Sequence Requirements for Ribosome Stalling by the Arginine Attenuator Peptide*§

Christina C. Spevak†, Iivaylo P. Ivanov*§, and Matthew S. Sachs†¶

From the †Department of Neurobiology, The Scripps Research Institute, La Jolla, California 92037, the §BioSciences Institute, University College Cork, Cork, Ireland, the ¶Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112-5330, and the †Department of Biology, Texas A&M University, College Station, Texas 77843-3258

The 5′ regions of eukaryotic mRNAs often contain upstream open reading frames (uORFs). The Neurospora crassa arg-2 uORF encodes the 24-residue arginine attenuator peptide (AAP). This regulatory uORF-encoded peptide, which is evolutionarily conserved in fungal transcripts specifying an arginine biosynthetic enzyme, functions as a nascent peptide within the ribosomal tunnel and negatively regulates gene expression. The nascent AAP causes ribosomes to stall at the uORF stop codon in response to arginine, thus, blocking ribosomes from reaching the ARG-2 initiation codon. Here scanning mutagenesis with alanine and proline was performed to systematically determine which AAP residues were important for conferring regulatory function. The minimal functional domain of the AAP was determined by positioning AAP sequences internally within a large polypeptide. Pulse-chase analyses revealed that residues 9–20 of the AAP composed the minimal domain that was sufficient to confer regulatory function. An extensive analysis of predicted fungal AAPs revealed that the minimal functional domain of the N. crassa AAP corresponded closely to the region that was most highly conserved among the fungi. We also observed that the tripeptide RGD could function similarly to arginine in triggering AAP-mediated ribosome stalling. These studies provide a better understanding of the elements required for a nascent peptide and a small regulatory molecule to control translational processes.

Nascent peptides can interact with the ribosome from within the tunnel to regulate gene expression. Often these are short regulatory peptides encoded as upstream open reading frames (uORFs) in eukaryotes (1–3) or as leader peptides in prokaryotes (4, 5). The requirements for their function are beginning to become understood. The Neurospora crassa arg-2 gene specifies the small subunit of arginine-specific carbamoyl phosphate synthetase, the first enzyme in fungal arginine (Arg) biosynthesis (6, 7). The arg-2 mRNA contains a regulatory uORF encoding the 24-residue Arg attenuator peptide (AAP). The uORF-encoded AAP is highly conserved in fungi (8). It is not found outside this group. Thus, although the transcript for mammalian carbamoyl phosphate synthetase I contains a uORF, the gene has taken a different evolutionary path (9, 10), and the uORF peptide does not function similarly to the AAP (11).

In vivo studies show that AAP regulates gene expression at the level of translation. In both N. crassa and Saccharomyces cerevisiae, analyses of mRNA containing the AAP-encoding uORF show that, with respect to transcript distribution in polysomes, adding Arg to the growth medium causes a shift of these transcripts to the monosome fraction (12–14). Pulse-labeling studies in N. crassa show that the rate of synthesis of ARG-2 is reduced in Arg-supplemented medium (12).

Reconstitution of translational regulation in vitro demonstrates that the N. crassa, S. cerevisiae, Aspergillus nidulans, and Cryptococcus neoformans uORF-encoded AAPs cause the translating ribosome to stall in response to Arg with the uORF termination codon in the peptidyl transferase center (8, 11, 15, 16). AAP-regulated stalling of ribosomes in response to Arg is observed in cell-free translation systems from plant, animal, and fungal sources, indicating the mechanism does not require fungal-specific factors (17).

AAP-mediated ribosome stalling at the uORF termination codon in response to Arg blocks scanning ribosomes from reaching the downstream start codon for the biosynthetic enzyme, thus, reducing gene expression at the level of translation (18). In vivo, the increased occupancy of the uORF termination codon by the ribosome when Arg levels are high also triggers degradation of uORF-containing mRNA by the nonsense-mediated mRNA decay pathway (14). Thus, AAP-mediated ribosome stalling regulates gene expression in cis by reducing translation from a downstream start codon in the mRNA and by reducing the stability of the mRNA which served as the template for its synthesis.

