ RNA (5μg) was hybridized to the end-labelled DNA at 56°C in 80% formamide, 0.4 M NaCl, 50 mM Na PIPES pH 6.8, 1 mM Na₃EDTA (FNPE) (34), then the hybrids were digested with 50 units of S1 nuclease (Sigma) or 50 units of mung bean nuclease (New England Biolabs) for 30 min. at 37°C (35). For primer extension, the end-labelled double-stranded DNA primer and the RNA were mixed in FNPE and the DNA denatured at 80°C. The hybridization mixture was then incubated overnight in a water bath which was allowed to slowly cool from 65°C to 32°C. The hybrids were recovered by ethanol precipitation and extended with 20 units AMV reverse transcriptase (35). The nuclease digestion and primer extension products were fractionated on an 8% polyacrylamide/7M urea gel and visualized by autoradiography.

Genetic Mapping using Restriction Fragment Length Polymorphisms

Genomic DNAs (provided by R.L. Metzenberg, University of Wisconsin) from 38 progeny from a cross between Oak Ridge (OR) and Mauriceville (M) strains of *N. crassa* were hybridized with a 10 kb *SacI* fragment 5' to the *crp-2* gene. This fragment detects a 4.8 kb *StyI* fragment at the M *crp-2* locus and a 5.4 kb *StyI* fragment at the OR *crp-2* locus. The *crp-2* allele present in each progeny strain was determined. The segregation of M and OR alleles of *crp-2* among the 38 progeny was compared to the known segregations at 55 other loci already on the RFLP map of *N. crassa*. The *crp-2* alleles showed a segregation pattern very similar to that of loci on the right arm of linkage group III, in particular *cyt-8*. Calculation of recombination frequencies with *cyt-8* and other loci on LG IIIIR placed *crp-2* 21 cM right of *thi-1*, 19 cM left of *trp-1*, and 5 cM from *cyt-8*, probably to the right of it.

RESULTS

Cloning and Sequencing a Ribosomal Protein Gene

The ribosomal protein designated rp59 in yeast is highly conserved among eukaryotes (36). Preliminary experiments

showed that a 500 bp *EcoRI*-*BglII* fragment from the yeast rp59 gene (*CRY1*) (37) strongly hybridized to a single 4.7 kb *EcoRI* fragment of *N. crassa* DNA. An *N. crassa* gene bank consisting of partial *Sau3AI* fragments cloned into the λ vector EMBL4 (38) was screened by hybridization with the *CRY1* *EcoRI*-*BglII* fragment. The gene bank (provided by J. Baum and R. Geever, University of Georgia) consisted of six independent sub-libraries each derived from 10⁴ phage clones. Nine strongly hybridizing plaques were identified from the screening, and four of these were selected from independent sub-libraries for further analysis. The restriction maps of the four clones are shown in Figure 1A. These maps clearly indicate that the four clones were derived from the same region of the genome. In particular, all four contained two *EcoRI* fragments of 0.6kb and 4.7kb, at the same position on the restriction map, which hybridized to the *CRY1* probe. The region hybridizing to *CRY1* was subcloned, subjected to detailed restriction mapping, and its DNA sequence determined. Figure 1B shows the restriction map of the gene and the sequencing strategy.

The sequence is shown in Figure 2. The nucleotide sequence of the region hybridizing to *CRY1* shows four segments of strong homology to the yeast rp59 (37) and mammalian S14 ribosomal protein genes (39), which are consistent with a continuous open reading frame separated by three intervening sequences (IVSs) (Figures 1B and 2). The protein encoded by the *N. crassa* sequence shows 69% homology to the yeast rp59 sequence and 83% homology to the hamster S14 sequence (identical amino acids are indicated by asterisks in Figure 2). Accordingly, the *N. crassa* gene is inferred to encode the homologue of the yeast

