Arginine-specific Negative Regulation of *Neurospora crassa arg-2* Mediated by the *arg-2* uORF and Arginine

Julie Marie Carroll
B.S., St. Martin's College, 1994

A thesis presented to the faculty of the
Oregon Graduate Institute of Science and Technology
in partial fulfillment of the
requirements for the degree
Master of Science
in
Molecular Biology

June 1997

The thesis "Arginine-specific Negative Regulation of *Neorospora crassa* *arg-2* Mediated by the *arg-2* uORF and Arginine" by Julie Carroll has been examined an approved by the following Examination Committee:

Matthew S. Sachs, Thesis Advisor
Associate Professor

James M. Cregg
Professor

Margaret Alic
Research Scientist

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Matthew Sachs, for his instruction and guidance while in his laboratory. Thanks also to Zhong and Zongli for help learning new techniques. I greatly appreciate the input of my committee regarding this manuscript, especially Dr. Margaret Alic whose comments and editing were extremely helpful. I also appreciate the help of the support staff for the department, who keep things running smoothly. Finally, thanks to my parents for their financial and emotional support and to Shula and Chris who let me vent my frustrations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Arginine-specific Negative Regulation of *Neurospora crassa arg-2* Mediated by the *arg-2* uORF and Arginine

Julie Carroll, M.S.

Supervising Professor: Matthew S. Sachs

Neurospora crassa arg-2 encodes the small subunit of Arg-specific carbamoyl phosphate synthetase and is negatively regulated by arginine. This regulation is mediated by a 24-codon upstream open reading frame (uORF). The sequence of this uORF is critical for Arg-specific regulation. Six mutated templates were used to examine which residues of the uORF are important for this regulation. Mutations were created using megaprimer PCR and a luciferase gene was used as a reporter in the *in vitro* translation studies. Mutations of Asp 12, Asp 16, and Ser 10 eliminate Arg-specific regulation. Leaky scanning is thought to be involved, and a hypothetical ribosome stalling model that mediates Arg-specific attenuation of translation is proposed.

CHAPTER 1
INTRODUCTION

1.1 Carbamoyl Phosphate Synthetase

Carbamoyl phosphate is the activated carbamoyl donor in the reaction where the carbamoyl group is transferred to ornithine to form citrulline in the urea cycle. Carbamoyl phosphate has a high transfer potential because of its anhydride bond. Carbamoyl phosphate is synthesized from NH_4^+, CO_2, H_2O, and ATP in a reaction catalyzed by carbamoyl phosphate synthetase. The reaction requires *N*-acetylglutamate, and since it consumes two molecules of ATP it is essentially irreversible (Figure 1.1).

Carbamoyl phosphate also plays a role in pyrimidine biosynthesis. CO_2, glutamate, and ATP are combined to make carbamoyl phosphate in a reaction catalyzed by carbamoyl synthetase. The committed step in pyrimidine nucleotide biosynthesis is the formation of *N*-carbamoylaspartate from carbamoyl phosphate and aspartate in a reaction catalyzed by aspartate transcarbamoylase.

The third role of carbamoyl phosphate is as an intermediate in the biosynthesis of arginine, which is the function of *arg-2* in *Neurospora crassa*. Eukaryotes contain two carbamoyl phosphate synthetases (CPSs), whereas most prokaryotes contain only one CPS. *Neurospora* CPS-A is a two-subunit enzyme that is located in the mitochondrial matrix and functions in arginine biosynthesis, and CPS-P is located in the nucleus and functions in pyrimidine biosynthesis. The CPS-A small subunit is encoded by *arg-2*, the CPS-A large subunit is encoded by *arg-3*, and CPS-P by *pyr-3* in *N. crassa* (Davis, 1986).

In general, the level of CPS-A activity is the rate-determining component in the arginine biosynthetic pathway (Davis and Ristow, 1987). The activity is

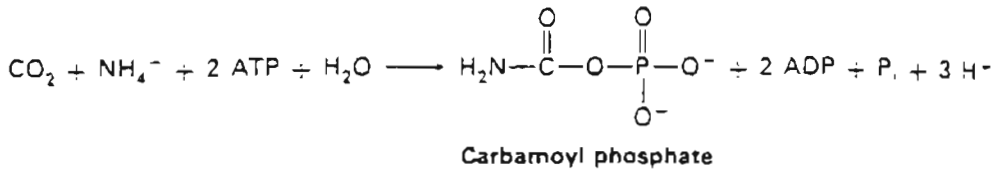


Figure 1.1. The carbamoyl phosphate synthesis reaction.

proportional to the level of the *arg-2* small subunit. *arg-2* is the only gene encoding a *N. crassa* arginine biosynthetic enzyme that is subject to negative regulation by arginine (Davis, 1986). Both the transcription and translation of *arg-2* are negatively affected by arginine (Orbach et al., 1990; Luo et al., 1995).

1.2 The *arg-2* Gene

The *N. crassa arg-2* gene encodes the small subunit of arginine-specific carbamoyl phosphate synthetase which functions as a glutamine amidotransferase. The level of *arg-2* expression regulates arginine biosynthesis in *N. crassa* (Davis and Ristow, 1987).

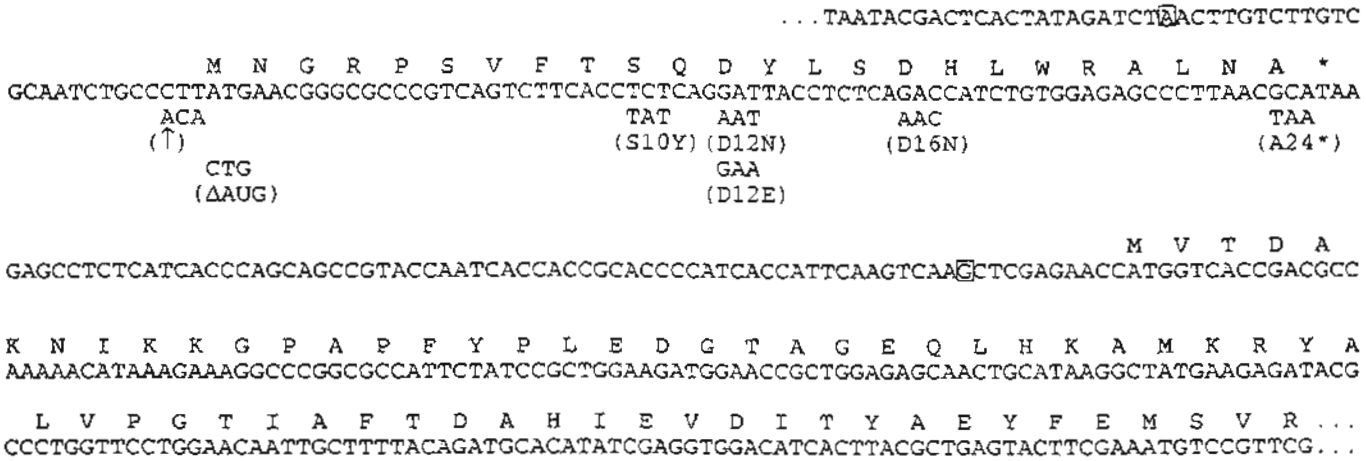
The *arg-2* polypeptide's predicted amino acid sequence of 453 residues is 56% and 36% identical with the sequences of homologous polypeptides of *Saccharomyces cerevisiae* and *Escherichia coli,* respectively (Orbach et al., 1990). The polypeptide also has an additional amino-terminal domain that contains the features of a mitochondrial signal sequence. The *arg-2* mRNA contains a 24-codon upstream open reading frame (uORF) that is nearly identical with the uORF of the homologous *S. cerevisiae* and *Magnaporthe grisae* transcripts (Figure 1.2) (Orbach et al., 1990; Shen and Ebbole, 1997), and is essential for translational regulation in response to arginine availability (Luo et al., 1995).