Previous studies explored some of the cis-acting sequence requirements for AAP-mediated regulation. Regulation does not require either the intercistronic region (11, 19) or a specific termination codon (15). In vivo, it was demonstrated that mutation of a conserved Asp residue to Asn abolishes regulation in both N. crassa and S. cerevisiae (13, 20). Analyses of the distribution of ribosomes on reporter mRNAs in vivo show these mutations (D12N in the N. crassa AAP and D13N in the S. cerevisiae AAP) abolish regulation at the translational level (13, 14, 21). Arg-specific translational regulation by the
AAP Sequence Requirements

AAP is dependent on the amino acid sequence and not the mRNA sequence (11, 19). Studies in both N. crassa and S. cerevisiae have identified some residues in the AAP as important (11, 15, 19, 22), but the importance of each residue has not been examined systematically.

AAP-mediated ribosome stalling is also observed when the AAP is fused at the N terminus of a larger polypeptide (11, 17) or placed as an internal domain within a larger polypeptide (17). This contrasts with several other eukaryotic uORFs whose nascent peptides cause ribosomes to stall at the uORF termination codon, because these other nascent peptides do not cause stalling of ribosomes involved in elongation (2, 3, 23).

Arg-specific regulation appears to be controlled by the level of the amino acid and not the level of arginyl-tRNA aminoacylation (16). Examination of a variety of Arg analogs for their regulatory effects showed that the chiral center, guanidino group, and primary amino group of Arg are most important for regulatory function. The carboxyl group did not appear important (24). Because production of carbamoyl phosphate is a limiting step for arginine biosynthesis in fungi and the level of the small subunit of carbamoyl phosphate synthetase can control enzyme activity in the cell (6), and because the AAP has a role in regulating expression in wild-type cells growing in minimal medium (14), the regulatory role of the AAP is likely to be in control of the rate of arginine biosynthesis in response to the level of the free amino acid in the cytoplasm.

Understanding the requirements for the regulation of protein synthesis by nascent peptides that, with small regulatory molecules, control synthesis from within the ribosome is important for understanding both mechanisms that control gene expression and the functions of the ribosome. Arg-specific regulation by the AAP represents a deeply conserved regulatory mechanism for fungi that acts directly on the translating ribosome, and thus, insight into its function provides a basis for understanding these processes. Here we systematically determined which AAP residues are important for translational regulation in response to Arg through substituting AAP residues with Ala or Pro and by determining the minimal domain of the AAP required for regulatory control. Because the carboxyl group of Arg appeared unimportant for regulation, we tested the regulatory effects of adding short synthetic peptides containing Arg residues; the tripeptide RGD exerted regulation similarly to Arg.

EXPERIMENTAL PROCEDURES

DNA Templates—For scanning mutagenesis to replace AAP residues with Ala, Pro, and selected other residues, synthetic oligonucleotides (Sigma Genosys, The Woodlands, TX) were designed for use in megaprimer PCR (25) to place specific mutations affecting the AAP in PCR-generated DNA fragments that contained BglII and NcoI restriction sites near their 5’ and 3’ ends, respectively (supplemental Table 1). After digestion with these enzymes, the PCR products were ligated to a reporter vector (pR401) for further studies. The PCR template was pR301 (26). Forward primers CS1 (5’-GAATGGTGCAT-3’) or RF1 (5’-CCGCAAG-3’) were used for first-round PCR, and the various mutagenic primers named in supplemental Table 1 were used as the reverse primers for first round PCR. The product from the first round PCR (megaprimer) was used as the forward primer, and primers CS12 (5’-CTTTTCTTTAT-GTTTTTG-3’) or RF12 (5’-GTGTTTTGCCGTGGTGTAAG-3’) were used as the reverse primers. The final megaprimer PCR products were digested with BglII and NcoI and ligated to similarly digested and gel-purified vector pR401, which differed from pR301 by its containing a unique (and diagnostic) Ndel site. This vector contains the T7 RNA polymerase promoter and firefly luciferase coding region.

For analyses of the minimal functional AAP domain, AAP sequences were ligated to a vector that was derived from plasmid pLL301 (see the supplemental data in Ref. 17). This vector, pCS801, contains a T7 RNA polymerase promoter to enable in vitro synthesis of a polyadenylated transcript containing a reporter ORF consisting of (i) nine methionine codons at the N terminus, (ii) an HA tag, (iii) a luciferase domain, (iv) a modified AAP region containing strategically placed unique restriction sites to enable mutation or truncation of the AAP region, and (v) another luciferase domain (see Fig. 4). Synthetic complementary oligonucleotides were annealed and then ligated to appropriately digested pCS801 vector to produce plasmids (supplemental Table 2 and Fig. 4) containing mutated and/or truncated AAP coding segments specifying the amino acid sequences depicted in Fig. 4. All mutations were verified by DNA sequencing. Plasmid DNA templates were purified by using either the Plasmid Midi Prep kit (Qiagen) or the Wizard Plus Midi Prep kit (Promega).