Figure 1. Cloning of the *crp-2* Ribosomal Protein Gene from *N. crassa*. A. Map of the four overlapping lambda clones of *crp-2*. E = *EcoRI* sites. S = *SacI* sites. Numbers indicate fragment sizes in kilobases. The black box indicates the *crp-2* gene. Dotted lines indicate a 4.7 kb *EcoRI* fragment and a 10 kb *SacI* fragment referred to in the text. B. Restriction map of *crp-2* and sequencing strategy of the 5' end of the gene. The map derives from the complete sequence of the gene (Figures 2 and 4). The exons homologous to the mammalian ribosomal protein S14 mRNA sequence (39) are shown by shaded boxes and the 5' untranslated region by the white box. The black arrow head designates the startpoint of transcription. Only the *TaqI*, *Sau3AI* and *HpaII* sites used for sequencing are shown (marked with asterisk). Black arrows indicate dideoxy sequencing of double-stranded plasmid DNA. Hatched arrows indicate chemical sequencing of 3' end-labelled fragments. Solid bars indicate the hybridization probe and primer used to map the 5' end of *crp-2* mRNA (see Figure 3).

rp59 and mammalian S14 ribosomal proteins, and is designated *crp-2* (Cytoplasmic Ribosomal Protein). The boundary sequences of the inferred IVSs match the consensus sequences for fungi (40), and the conserved heptamer found in *N.crassa* IVSs (40) is also present in each case (underlined in Figure 2). Two of the IVSs occur 7 bp from the sites of IVSs in the human gene. There is no IVS at the site of the single IVS in the yeast *CRY1* gene.

The codon usage of *crp-2* is strongly biased toward C in the third position (or G where C is unavailable) (not shown), as previously described for strongly expressed *N. crassa* genes (40, 41).

The 5' end of the mRNA was mapped from the *Sac*I site in the first exon using S1 and mung bean nuclease digestion of RNA-DNA hybrids and by primer extension. The results are shown in Figure 3. Each of the three mapping experiments

Figure 2. Nucleotide sequence of the *Crp-2* Gene and Comparison with Yeast *rp59* and Human *S14* Protein Sequences. Black dots indicate major transcription startpoints and open diamonds indicate minor starts (see Figure 3). Slashes indicate inferred IVS boundaries. Underlined sequences are conserved heptamers found in *N. crassa* IVSs (40). Black and outline vertical arrows indicate positions of IVSs in the human *S14* gene (52) and the yeast *rp59* gene (37) respectively. Asterisks indicate identities between the predicted *crp-2* protein sequence and the human *S14* (39) and yeast *rp59* (37) protein sequences. The sequence of the 5' flanking region of *crp-2* is presented in Figure 4. The GenBank/EMBL accession number of this sequence is X53734.

Figure 3. Mapping the 5' End of *Crp-2* mRNA by Nuclease Digestion and Primer Extension. *Crp-2* DNA was 5' end-labelled at the *SacI* site (+133), then cleaved upstream at *SmaI* (−154) to create nuclease digestion probes, or at *Fnu4H I* (+81) to create a primer for reverse transcriptase. The fragments were then hybridized to RNA and the hybrids treated as indicated. Tracks 1–3, *E. coli* RNA (5μg); tracks 4–6, *N. crassa* mRNA (5μg); tracks 1 & 4, S1 nuclease digestions; tracks 2 & 5, mung bean nuclease digestions; tracks 3 & 6, primer extensions; tracks AG & CT, A+G & T+C chemical sequencing reactions of the *SacI*–*SmaI* fragment. Transcription start sites were inferred from the primer extension products where they matched nuclease digestion products.

revealed a very similar collection of heterogeneous 5' ends for the *crp-2* mRNA. These were located 25–70 nucleotides (nt) upstream from the presumptive initiation codon, with the major startpoint 64 nt from the AUG. Fungal mRNAs commonly have very heterogeneous 5' ends (35, 42, 43). The agreement between the nuclease digestions and the primer extension indicates that no additional intervening sequences are present in the 5' end of the gene.