1.3 Translational Regulation

1.3.1 Introduction to translational regulation

Gene expression can be regulated at many levels. It can be controlled through gene rearrangements, transcription, translation and post-translational processing. Control at the level of protein synthesis allows cells to respond rapidly to changes in physiological conditions, since activation or repression of mRNAs can occur essentially instantaneously. It seems logical that the simplest place to control protein synthesis would be at the level of initiation. In most cases, regulation of translation

A



B

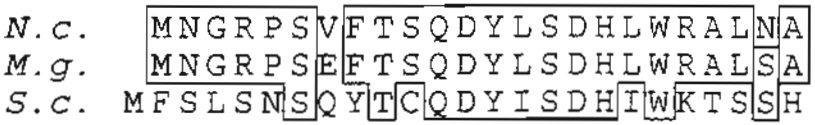


Figure 1.2. Sequences of the 5′-region of the *arg-2-LUC* gene and comparisons of *arg-2* uORF-related peptides. **(A)** 5′-regions of wild-type and mutant *arg-2-LUC* templates. The sequence shown begins with the T7 RNA polymerase binding site and ends within the luciferase coding region. The 5′- and 3′- boundaries of the *arg-2* region are indicated by boxes. The amino acid sequences of the *arg-2* uORF and the amino-terminus of luciferase are shown above the nucleotide sequence. The positions of specific mutations and their predicted consequences for uORF translation are indicated below the sequence of the wild-type template. The (↑) mutation improves the initiation context for uORF translation (Wang & Sachs, 1997). **(B)** Alignment of the peptide sequences encoded by the uORFs in the 5′-leader of transcripts of the homologous genes *N. crassa arg-2* (N.c.; Orbach et al., 1990), *M. grisae arg-2* (M.g.; Shen & Ebbole, 1997), and *S. cerevisiae CPA1* (S.c.; Werner et al., 1987) (Wang & Sachs, 1997a).

begins at the level of initiation. It is more efficient to control a pathway at its point of origin than to interrupt it midway and have to deal with the buildup of intermediates and other components. However, there are cases where regulation occurs at later steps in the translational pathway, especially at the elongation level (Hershey et al., 1996).

Translational control is the change in the efficiency of translation, or the number of amino acids polymerized per unit time per mRNA molecule. A change in the efficiency of translation will affect the overall amount of proteins synthesized or the species of proteins produced (Hershey et al., 1986). The overall rate of protein synthesis is proportional to the number of actively synthesizing ribosomes and the average rate at which they polymerize amino acids (elongation rate) (Hershey et al., 1986).

1.3.2 Translational control by uORFs

Translational control includes the control of the synthesis of polypeptides from mRNA and the control of the stability of mRNA that is mediated through translational effects. In comparison to transcriptional control mechanisms, translational control mechanisms are poorly understood.

A uORF is an uAUG codon within the 5′-leader of an mRNA, followed by an open reading frame and a termination codon different from that of the major coding ORF. Non-overlapping and overlapping uORFs are distinguished by the position of the termination codon relative to the initiation codon of the downstream ORF. The potential peptide product of the uORF will be entirely different from that of the downstream ORF (Geballe, 1996).

Approximately 10% of eukaryotic gene transcripts contain uAUG codons (Kozak, 1989), and uAUGs are more frequent in certain subsets of genes (e.g., 42% of *Drosophila* genes have transcripts containing uAUG codons). Despite the increase in nucleotide sequence data, there has been no recent collection of uAUG codons among eukaryotic genes for a number of reasons:

1. Entries in sequence databases do not consistently include the entire or correct 5′-leader sequences.

2. Numerous genes express more than one transcript, either by alternative transcript initiation sites or post-transcriptional processing events.
3. Non-AUG codons sometimes function as initiation codons.

Therefore, upstream non-AUG initiation codons may be more common than is currently appreciated and eukaryotic transcripts that are now considered simple monocistronic mRNAs may have a more complex polycistronic structure (Geballe, 1996).

Some uORFs have little or no effect on the translation of the downstream uORF, although the uAUG codon is recognized efficiently (Geballe, 1996). Others, like cytomegalovirus gp48 uORF2, have the ability to reduce translation from the downstream AUG even though initiation at the uAUG is not efficient (Cao and Geballe, 1995). Therefore, it would seem that other events during translation in addition to initiation determine the overall effect of the uORF (Cao and Geballe, 1995). The inhibitory effect of gp48 uORF2 depends on its amino acid coding sequence. It also requires a termination codon at a specific position relative to the uORF coding sequences; extension at the 3′ end by a single codon eliminates uORF inhibition (Cao and Geballe, 1995).

1.3.3 The scanning model of eukaryotic translation

The 40S ribosomal subunit with its associated factors initially binds the mRNA at or near the cap and then scans in a 3′ direction. When the first initiation codon is encountered, the 60S subunit joins the 40S subunit to form a complete 80S ribosome.