RNA Synthesis, Cell-free Translation Reactions, and Primer Extension Inhibition (Toeprint) Assays—Capped, polyadenylated RNA was synthesized with T7 RNA polymerase from plasmid DNA templates that were linearized with EcoRI, and the yield of RNA was quantified (27). The reaction conditions for in vitro translation using N. crassa extracts, the preparation of extracts, and the measurements of luciferase enzyme activity were as described (11, 15, 27, 28). Translation reaction mixtures were incubated at 25 °C for 30 min, and translation was halted by freezing reaction mixtures in liquid nitrogen. For firefly luciferase activity measurements, equal amounts (12 ng) of mRNA were used to program extracts; 2.5 ng of synthetic mRNA encoding Renilla (sea pansy) luciferase was added to all reactions to serve as an internal control (15).

To assess Arg-specific ribosome stalling by monitoring the transient accumulation of stalled intermediates in polypeptide synthesis in response to Arg, the previously described [35S]Met pulse-chase protocol (17) was used with micrococal nuclease-treated N. crassa extracts or micrococal nuclease-treated wheat germ extracts (Promega). Translation reaction mixtures (20 μl) were programmed with 60 ng of RNA, and [35S]Met was used at a final concentration of 0.5 μCi/μl. Translation reactions were incubated for 2 min and then edeine was added to block additional rounds of translation initiation.

The primer extension inhibition (toeprint) assays were accomplished using 32P-labeled primer ZW4 as described (27, 29). The gels were dried and exposed to screens for ~24 h
before analysis with a Molecular Dynamics PhosphorImager. All toeprint data shown for the AAP residue substitutions were representative of two or more independent experiments with two independently prepared RNAs. The effects of RGD and GRGDS on ribosome stalling were representative of two independent experiments with independently prepared RNAs.

**Bioinformatic Analysis of AAP Conservation** — For comparative sequence assembly and analysis, 76 new AAP sequences were compiled using methodology described previously (30). These sequences and 44 known AAP sequences (8) were compared as described (30). Specific details are indicated in the legends of Fig. 6 and supplemental Fig. 2.

**RESULTS**

**Systematic Mutagenesis Revealed Key Residues in the AAP** — Synthesis of the 24-residue evolutionarily conserved AAPs caused ribosomes to stall on the mRNA in response to high Arg concentrations. The amino acid sequence of the AAP, not the nucleotide sequence specifying it, is responsible for regulation (11). To better understand the sequence requirements for AAP-mediated ribosomal stalling, we performed alanine-scanning mutagenesis to determine which residues of the AAP were important for Arg-specific regulation. Proline-scanning mutagenesis was also performed to determine whether substitution of each residue with Pro, which should constrain peptide folding, had different effects on the AAP regulatory function than did Ala.

Because AAP residues 2–5 are entirely dispensable for regulation (11), we started scanning mutagenesis on residue Ser-6 (Figs. 1 and 2). Mutations were placed into the N. crassa AAP coding region (Fig. 1, supplemental Table 1) using mega-primer PCR (25). To assess the effects on uORF constructs containing individual alanine or proline substitutions, the effects of each mutation on regulation were compared with regulation mediated by the wild-type uORF and by a uORF in which Asp-12 was mutated to Asn (D12N). The D12N mutation abolishes regulation in vivo (13) and in vitro (11, 15, 27).

Equal amounts of capped and polyadenylated synthetic RNAs representing each construct were used to program N. crassa extracts. The results of analyzing the effects of adding 10 or 500 μM Arg to translation reaction mixtures programed with three independently prepared batches of synthetic mRNA are shown in supplemental Table 3. Because the mutated AAPs might be expected to exhibit reduced function...
and not simply show loss of function, we also compared the effects of two different high Arg concentrations (500 μM or 2 mM) to a low Arg concentration (10 μM) using a fourth independent batch of RNA in a second independent batch of extract (supplemental Table 3 and Fig. 2). The results with different batches of extract and RNA were overall similar with respect to the relative effects of each mutation on regulation.