A sequence required for 5S and 40S rRNA transcription occurs in the *crp-2* 5' region

The 5' flanking sequence of *crp-2*, shown in Figure 4A, contains three sequences, located at −271, −89 and −9, which are strongly homologous to the Ribo box sequences of the 5S and 40S rRNA genes (28, 29). The orientation of the first two with respect to the direction of transcription is opposite to that in the two rRNA genes. The third Ribo box, which spans the major transcription startpoints of *crp-2*, has the same orientation as in the rRNA genes and displays 15 out of 17 identities with the Ribo box of α 5S genes. This is a stronger homology than that between the 5S and 40S Ribo boxes. The Ribo boxes of the 5S and 40S rRNA genes are directly compared with the *crp-2* −9 Ribo box in Figure 5A. The 5S and *crp-2* Ribo boxes are 17 bp in length while the 40S Ribo box contains an extra residue toward the center of the sequence. The −271 and −89 Ribo boxes of *crp-2*, shown in Figure 5B, are 18 bp in length, like the 40S Ribo box, but each diverges from the consensus at three positions. One other *N. crassa* ribosomal protein gene has been sequenced to date, which is *crp-1*, homologous to yeast *CYH2* (44). Figure 4B shows that the *crp-1* 5' region also contains three Ribo box homologies. Each displays four mismatches to the consensus. Although the *crp-2* Ribo boxes only have up to three mismatches, the occurrence of the Ribo box homologies in *crp-1* is still statistically significant ($p < 0.01$).

Repeated and Conserved Sequences in the *Crp-2* 5' region

Two other sequences are repeated three times in the *crp-2* 5' region, and represent additional candidates for elements involved in transcription of *crp-2*. One, designated the CG repeat, is an 11 bp GC rich sequence (consensus: CCCGCC^A/_GA^A/_GTT) with a strong similarity to the binding site for the human transcription factor Sp1 (^G/_A^C/_T^C/_TCCGCC^G/_T, ref 45). Human Sp1 binds to the CG repeats *in vitro* (Yuguang Shi and BMT, unpublished data), but it is not known if there is an Sp1 analog in *N. crassa*. The CG repeats occur at −192, −74 and −26, all with the same orientation (striped arrows in Figure 4A). Each CG repeat shows a 9/11 match to the other repeats. Two additional CG repeats, at −145 and +55 show homology just to an 8 bp core sequence (CCGCCAAG). The *crp-1* 5' region also shows three sequences with some homology to the CG repeats of *crp-2*, though the homology is less than among the *crp-2* CG repeats. None of fifteen other *N. crassa* genes had significant homology to the CG repeat (one or no mismatches) or to the 8 bp core (no mismatches), though the constraints placed on the computer search may be different than those required by the hypothetical protein which binds the CG repeats. GC-rich sequences are also associated with the promoter sequences of *E. coli* and vertebrate ribosomal protein genes (3, 21, 22). The second repeated sequence, arbitrarily designated the MG repeat, is 17 bp long and occurs at −297, −252, and −44, in both orientations (grey arrows in Figure 4A). Each MG repeat shows 14/17 matches to the other MG repeats. There are no sequences in *crp-1*, or fifteen other *N. crassa* genes examined, homologous to the *crp-2* MG repeats (four or less mismatches). Neither *crp-2* nor *crp-1* contain sequences with significant homology to the yeast UAS_{pg} sequence, to the UAS_T sequence of the yeast *TCM1* ribosomal protein gene (46) or to the GF1 site of ribosomal protein genes S33 and L3 (47). *Crp-2* contains no sequence like a TATA box except an AT rich sequence at −19 (Figure 4A), while *crp-1* contains a TATA-like sequence (TAAATTT) at −42 (44).

Besides the Ribo box and CG repeat homologies, there are

A

B

Figure 4. Potential Transcriptional Elements in the 5' Flanking Regions of *Crp-2* and *Crp-1*. A. *Crp-2* 5' flanking region. B. *Crp-1* 5' flanking region. *Crp-2* sequence is from this study. *Crp-1* sequence is from Kreader and Heckman (44). Black arrows with white stripes indicate Ribo boxes. White arrows with black stripes indicate CG repeats in *crp-2* and sequences with partial homology to CG repeats in *crp-1*. Thin dashed arrows indicate additional sequences in *crp-2* with homology to the eight bp core of the CG repeats. Grey arrows indicate MG repeats. Boxed sequences are conserved between *crp-2* and *crp-1*. Within the boxes, lower case letters indicate mismatches between the *crp-2* and *crp-1* sequences. The GenBank/EMBL accession number of the *crp-2* sequence is X53734.

three single copy sequence elements that are conserved in DNA sequence and position in the 5' regions of the two genes (boxed in Figures 4A and B). The first, located at -111 in *crp-2* and -105 in *crp-1*, shows 9/11 bp identities between the *crp-2* and *crp-1* sequences. The second, at -83 in *crp-2* and -75 in *crp-1*, shows 14/17 identities. This overlaps a Ribo box and a CG repeat in *crp-2* and a CG repeat in *crp-1*. The third, at -27 in *crp-2* and -22 in *crp-1*, shows 7/8 identities and spans a CG repeat from both genes.