The scanning model predicts that sequence elements within mRNA 5′-leaders have the potential to control the access of ribosomes to the downstream coding ORF. In fact, many systems have a variety of regulatory elements within 5′-leaders. However, the scanning mechanism, while it accounts for many observations, is not universally used, and alternative mechanisms are clearly involved in translation of some eukaryotic mRNAs (Geballe, 1996). The scanning model is compatible with many methods of negative regulation including inhibition by upstream AUG codons, repressor proteins, and secondary structure within the 5′-noncoding region (Geballe, 1996). The scanning model predicts that expression of a downstream ORF will vary

inversely with initiation at the uAUG codon. However, several refinements and additions to the scanning model describe mechanisms that counter the expected inhibitory impact of uAUG codons (Figure 1.3).

1.3.4 Mechanisms that counter the expected inhibitory impact of uAUG codons

1.3.4.1 Leaky scanning. Ribosomes may ignore the 5′-proximal AUG codon. If the nucleotides flanking the AUG codon are unfavorable for initiation or if the uAUG codon is less than 10 nucleotides from the 5′ end of the mRNA, the 40S subunit may fail to recognize the uAUG codon and continue scanning to the next AUG codon (Kozak, 1989).

Parameters other than the primary sequence may influence efficiency of uAUG recognition. For example, secondary structure in the RNA immediately downstream from a poor context uAUG codon may retard ribosomal scanning and thereby facilitate initiation at the uAUG codon (Kozak, 1989)

1.3.4.2 Ribosome reinitiation. Efficient reinitiation by the ribosomes could occur if:

1. The start codon upstream of the major coding sequence is in a good initiation context.
2. A termination codon is in frame with the upstream AUG codon.
3. The termination codon lies upstream from the downstream AUG.

The ribosomes can translate the first ORF and then continue scanning and reinitiate downstream. The efficiency of reinitiation increases as the distance from the downstream start codon and the uORF lengthens (Kozak, 1989).

1.3.5 *GCN4*

A unique translational control mechanism involving reinitiation occurs in the *S. cerevisiae GCN4* gene (Grant et al., 1995). The *GCN4* RNA has four uORFs located upstream of the major coding sequence. When amino acids are abundant, *GCN4* translation is repressed, mainly through uORF4 (Grant et al., 1995). When the yeast is starved for one or more amino acids, the repression is lifted, allowing translation of

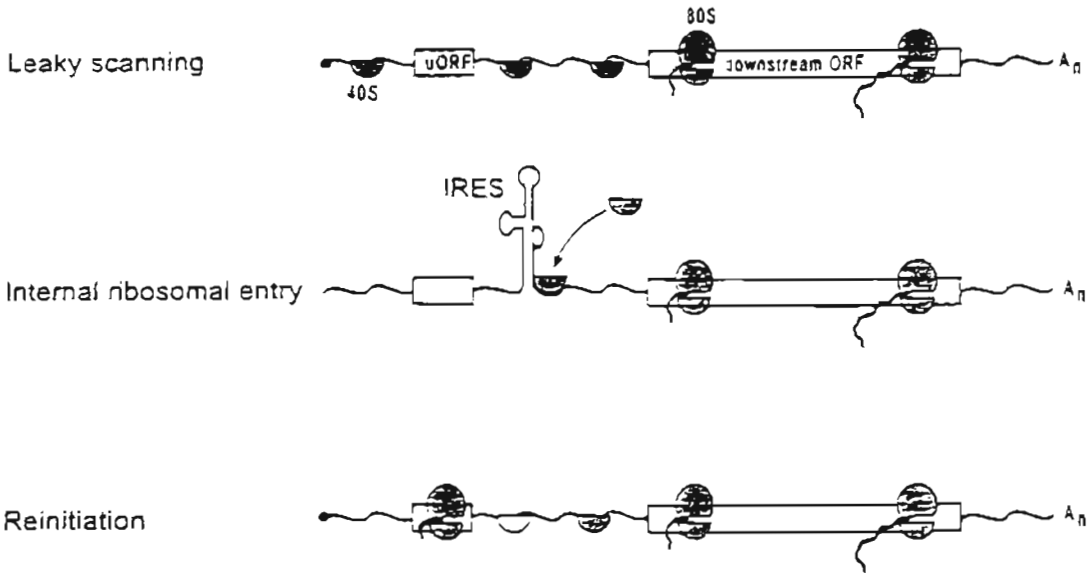


Figure 1.3. Translational mechanisms of avoiding the inhibitory effects of an uAUG codon. Ribosomes may (1) scan past an uAUG codon without initiating until reaching the downstream AUG codon (leaky scanning), (2) initiate at an internal ribosomal entry site (IRES), or (3) initiate at an uAUG codon and translate the uORF, but then reinitiate at a downstream AUG codon (Geballe, 1996).

GCN4. The *GCN4* protein is a transcription factor that activates transcription of a variety of amino acid biosynthetic genes to restore the amino acid supply (Grant et al., 1995). Ribosomes that translate uORF1 continue scanning and reinitiate downstream. Under non-starvation conditions, reinitiation occurs at one of the remaining three uORFs and *GCN4* is then repressed. When grown under amino acid limitation, ribosomes bypass the uORFs and reinitiate at *GCN4*. There is strong evidence that the sequence 5′ to uORF1 enhances reinitiation. Deletions and insertions in the leader 5′ to uORF1 decreased reinitiation at *GCN4* (Grant et al., 1995). The sequence may influence the probability of ribosome release following peptide termination at uORF1, or it may assist in the rebinding of an initiation factor required for reinitiation prior to the continuation of scanning (Grant et al., 1995). Regulation of *GCN4* is independent of the sequence of the uORFs. Instead, *GCN4* regulation depends on the spacing between the uORFs and on sequences following the termination codon (Geballe, 1996).

1.3.6 *CPA1*

5′-leader sequences flanking the sequence-dependent uORFs do not appear to be involved in the translational regulation. *CPA1* is the *S. cerevisiae* homolog of *arg-2*. The expression of this gene is presumed to be repressed by arginine at a translational level. The *CPA1* mRNA contains a 25-codon uORF which confers negative regulation by arginine, as shown by a series of mutations made in the uORF. The primary sequence of the uORF is essential for this regulation (Delbecq et al., 1994). In contrast to the uORF of *gp48,* the *CPA1* uORF termination codon does not appear to be necessary for the uORF-mediated inhibition. The expression of an in-frame fusion of the uORF to a downstream reporting ORF, eliminating the termination codon, was inhibited (Delbecq et al., 1994). However, effects on RNA levels were not considered in these studies.

Analyses of the *CPA1* 5′-leader provided elegant confirmation of the regulatory significance of this uORF (Werner, 1987). In the presence of arginine, the *CPA1* 5′-leader inhibits downstream expression. All *cis*-acting mutants in strains

selected *in vivo* for constitutive expression of the downstream cistron mapped to the uORF. In addition to nonsense and uAUG codon mutations, missense mutations of codons 11 and 13 were isolated (Geballe, 1996).