Examination of the luciferase reporter enzyme activity showed that three fundamentally different effects of mutations could be observed (Fig. 2). The first class of mutations eliminated regulation at either 500 μM Arg (Fig. 2A) or 2000 μM Arg (Fig. 2B). This was observed for Ala and Pro substitutions at residues Asp-12, Tyr-13, Leu-14, and Trp-19, which are each highly conserved in fungi (8). The second class of mutations showed reduced regulation at 500 μM Arg relative to 2 mM Arg. Thus, comparison of Fig. 2, A to B, and analysis of the data in supplemental Table 3 showed that specific mutations at residues including Ser-6, Thr-9, Ser-15, Asp-16, His-17, Leu-18, Arg-20, and Ala-21 resulted in no regulation or reduced regulation in response to 500 μM Arg, but the level of regulation increased when the Arg concentration was raised to 2 mM (i.e., residues marked in red or yellow in Fig. 2A are yellow or green, respectively, in Fig. 2B). The final class of mutations showed no obvious regulatory consequences by this assay (they looked nearly wild type at 500 μM Arg). These substitutions were primarily Ala substitutions near the N and C termini of the AAP (e.g., Val-7, Phe-8, Leu-22, and Asn-23).

The effects of Pro and Ala substitutions on regulation at a specific residue were generally similar, with Pro having a greater impact on regulation in most instances. A case where there was a large difference between Ala and Pro substitution was Ser-10, in which substitution with Pro abolished regulation independent of Arg concentration, but substitution with Ala did not. Interestingly, the highly conserved residue Asp-16 (8) could be substituted with Ala or Pro without eliminating AAP function; however, the D16N mutation eliminated function (Ref. 27 and data not shown).

To determine whether any of the mutations had significant effects on the synthesis of luciferase that were independent of their effects on Arg-specific regulation, we assessed the relative amount of luciferase produced per unit amount of input RNA (data not shown). In all cases luciferase activity was at least 80% that of the wild-type mRNA control in extracts supplemented with 10 μM Arg (low Arg). Thus, none of these mutations appeared to increase stalling in low Arg (i.e., cause constitutive stalling), a result that was borne out by the toeprint assays described next.

To verify that the differences in regulatory effects of these and additional mutations that affected regulation based on luciferase assays arose from differences in the capacity of the specific AAPs to stall ribosomes, the positions of ribosomes on each RNA were assayed by primer-extension inhibition assay (toeprinting) (Fig. 3). In these studies sequencing markers (not shown) were used to identify the positions of toeprint signals corresponding to stalled ribosomes; these signals were absent on toeprint analyses of RNA alone (−EXT, Fig. 3A, lane 5) or extract not programmed with RNA (−RNA, Fig. 3A, lane 6). Data for representative substitutions are shown in Fig. 3. The level of Arg-specific regulation, as determined by luciferase activity measurements in parallel studies, is shown in Fig. 3 above the toeprint data; there is good qualitative correspondence between the extent that Arg caused ribosome stalling at the uORF termination codon and the degree of regulation observed by luciferase assay. As expected (11), in extracts containing high Arg (500 μM) compared with low Arg (10 μM), an increased toeprint signal corresponding to ribosomes stalled at the wild-type uORF termination codon was observed (Fig. 3A, compare lanes 2 to lane 1). Also as expected, the D12N mutation did not show an increased toeprint signal at the uORF stop codon in response to high Arg (Fig. 3A, compare lanes 4 and 3). Mutation of residues Asp-12, Tyr-13, or Lys-14 to Ala or Pro abolished regulation, as seen by assay of luciferase activity and by toeprinting (Figs. 2 and 3B). The toeprint results from D16P, D16A, and D16E were also consistent with the luciferase data (Fig. 3C). Substituting the conserved Trp-19 residue with Ala or Pro abolished regulation, but changing it to Tyr did not, as established by luciferase assay and by toeprinting (Figs. 2 and 3D). An alternative aromatic residue substitution, W19F, also did not affect regulation as determined by luciferase assay. These results are consistent with the comparative analysis data across fungi showing that conserved AAPs contained an aromatic residue at this position but not always Trp (Ref. 8 and see below). Additional toeprinting data were obtained for all of the Ala and

![Figure 3](image-url)
Pro substitutions shown in Fig. 2; these data (not shown) were also consistent with the luciferase assays.

Because Ser-15, Asp-16, and His-17 are highly conserved in Pezizomycotina and Saccharomycotina (Ref. 8 and see below) and each can be individually changed to Ala with only minimal impact on function in response to 2 mM Arg (Fig. 2B), we substituted Ala for Ser-15, Asp-16, and His-17 in all possible combinations. All three residues could be converted to Ala simultaneously and regulation maintained, albeit at a reduced level. The effects of mutations were additive in that the effects of single mutations generally had less of an impact on regulation than double mutations, which had less impact than the triple mutation (data not shown).