Genetic Mapping of *crp-2*

There is a single copy of the *crp-2* gene in the *N. crassa* genome, based on gel blot hybridization experiments (data not shown). To place *crp-2* on the *N. crassa* genetic map, a 10 kb *SacI* fragment 5' to the *crp-2* gene (Figure 1A) was hybridized to DNAs from 38 progeny of a cross between *N. crassa* Oak Ridge and Mauriceville strains (DNAs generously provided by R.L. Metzenberg, University of Wisconsin, Madison). This progeny set has previously been used to construct a restriction fragment length polymorphism (RFLP) genetic map of *N. crassa* (48). The

10 kb *SacI* fragment 5' to *crp-2* detects a polymorphism among *StyI* digests of the DNAs (Figure 6A). The genetic segregation of this polymorphism placed *crp-2* on the right arm of chromosome III, 21 cM right of *thi-1*, 19 cM left of *trp-1*, and 5 cM from *cyt-8*, probably to the right of it. Alignment of the RFLP map (48) with the genetic map (49), as shown in Figure 6B, suggests that *crp-2* may be very close to *ad-4*, but the two loci have not been directly mapped with respect to each other.

DISCUSSION

We have cloned a ribosomal protein gene, *crp-2*, from *N. crassa* that is strongly homologous to the *CRY1* (rp59) gene of yeast and the S14 gene of mammals. *Crp-2* is present in a single copy in the *N. crassa* genome, like the other cloned *N. crassa* ribosomal protein gene, *crp-1* (44). This contrasts with yeast where many ribosomal protein genes are duplicated (4). An absence of duplicated genes may result from the efficient mechanism *N. crassa* possesses for inactivating duplicated sequences (50).

The ribosomal proteins homologous to *crp-2* are highly

A	α 5S rRNA (pol III)	+18	CTGGAAACTCG-GGATC >
	β 5S rRNA (pol III)	+18	GTAGAAACTCG-GGATC >
	40S rRNA (pol I)	-65	GTGGAAAGTCGGGCTC >
	<i>crp-2</i> (pol II)	-9	CTGGAAACTGG-AGATC >
	consensus		$\begin{matrix} \text{G} \\ \text{C} \end{matrix} \text{TAGAAA} \begin{matrix} \text{C} \\ \text{G} \end{matrix} \text{C} \begin{matrix} \text{G} \\ \text{G} \end{matrix} \text{G} \begin{matrix} \text{A} \\ \text{C} \end{matrix} \text{TC} >$
B	<i>crp-2</i>	-271	< CTGGcAtcGTCGCAGATC
	<i>crp-2</i>	-89	< GgGGAAAGTGGCAGaa
	<i>crp-1</i>	-457	CT ^g GGAAGACaGG-GGAac >
	<i>crp-1</i>	-384	< GcTGcAAGaTCG-GGCTC
	<i>crp-1</i>	-270	GaAGgcAAGaGG-GGCTC >

Figure 5. Consensus Sequence for the Ribo Box. A. Derivation of a Ribo box consensus from the Ribo boxes of α and β 5S rRNA genes, the 40S rRNA gene and the *crp-2* Ribo box at -9. α and β are different sub-families of *N.crassa* 5S genes (57) which are transcribed with equal efficiency *in vitro* (25). The consensus includes all bases which appear among the four Ribo box sequences. '-' indicates the position of an additional base pair in the 40S Ribo box. B. Additional Ribo boxes in the *crp-1* and *crp-2* 5' flanking sequences. Lower case letters indicate mismatch with the consensus. < and > indicate relative orientations. The putative -457 Ribo box of *crp-1* has an extra G residue at position 3.