1.3.7 Possible mechanisms of inhibition by sequence-dependent uORFs

It is reasonable to think that for sequence-dependent uORFs, the peptide product of the uORF mediates regulation of the downstream ORF. However, peptide products of sequence-dependent uORFs have not been isolated. If, in fact, it is the peptide products that mediate inhibition, then the observation that sequence-dependent uORFs seem to function only in *cis* is confusing. However, the peptide product might be synthesized and released, but only in a concentration sufficient enough to inhibit translation in its local microenvironment (Geballe, 1996).

Another way the peptide could cause inhibition of the major coding sequence is that the nascent peptide may mediate repression while still connected to the ribosome. For example, if the peptide interacted with the ribosome or a ribosome-associated translation factor, termination could be prevented or delayed due to ribosome stalling or another similar mechanism. Termination is thought to be relatively slow and so it could conceivably be subject to regulation (Geballe, 1996). Alternatively, the translation of certain codons may be sufficient for repression with the peptide being a byproduct, as in transcriptional attenuation in bacteria (Geballe, 1996).

1.4 Arginine Regulation of *arg-2*

Three separate control mechanisms regulate *arg-2* expression:

1. *arg-2* is the only component of the arginine biosynthetic pathway known to be negatively regulated by arginine (Davis, 1986).
2. Like many genes involved in *Neurospora* amino acid metabolism, *arg-2* is positively regulated in response to amino acid starvation. The level of *arg-2* mRNA increases with amino acid starvation. Ccp1p, the product of the cross-

pathway control gene, *cpc-1*, is important for this response (Orbach et al., 1990).

3. *arg-2* is developmentally regulated in that the highest level of *arg-2* mRNA is found during spore germination and early exponential growth (Sachs and Yanofsky, 1991).

arg-2 mRNA levels are reduced by arginine supplementation and increased by amino acid limitation. This increase is not observed in a strain containing the *cpc-1* mutation (Orbach et al., 1990). *cpc-1*, the *N. crassa* homolog of *GCN4*, encodes a transcriptional activator that positively regulates gene expression in response to amino acid starvation. *cpc-1* also contains a uORF. The lack of an increase in mRNA levels in the Δ*cpc-1* strain indicates that the cross-pathway control system participates in *arg-2* regulation. Four copies of the sequence TGACTC, the binding site for the CPC1 regulatory protein, are found in the *arg-2* genetic region. Two copies are located upstream of the mRNA start sites and two are present within introns in the *arg-2* uORF (Orbach et al., 1990). Synthesis of both the small and large subunits (*arg-2* and *arg-3*) is subject to cross-pathway-mediated induction under conditions of amino acid limitation (Davis, 1986). Synthesis of the small subunit is also subject to arginine-specific negative regulation. Growth of *N. crassa* in the presence of excess Arg results in a 5- to 10-fold decrease in small subunit levels (Davis and Ristow, 1987).

arg-2 and *cpc-1* were examined for their response to amino acid availability. Negative translational effects on *arg-2* and positive translational effects on *cpc-1* were discovered soon after cells were provided with excess arginine (Luo et al., 1995). Arginine-specific negative regulation of *arg-2* affected the levels of both *arg-2* mRNA and *arg-2* mRNA translation. This was discovered by examining mRNA levels, polypeptide pulse-labeling, and the distribution of mRNA in polysomes (Luo et al., 1995).

The *arg-2* mRNA contains a 24-codon uORF. The uORF sequence is important in negative Arg-specific translational control and its sequence is evolutionarily conserved (Figure 1.2). *arg-2-lacZ* fusion genes were used to examine

the effects of mutations that alter the *arg-2* uORF (Luo and Sachs, 1996). The uORF conferred Arg-regulated expression on *arg-2-lacZ* fusion genes when integrated into the *N. crassa* genome. Elimination of the start codon resulted in increased, constitutive gene expression and loss of translational control (Luo and Sachs, 1996). The primary sequence of the uORF was important for translational control. A UV-induced mutation at codon 12 of the uORF changing an Asp to an Asn (D12N) eliminated Arg-specific negative regulation (Freitag et al., 1996).

In vitro studies using an amino acid-dependent translation system that requires cap and poly(A) tails for maximal translation showed that the *arg-2* uORF conferred Arg-specific negative regulation on capped and polyadenylated luciferase RNAs when it was placed in the 5'-leader regions of synthetic RNAs (Wang and Sachs, 1997a). Creating a favorable uORF initiation context (↑ AUG) caused reduced luciferase production and slightly increased the magnitude of Arg-specific regulation. Mutation of Asp codon 12 to Asn (D12N) eliminates Arg-specific regulation *in vitro* as it did *in vivo*. Elimination of the uORF translation initiation codon (AUG → GUG) also eliminated Arg-specific regulation. In addition, RNA stability does not seem to be a factor in Arg-specific regulation (Wang and Sachs, 1997a). These studies also showed that this negative regulation is specific for arginine; histidine, lysine, and other closely related metabolites or analogs of arginine do not confer negative regulation on *arg-2* (Wang and Sachs, 1997a).

1.5 Thesis Overview

The goal of this study was to find out which residues of the uORF, in addition to Asp 12, are important for the regulation of *arg-2*. Six mutated uORFs were created.

1. ↑ D12N is a double mutant containing the D12N uORF in a better initiation context.
2. ΔAUG eliminates the uORF start codon (AUG → CUG).

3. S10Y changes serine 10 to a tyrosine, and corresponds to the *CPA1-O-5* mutation where a cysteine in the same relative location in the *CPA1* uORF as the serine in the *N. crassa arg-2* uORF was mutated to a tyrosine. This mutation eliminated arg regulation (Werner et al., 1987).
4. D16N is a second Asp to Asn mutation designed to determine if Arg regulation is lost as in D12N.
5. A24* is a truncated version of the uORF.
6. D12E changes Asp 12 to glutamine to see if it is the -COOH that is important for the regulation.

CHAPTER 2
MATERIALS AND METHODS

2.1 Preparation of Templates Containing Mutant *arg-2* uORF Sequences

Megaprimer PCR (Sarker and Sommer, 1990) was used to obtain mutant arg-2 uORF DNA fragments to which 5′-*Bgl*II and 3′-*Xho*I sites were added. Templates for PCR reactions were plasmids pMF11-wt and pMF11-D12N (Freitag et al., 1996), which contain the wild-type *arg-2* uORF and the D12N mutant uORF, respectively. The region amplified by PCR is shown in Figure 1.2A and the primers used in Table 2.1. Conditions for PCR are as described (Freitag et al., 1996). Predicted sizes for first PCR products were between 155-162 base pairs, and the predicted size of the second PCR product was 182 base pairs.

*Bgl*II- and *Xho*I-digested PCR products were gel-purified and ligated to *Bgl*II- and *Xho*I-digested vector pHLuc+NFS4 (Wang and Sachs, 1997a) (Figure 2.1). This plasmid contains Promega's firefly luciferase gene (minus a few restriction sites) behind a T7 promoter, a poly(A) tail, and a gene for ampicillin resistance. The sequences of the plasmid constructs were confirmed by sequencing both strands of the template plasmids.

2.2 Preparation of Synthetic RNA Transcripts

Plasmid DNA templates were purified by equilibrium centrifugation using cesium chloride gradients and linearized using *Ppu* 101. Capped, polyadenylated RNAs were synthesized using T7 RNA polymerase and the yield measured as described (Wang and Sachs, 1997a).

TABLE 2.1

Sequences of the Oligonucleotides Used for Megaprimer Mutagenesis[a]

Oligonucleotide	Length
ZW1 5′-CTGAGATCTAACTTGTCTTGTCGC-3′ contains *Bgl*II site	24 mer
ZW2 5′-CGCTCGAGCTTGAC TTGAATGGT-3′ contains *Xho*I site	23 mer
ZL19 5′-TTGTCGCAATCTGCCacaATGAACGGGCGCCC-3′ creates a good context D12N	32 mer
ZL20 5′-ATCTGCCCTTcTGAACGGGC-3′ eliminates start codon	20 mer
oJC102 5′-GTCAGTCTTCACCtaTCAGGA-3′ S10Y	21 mer
oJC104 5′-TACCTCTCAaACCATCTGTGG-3′ D15N	21 mer
oJC108 5′-GCCCTTAACtaATAAGAGCCTC A24* (stop)	22 mer
oJC103 5′-ACCTCTCAGGAaTACCTCTCA-3′ D12E	21 mer

[a] The nucleotide(s) shown in lower case are those that differ from the wild-type sequence.

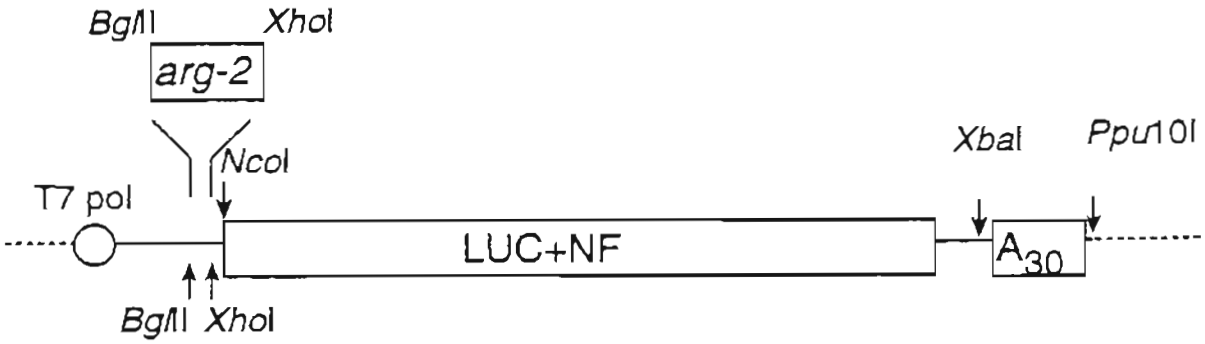


Figure 2.1. The *arg-2-LUC* construct. The arg-2 uORF was placed in front of the luciferase gene (Promega) and behind the T7 promoter. A poly(A) tail was added, and an ampicillin resistance gene was included (not shown).

2.3 Cell-Free Translation and Analyses of Translation Products

Amino acid-dependent *N. crassa* cell-free translation extracts were used and assayed for luciferase enzyme activity as described (Wang and Sachs, 1997a).

CHAPTER 3
RESULTS

A cap-, poly(A)- and amino acid-dependent *N. crassa* cell-free translation system was developed to analyze the mechanism of Arg-specific translational control (Wang and Sachs, 1997a). Arg-specific translational control via the *arg-2* uORF is reconstituted as measured by the activity of luciferase (LUC) RNA containing the uORF (Wang and Sachs, 1997a). This *in vitro* system closely resembles the *in vivo* system in the degree of Arg-specific regulation observed and the effects of uORF mutations on regulation (Wang and Sachs, 1997a).

When the Arg concentration is increased from 10 μM (low Arg) to 500 μM (high Arg) in reaction mixes containing dicistronic arg-2-LUC RNA possessing the wild-type uORF in the 5′ leader upstream of the LUC coding region (Figure 2.1), translation of luciferase is diminished, as determined by measuring luciferase enzyme activity (Figure 3.1). uORF translation is eliminated in a mutation which removes the *arg-2* uORF initiation codon (AUG → GUG). A change in codon 12 of the uORF, a critical Asp residue (D12N), eliminates the negative regulatory effect of Arg (Wang and Sachs, 1997a), but not the translation of the uORF (Freitag et al., 1996).

Analysis of the effects of the six mutations created showed that Asp 12 is not the only residue of the *arg-2* uORF critical for negative Arg-specific regulation of *arg-2*. The luciferase production of the six mutated templates are shown in Figure 3.1.