Figure 6. Genetic mapping of *Crp-2*. A. Segregation of an RFLP adjacent to *crp-2*. DNAs from 38 progeny of an Oak Ridge (OR) x Mauriceville (M) cross (48) (kindly provided by R.L. Metzberg, University of Wisconsin, Madison) were digested with *SfiI* and hybridized with a 10 kb *SacI* fragment 5' to *crp-2* (see Figure 1). Most digestions show additional bands due to partial digestion of the DNA. The polymorphic products of complete digestion are indicated by the arrows. OR and M indicate DNA from the Oak Ridge and Mauriceville parents. A1, A4, B6 etc indicate progeny DNAs. Progeny with same letter code (e.g. A1 and A4) originate from the same ordered tetrad (48). B. Alignment of the genetic map (linkage group IIIR) (49) with the RFLP map (48) containing *crp-2*. The circle indicates the centromere. The maps were aligned at *thi-4* and *trp-1*. *Crp-2* and *ad-4* appear to precisely coincide, but errors in the two maps may place them several centimorgans apart.

conserved throughout evolution (36). Intriguingly, the *N.crassa* *crp-2* DNA and protein sequences are substantially more homologous to the mammalian sequences than to those of the related ascomycete fungus *Saccharomyces cerevisiae* (83% compared to 69%, respectively, at the protein level; 71% compared to 60%, at the DNA level). The glyceraldehyde-3-phosphate dehydrogenase genes of filamentous fungi also are more closely related to vertebrates than to yeast (51). In contrast however, the *N.crassa* *crp-1* (44) and *crp-3* (Y. Shi and B.M. Tyler, unpublished) protein sequences are more similar to

their yeast counterparts (*CYH2* and *rp51*, respectively) than their mammalian counterparts (L27a and S17, respectively). Two IVSs occur in similar positions in the *N.crassa* *crp-2* and human S14 (52) genes. Highly conserved IVS positions have been reported in other systems (53). However, in this case, the positions of both IVSs are offset by 7 bp in the respective *N.crassa* and human mRNA sequences (Figure 2). This unusual offset may have resulted if a small sequence duplication occurred at the ends of each IVS when it originated in the ancestral *crp-2*/S14 gene (or subsequently to its origination). A different evolutionary choice of splice site within the duplications in fungi and mammals could have led to the present differences. These two IVSs are completely absent from both the yeast *CRY1* gene (37) and a *Drosophila* S14 gene (54).

A key interest in cloning this ribosomal protein gene was to look for possible sequences involved in coordinating the transcription of *N.crassa* r-protein genes with each other and with ribosomal RNA genes. In yeast, most ribosomal protein genes are coordinated via the UAS_{rpg} element (8, 11, 37, 55, 56). However, the *crp-1* and *crp-2* genes contain no sequences with statistically significant homology to the UAS_{rpg} consensus; the best homologies have only 11/15 matches. Instead, three sequences different to the UAS_{rpg} consensus are conserved between the *crp-1* and *crp-2* promoter regions in both DNA sequence and position (Figure 4). The longest of these shows 14/17 matches and is located at -75 in *crp-1* and -83 in *crp-2*. The strong conservation of the DNA sequences and relative positions of the three elements suggest they may be involved in the transcription of the *crp-1* and *crp-2* genes. Examination of further *N.crassa* ribosomal protein genes, along with functional assays, will be required to test whether this is correct and whether the sequences are involved in coordinating *N.crassa* r-protein gene transcription in general.

Sequences directly coordinating transcription of ribosomal protein genes with that of rRNA genes have not been reported in eukaryotes (see Introduction). We showed previously that a sequence, designated the 'Ribo box', is present in the promoters of both 5S rRNA genes and 40S rRNA genes (25, 26, 28, 29). We showed that the Ribo box either is required directly for transcription of these genes *in vitro*, or overlaps with sequences required for transcription (28, 29). The 5S and 40S Ribo boxes, though differing at five positions, were functionally interchangeable (29). The cloning of the *crp-2* gene, described here, was initiated to test the hypothesis that, in *N.crassa*, the Ribo box was also involved in ribosomal protein gene transcription. In fact, we have found that the *crp-2* gene contains three sequences with strong homology to the Ribo box, at -271, -89 and -9. The -9 Ribo box spans the transcription initiation sites and has 15/17 matches to the α 5S Ribo box. The -271 and -89 Ribo boxes show three mismatches to a Ribo box consensus derived from the 5S, 40S and *crp-2* -9 Ribo boxes. The probability that there would be at least one sequence with 15/17 matches to either the 5S or 40S Ribo boxes anywhere in the 600 bp surrounding the initiation site, is 3×10^{-7} . The probability of finding two additional sequences with only three mismatches to the consensus is 3×10^{-6} . *Crp-1* also contains three Ribo boxes. Although these show less homology to the consensus (4 mismatches) than the *crp-2* Ribo boxes, their occurrence is statistically significant ($p = 0.008$). These statistics are persuasive that the Ribo box is involved in the transcription of *crp-2*, and possibly other ribosomal protein genes. In particular, the Ribo box may coordinate transcription of r-protein