Removing the start codon (ΔAUG, AUG → CUG) eliminated Arg-specific regulation as in the AUG → GUG mutation (Wang and Sachs, 1997a), but did not increase the overall expression of LUC as the scanning model would predict, since the uORF is not being translated. However, in some organisms, CUG is recognized as a start codon, so that could explain the results. Also, although all efforts were made to

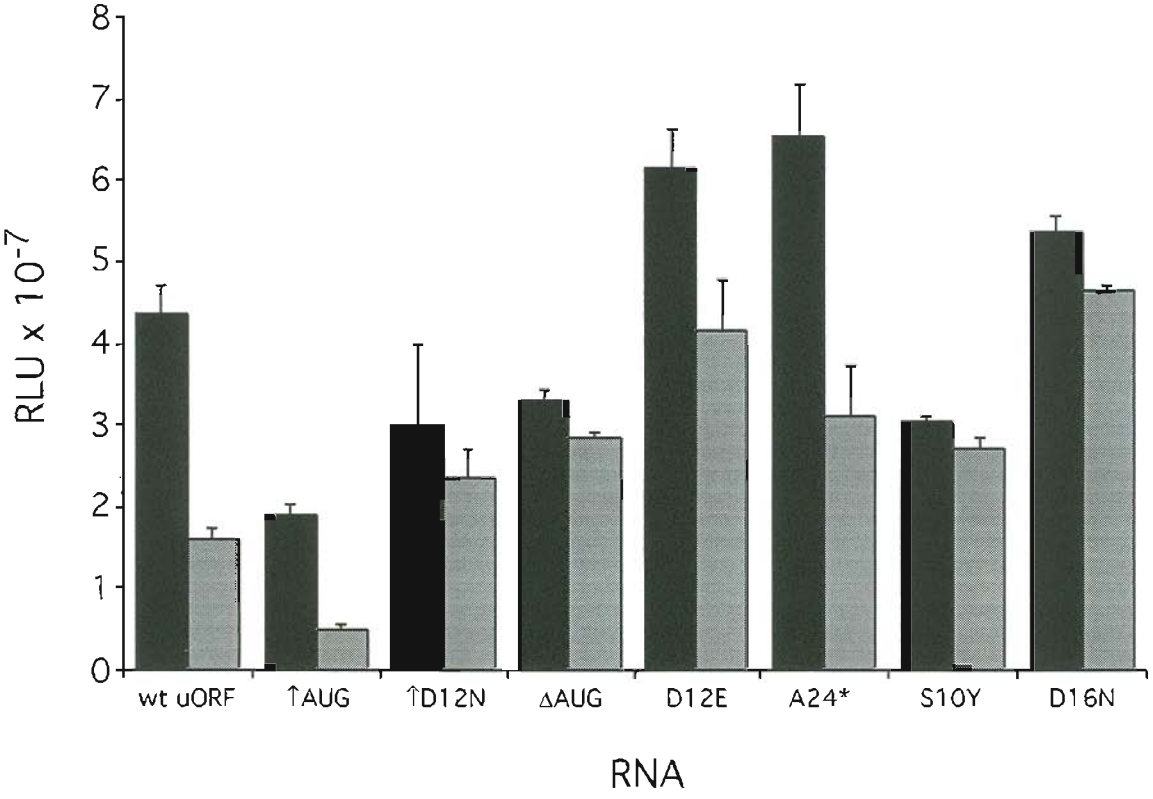


Figure 3.1. Effects of uORF mutations on Arg-specific regulation in the *N. crassa in vitro* translation system. Equal amounts of each RNA (1.2 ng) were translated in extracts containing 10 μM Arg (black) or 500 μM Arg (gray) and loaded on the mRNA downstream from an uAUG codon, as described (Wang and Sachs, 1997a). The luciferase production was measured in relative light units (RLU). Standard deviations from the mean values obtained from two independent translation reactions are indicated by error bars.

ensure equal distribution of mRNAs per sample, a slight difference from one sample to the next could occur.

The effects of changing codon 10 (S10Y) are shown (Figure 3.1, lane 7). *N. crassa* and *M. grisae arg-2* uORFs encode a Ser at codon 10 while the corresponding *CPA1* uORF encodes Cys. Mutation of *CPA1* Cys to Tyr eliminates Arg-specific regulation (Werner et al., 1987). We observed that the mutation of Ser to Tyr (S10Y) in the *N. crassa* uORF eliminates regulation.

Both *arg-2* and *S. cerevisiae CPA1* uORFs contain another conserved Asp, codon 16 in *N. crassa*. A change from Asp to Asn (D16N) eliminates Arg-specific regulation (Figure 3.1, lane 8). The D12E mutation, designed to determine if it is the -COOH that is necessary for Arg-regulation, still conferred Arg-specific regulation. However, the regulation was not as strong as in the wild-type (Figure 3.1, lane 5). Improving the initiation context of the wild-type (↑ AUG) decreases translation of luciferase and slightly increases the degree of Arg-specific regulation (Wang and Sachs, 1997a) (Figure 3.1, lane 2). In sharp contrast, improving the initiation context of the D12N template (↑ D12N) decreases translation of LUC, but does not confer Arg-specific regulation (Figure 3.1, lane 3) (Wang and Sachs, 1997a). Therefore, these Asp codons are critical for regulation.

Another mutation was made to examine the effects of shortening the uORF. A change in codon 24, an alanine to a stop codon (A24*), was made and studied in the *in vitro* system. In high Arg, Arg-specific regulation remained (Figure 3.1, lane 6). This indicates that the full-length transcript is not critical for Arg-specific regulation.

CHAPTER 4
DISCUSSION

4.1 Luciferase Assay Results

The above results indicate that at least three residues are important for the regulation of *arg-2*, Asp 12 and 16 and Ser 10. It also seems that a full length transcript is not necessary for Arg-specific regulation. Further experiments need to be done to see how short the uORF can be and still confer regulation. The fact that the good context constructs decreased overall expression of LUC indicates that leaky scanning is most likely involved in the translation of *arg-2* and its uORF.

4.2 Possible Mechanisms of Translational Regulation by Arg and the *arg-2* uORF

Other work in our laboratory supports the luciferase assay results and also indicate that ribosome stalling is occurring (Wang and Sachs, 1997b). Primer-extension inhibition, or "toeprinting," was used to test the hypothesis that Arg exercises its negative effects on translation of *arg-2* by causing the ribosomes actively translating the uORF to stall, thereby obstructing the ribosome from reinitiating at the downstream initiation codon (Lovett and Rogers, 1996). In the toeprinting technique, reverse transcriptase is used to extend a radiolabeled primer on an RNA template with or without cellular factors (Hartz et al., 1988). The toeprint should show the positions of translational components: ribosomes, ribosomal subunits, or other factors at sites where they accumulate on RNA, as they might at sites of rate-limiting steps in translation.

The results of the toeprints showed that when surplus arginine was added to translation reactions:

1. The intensity of the toeprint at a site corresponding to the uORF termination site increased.
2. Another toeprint at a location upstream of the termination site appeared. This toeprint likely represents additional ribosomes stalling behind the ones at the termination site.
3. The toeprint corresponding to ribosomes at the downstream luciferase start codon decreased.

The toeprints also support the luciferase assay results in that Arg-specific regulation is eliminated in D12N, D16N and S10Y (Wang and Sachs, 1997b).

Many uORFs reduce translation at downstream initiation codons. In *S. cerevisiae GCN4*, the primary sequence of the uORF appears to be generally unimportant. However, the primary sequences of the uORFs are critical for reducing gene expression. The action of these uORFs have been hypothesized to involve the sequence-dependent arrest of ribosomes translating them. This blockade reduces ribosomal loading at the downstream AUG codon (Geballe and Morris, 1994). This arrest of ribosomes at a eukaryotic uORF has been directly demonstrated for the uORF2 of the cytomegalovirus gp48 transcript (Cao and Geballe, 1996). For gp48, ribosomes seem to arrest at the uORF translation termination site by an unregulated mechanism.

For the *arg-2* uORF, the question is whether ribosomes reach the downstream AUG codon mainly by leaky scanning of the 40S subunit past the uORF or by translation reinitiation following uORF translation (Figure 4.1) (Wang and Sachs, 1997b). The conditions for leaky scanning, in that translation initiation can occur at a second AUG codon particularly if the initiation context for the first AUG codon is poor, hold true for *N. crassa arg-2*. Leaky scanning past the *arg-2* uORF initiation codon was consistent with the expression *in vivo* of an *arg-2-LacZ* gene in which the UORF and the *LacZ* coding region overlapped (Luo and Sachs, 1996). This is further supported by the ↑ AUG uORF and ↑ D12N data in that the translation of the uORFs

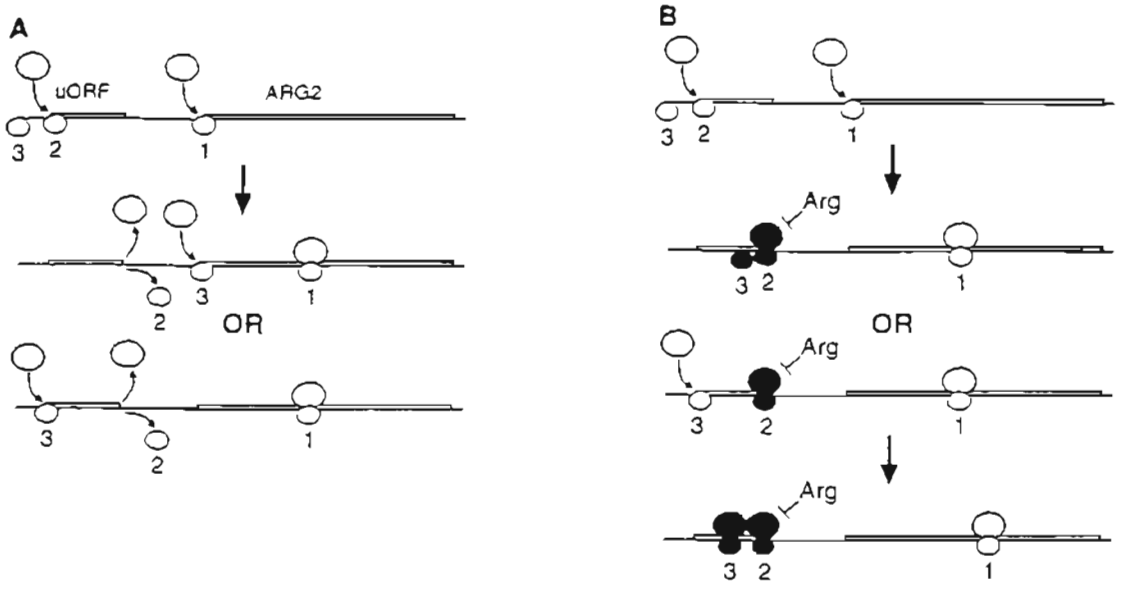


Figure 4.1. Hypothetical illustration of how ribosome stalling could mediate Arg-specific attenuation of translation from a downstream start codon. White ribosomes are not obstructed; black ribosomes are obstructed. (A) In low Arg, ribosomes do not stall at the uORF termination site. 40S ribosomal subunits (1–3) loaded from the 5′ end scan for initiation codons. 1 joins a large subunit and initiates translation at the ARG2 start codon; 2 similarly initiates translation at the uORF start codon; 3 begins scanning from the 5′-end. As time elapses (thick arrow), 1 elongates ARG2; 2 terminates uORF translation and dissociates; 3 either scans past the uORF or initiates translation at the uORF start codon. **(B)** In high Arg, ribosomes stall at the uORF termination site. 40S ribosomal subunits (1–3) are loaded from the 5′ end scan for initiation codons. 1 joins a large subunit and initiates translation at the ARG2 start codon; 2 similarly initiates translation at the uORF start codon; 3 begins scanning from the 5′-end. As time elapses, 1 elongates ARG2; 3 reaches the uORF termination codon but Arg blocks termination and/or dissociation; 3 either scans past the uORF initiation codon but its further progress is arrested by stalled ribosome 2, or it initiates translation at the uORF start codon and stalls behind 2. This model predicts that the reduced ability of ribosomes to complete uORF translation in high Arg prevents ribosomes from loading at the ARG2 initiation codon (Wang and Sachs, 1997b).

increased and the translation of the luciferase gene decreased (Wang and Sachs, 1997a,b).

A hypothetical model for Arg-specific translational control mediated by the *arg-2* uORF can be derived from the luciferase activity measurements and the toeprinting results. In this model, the movement of ribosomes past the uORF is not obstructed in low Arg, and translation at the downstream initiation codon is relatively high (Figure 4.1A) (Wang and Sachs, submitted). In high Arg, ribosomes stall while translating the uORF, hindering ribosome movement and reducing translation at the downstream initiation codon (Figure 4.1B) (Wang and Sachs, submitted).

Arginine might act directly as a regulator by interacting with the ribosomal peptidyl transferase center and inhibiting transpeptidation as it was shown to do in *E. coli* in a puromycin-based assay (Palacian and Vasquez, 1979). Since arginine can bind to polypeptides, it is possible that it could mediate regulation in this manner. Also, in *E. coli*, the arginine repressor is a DNA-binding protein that changes conformation when bound to an arginine corepressor (Maas, 1994). The crystal structure of the Arg-binding domain complexed to arginine shows that arginine binds to the protein through aspartate residues in the repressor (Van Duyne et al., 1996). This is important because aspartate residues are critical in arginine-specific regulation in *arg-2* (codons 12 and 16 of the uORF) and *CPA1* (codon 13 of the uORF) (Werner, 1987; Freitag et al., 1996; Wang and Sachs, 1997a).