genes with that of 5S and 40S rRNA genes. We are currently carrying out functional assays to test this hypothesis.

ACKNOWLEDGEMENTS

I thank John Woolford (Carnegie-Mellon University) for the gift of the yeast *rp59* gene; Jim Baum and Robert Geever (University of Georgia) for the *N. crassa* λ phage library; Robert Metzner (University of Wisconsin, Madison) for DNAs from RFLP-mapping progeny; Eva Steinberger, Yuguang Shi and Layne Huiet for helpful discussions; Yuguang Shi, Eva Steinberger, Tom Cujec and Julie Billington for criticizing the manuscript; and Jeff Hall for photography. This work was supported by National Institutes of Health grants GM36964 and GM42178 (to B.M.T.) and by the Research School of Biological Sciences, Australian National University.

REFERENCES

- Alberghina, L. & Sturani, E. (1981) *Microbiol. Rev.* **45**, 99–122.
- Kief, D.R. & Warner, J.R. (1981) *Mol. Cell. Biol.* **1**, 1007–1015.
- Mager, W.H. (1988) *Bioch. Bioph. A.* **949**, 1–15.
- Warner, J.R. (1989) *Microbiol. Rev.* **53**, 256–271.
- Wormington, W.M. (1988) *Devel. Biol.* **5**, 227–240.
- Wormington, W.M. (1989) *Mol. Cell. Biol.* **9**, 5281–5288.
- Alberghina, L., Sturani, E., Costantini, M.G., Martegani, E. & Zippel, R. (1978) In Burnett, J.H. & Trinci, A.P.J. (ed.), *Fungal Walls and Cell Growth*. Cambridge University Press, London, pp. 295–318.
- 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.
- Rotenberg, M.O. & Woolford, J.L.J. (1986) *Mol. Cell. Biol.* **6**, 674–687.
- Woudt, L.P., Smit, A.B., Mager, W.H. & Planta, R.J. (1986) *EMBO J.* **5**, 1037–1040.
- Nieuwint, R.T.M., Mager, W.H., Maurer, C.T. & Planta, R.J. (1989) *Curr. Genet.* **15**, 247–251.
- 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.
- 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.
- Warner, J.R., Mitra, J.R., Schwindinger, W.F., Studeny, M. & Fried, H.M. (1985) *Mol. Cell. Biol.* **5**, 1512–1521.
- Tsay, Y.F., Thompson, J.R., Rotenberg, M.O., Larkin, J.C. & Jr, W.J.L. (1988) *Genes and Devel.* **2**, 664–667.
- Wittekind, M., Kolb, J.M., Dodd, J., Yamagishi, M., Memet, S., Buhler, J.M. & Nomura, M. (1990) *Molecular and Cellular Biology* **10**, 2049–2059.
- Dabeva, M.D., Post-Beittenmiller, M.A. & Warner, J.R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5854–5857.
- Hariharan, N., Kelley, D.E. & Perry, R.P. (1989) *Genes and Devel.* **3**, 1789–1800.
- Hariharan, N. & Perry, R.P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1526–1530.
- Caffarelli, E., Fragapane, P., Gehring, C. & Bozzoni, I. (1987) *EMBO J.* **6**, 3493–3498.
- Amaldi, F., Bozzoni, I., Beccari, E. & Pierandrei-Amaldi, P. (1989) *Trends Bioc. Sci.* **14**, 175–178.
- Tyler, B.M. & Giles, N.H. (1984) *Nucl. Acids Res.* **12**, 5737–5757.
- Tyler, B.M. & Giles, N.H. (1985) *Nucl. Acids Res.* **13**, 4311–4332.
- Tyler, B.M. & Giles, N.H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5450–5454.
- Tyler, B.M. (1987) *J. Mol. Biol.* **196**, 801–811.
- Tyler, B.M. (1990) *Nucl. Acids Res.* **18**, 1805–1811.
- Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Chen, E.Y. & Seeburg, P.H. (1985) *DNA* **4**, 165.
- Maxam, A.M. & Gilbert, W. (1980) *Meths. Enzymol.* **65**, 455–560.
- Huiet, R.L., Tyler, B.M. & Giles, N.H. (1984) *Nucl. Acids Res.* **12**, 5757–5765.
- Favaloro, J., Treisman, R. & Kamen, R. (1980) *Meths. Enzymol.* **65**, 718–749.
- Tyler, B.M., Geever, R.F., Case, M.E. & Giles, N.H. (1984) *Cell* **36**, 493–502.
- Chen, I.-T., Dixit, A., Rhoads, D.D. & Roufa, D.J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6907–6911.
- Larkin, J.C., Thompson, J.R. & Woolford, J.L., Jr. (1987) *Mol. Cell. Biol.* **7**, 1764–1775.
- Frischauf, A.M., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170**, 827–842.
- Rhoads, D.D. & Roufa, D.J. (1985) *Mol. Cell. Biol.* **5**, 1655–1659.
- Rambossek, J.A. & Leach, J. (1987) *CRC Critical Reviews in Biotechnology* **6**, 357–393.
- Kinnaird, J.H. & Fincham, J.R.S. (1983) *Gene* **26**, 253–260.
- Faye, G., Leung, D.W., Tatchell, K., Hall, B.D. & Smith, M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2258–2262.
- Boel, E., Hansen, M.T., Hjort, I., Hoegh, I. & Fiil, N.P. (1984) *EMBO J.* **3**, 1581–1585.
- Kreder, C.A. & Heckman, J.E. (1987) *Nucl. Acids Res.* **15**, 9027–9042.
- Dynan, W.S. & Tjian, R. (1985) *Nature* **316**, 774–778.
- Hamil, K.G., Nam, H.G. & Fried, H.M. (1988) *Mol. Cell. Biol.* **8**, 4328–4341.
- Dorsman, J.C., Doorenbosch, M.M., Maurer, C., de Winde, J., Mager, W.H., Planta, R.J. & Grivell, L.A. (1989) *Nucl. Acids Res.* **17**, 4917–4923.
- Metzenberg, R.L. & Groteluschen, J. (1987) *Fungal Genet. Newsl.* **34**, 39–44.
- Perkins, D.D., Radford, A., Newmeyer, D. & Björkman, M. (1982) *Microbiol. Rev.* **46**, 426–570.
- Selker, E.U., Cambareri, E.B., Jensen, B.C. & Haack, K.R. (1987) *Cell* **51**, 741–752.
- Smith, T.L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7063–7066.
- Rhoads, D.D., Dixit, A. & Roufa, D.J. (1986) *Mol. Cell. Biol.* **6**, 2774–2783.
- McKnight, G.L., O'Hara, P.J. & Parker, M.L. (1986) *Cell* **46**, 143–147.
- Brown, S.J., Rhoads, D.D., Stewart, M.J., B., v.S., Chen, I.T., Johnson, T.K., Denell, R.E. & Roufa, D.J. (1988) *Mol. Cell. Biol.* **8**, 4314–4321.
- Schwindinger, W.F. & Warner, J.R. (1987) *J. Biol. Chem.* **262**, 5690–5695.
- Woudt, L.P., Mager, W.H., Nieuwint, R.T.M., Wassenaar, G.M., Van der Kuyl, A.C., Murre, J.J., Hoekman, M.F.M., Brokhoff, P.G.M. & Planta, R.J. (1987) *Nucl. Acids Res.* **15**, 6037–6048.
- Selker, E.U., Yanofsky, C., Driftmeier, K., Metzner, R.L., Alzner-DeWeerd & RajBhandary, U.L. (1981) *Cell* **24**, 819–828.