Arginine may bind to a repressor that mediates control at the level of translation in combination with translation of the uORF. For example, Trp binds to an RNA binding protein that alters the structure of the trp operon leader RNA to regulate operon expression of *Bacillus* (Antson et al., 1995). No repressors have been identified that could affect the specific arginine regulation of either *arg-2* or *CPA1*, although *trans*-acting mutants that could potentially affect such regulation have been obtained (Freitag et al., 1996).

4.3 Conclusion

From the results presented here, it is clear that the primary sequence of the *arg-2* uORF is critical for this regulation. It also seems that at least three residues are important: Asp 12, Asp 16 and Ser 10. The -COOH of the Asps may be important for Arg binding. The full-length transcript does not appear to be important. It would be interesting to see how short the uORF could be and still confer regulation. Translational regulation mediated by uORFs is a subject that is not well understood, and gaining more knowledge about the control of translation initiation is probably the key to a better understanding of the effects of uAUG codons.

LITERATURE CITED

Antson, A.A., Otridge, J., Brzozowski, A.M., Dodson, E.J., Dodson, G.G., Wilson, K.S., Smith, T.M., Yang, M., Kurecki, T., and Gollnick, P. (1995) The structure of *trp* RNA-binding attenuation protein. *Nature* **374**, 693-700.

Cao, J. and Geballe, A.P. (1995) Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG codon. *J. Virol.* **69**, 1030-1036.

Cao, J. and Geballe, A. (1996) Coding sequence-dependent ribosomal arrest at termination of translation. *Mol. Cell. Biol.* **16**, 603-608.

Davis, R.H. (1986) Compartmental and regulatory mechanisms in the arginine pathways of *Neurospora crassa* and *Saccharomyces cerevisiae*. *Microbiol. Rev.* **50**, 280-313.

Davis, R. and Ristow, J. (1987) Arginine-specific carbamoyl phosphate metabolism in mitochondria of *Neurospora crassa*. *J. Biol. Chem.* **262**, 7109-7117.

Delbecq, P., Werner, M., Feller, A., Filipowski, R.K., Messenguy, F., and Pierard, A. (1994) A segment of mRNA encoding the leader peptide of the *CPA1* gene confers repression by arginine on a heterologous yeast gene transcript. *Mol. Cell. Biol.* **14**, 2378-2390.

Freitag, M., Dighde, N., and Sachs, M.S. (1996) A UV-induced mutation that affects translational regulation of the *Neurospora arg-2* gene. *Genetics* **142**, 117-127.

Geballe, A. (1996) Translational control mediated by upstream AUG codons. In *Translational Control* (Hershey, J.W.B., Mathews, M.B., and Sonenberg N., eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 173-197.

Geballe, A. and Morris, D.R. (1994) Initiation codons within 5′ leaders of mRNAs as regulators of translation. *Trends Biochem. Sci.* **19**, 159-164.

Grant, C., Miller, P., and Hinnebusch, A. (1995) Sequences 5′ of the first upstream open reading frame in *GCN4* mRNA are required for efficient translational reinitiation. *Nucleic Acids Res.* **23**, 3980-3988.

Hartz, D., McPheeter, D.S., Traut, R., and Gold, L. (1988) Extension inhibition analysis of translation initiation complexes. *Methods Enzymol.* **164**, 419–425.

Hershey, J.W.B., Duncan, R., and Mathews, M.B. (1986) Introduction: Mechanisms of translational control. In *Translational Control: Current Communications in Molecular Biology* (Mathews, M.B., ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 1–18.

Hershey, J.W.B., Mathews, M.B., and Sonenberg, N. (1996) *Translational Control*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Kozak, M. (1989) The scanning model for translation: an update. *J. Cell Biol.* **108**, 229–241.

Lovett, P.S. and Rogers, E.J. (1996) The nascent peptide as a *cis*-regulator of ribosome function. *Microbiol. Rev.* **60**, 366–385.

Luo, Z., and Sachs, M.S. (1996) Role of an upstream open reading frame in mediating arginine-specific translational control in *Neurospora crassa. J. Bacteriol.* **178**, 2172–2177.

Luo, Z., Freitag, M., and Sachs, M.S. (1995) Translational regulation in response to changes in amino acid availability in *Neurospora crassa. Mol. Cell. Biol.* **15**, 5235–5245.

Maas, W.K. (1994) The arginine repressor of *Escherichia coli. Microbiol. Rev.* **58**, 631–640.

Orbach, M., Sachs, M., and Yanofsky, C. (1990) The *Neurospora crassa arg-2* locus. *J. Biol. Chem.* **265**, 10981–10987.

Palacian, E., and Vasquez, D. (1979) Interaction of arginine with the ribosomal peptidyl transferase centre. *Eur. J. Biochem.* **101**, 113–127.

Sachs, M.S. and Yanofsky, C. (1991) Developmental expression of genes involved in conidiation and amino acid biosynthesis in *Neurospora crassa. Dev. Biol.* **148**, 117–128.

Sarker, G. and Sommer, S. (1990) The "megaprimer" method of site-directed mutagenesis. *BioTechniques* **8**, 404–407.

Shen, W.C. and Ebbole, D.J. (1997) Cross-pathway and pathway-specific control of amino acid biosynthesis in *Magnaporthe grisae. Fungal Genet. Biol.*, in press.

Van Duyne, G.D., Ghosh, W.K., Maas, W.K., and Sigler, P.B. (1996) Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J. Mol. Biol.* **256**, 377–391.

Wang, Z. and Sachs, M.S. (1997a) Arginine-specific regulation mediated by the *Neurospora crassa arg-2* upstream open reading frame in a homologous cell-free *in-vitro* translation system. *J. Biol. Chem.* **272**, 255–261.

Wang, Z. and Sachs, M.S. (1997) Arginine-specific translational attenuation. Submitted to *Mol. Cell. Biol.*

Werner, M., Feller, A., Messenguy, F., and Pierard, A. (1987) The leader peptide of yeast gene *CPA1* is essential for the translational repression of its expression. *Cell* **49**, 805–813.

BIOGRAPHICAL SKETCH

Julie Marie Carroll was born in Salem, Oregon on December 9, 1971. She attended St. Martin's College in Olympia, Washington, where she received a National Student Government Award, was named Biology Student of the Year, and was listed in Who's Who in America's Colleges and Universities. She was also a charter member of Beta Beta Beta. She graduated in 1994 with a Bachelor of Science degree in Biology. In 1995, she began the Master's program at the Oregon Graduate Institute of Science and Technology in the Department of Chemistry, Biochemistry, and Molecular Biology. She joined Dr. Matthew S. Sachs' lab, where she worked on translational control in *Neurospora*.