Additional studies were performed to investigate whether altering the distance between the conserved residues Asp-12/ Tyr-13 and the conserved aromatic residue Trp-19 affected regulation. Insertion of a single Ala residue after Lys-14 or His-17 or deletion of residue Lys-14 or Leu-18 abolished regulation as determined by luciferase assays (data not shown).

**Determination of the Minimal AAP Sequence Required for Regulation**—What is the minimal functional domain of the AAP? Previous studies in which the AAP was specified as an independent uORF showed that the N-terminal segment (residues 2–5) of the AAP could be deleted without affecting regulation, but further deletions eliminated regulation (11). Shortening the C terminus significantly impacted regulation (24). It was not established from these studies whether the loss of regulation occurred because the length of the nascent peptide was crucial or specific residues were crucial. Because the AAP (residues 2–24) could be placed internally within a larger polypeptide and the AAP function to stall ribosomes was retained using ribosomes from fungal, plant, and animal sources (17), we adapted this approach to determine the shortest contiguous region of the AAP that could retain stalling function.

We examined the minimal sequences required for the AAP using constructs to systematically remove the N- and/or C-terminal residues of the AAP when it was positioned internally in a larger polypeptide. The effects of removing these residues were assessed by [35S]Met pulse-chase experiments in which a stall in translation is detected by the appearance of a polypeptide intermediate in SDS-PAGE analyses (17). Synthetic mRNAs were derived from plasmid pLL301 (17) and a polypeptide intermediate in SDS-PAGE analyses (17). Synthesis of systematic deletions from each terminus to produce a construct was created that combined the largest functional domains had compensated or influenced the deleted residues are 

The constructs used to analyze the minimal sequence required for an internal AAP to function. Deletions from the N- and C-terminal regions are compensated by the LUC sequences on both sides (in lowercase). The LUC residues that are at the same positions as the original AAP residues are capitalized. Luc residues bordering AAP-residues in pCS914 were changed to Ala-residues in pCS914 as indicated. Pulse-chase time course experiments and SDS/PAGE were performed to examine whether the modified internal AAP continued to cause ribosome stalling. Green, Arg-specific stalling observed; red, Arg-specific stalling not observed.

stabilized when the concentration of Arg is high. The results of testing constructs containing various N- and C-terminal systematic deletions of the AAP and examining the radiolabeled translation products using the pulse-chase protocol established the minimal sequence required for Arg-specific regulation are summarized in Fig. 4. Systematic deletion from the N terminus showed that deletion of residues 2–8 (ΔNGRPSVF) did not affect regulation as seen by the presence of the intermediate (I) when Arg concentrations were high, similar to the situation of the full-length AAP (Fig. 5, compare panel A to panel B). Removal of residues 2–9 affected regulation (ΔNGRPSVFT) as no visible intermediate (I) product in the presence of high Arg concentrations was observed (Fig. 5, compare panel C with B and A). Similarly, systematic deletions from the C terminus showed that residues 21–24 (ΔALNA) could be deleted without eliminating regulation (Fig. 5D), whereas deletion of codons 20–24 (ΔRALNA) considerably diminished regulation (Fig. 5E).

After determining the essential residues for Arg-specific regulation by systematically deleting the AAP C and N termini, a construct was created that combined the largest functional deletions from each terminus to produce a construct encoding only 12 of the AAP residues (Thr-9—Arg-20). To eliminate the possibility that residues from the adjacent luciferase domains had compensated or influenced the deleted AAP residues, alanine was substituted for all amino acid positions of the AAP that were shown to be dispensable for function (pCS914, Fig. 4). This transcript was used to program N. crassa and wheat germ extracts, and polypeptide chain elongation was monitored by pulse-chase analysis (Fig. 5F and supplemental Fig. 1). An intermediate (I) corresponding to
polypeptide synthesis stalled at the internal AAP was observed to be enhanced in the presence of high Arg in both systems. Thus, pulse-chase analysis showed that the 12 core AAP residues (Thr-9—Arg-20) were sufficient to cause ribosome stalling in response to Arg.

The Ala replacement construct, pCS914, shares two residues within the full-length AAP outside the core (residues Ala-21 and Ala-24). The construct pCS913 (Fig. 4), in which the 12 core AAP residues were bounded by adjacent LUC residues, which eliminated any similar residues with the wild-type AAP, also stalled ribosomes in response to Arg (data not shown), providing evidence that these Ala residues were not crucial for stalling.

The Minimal Functional Domain of the AAP Is Functionally Important in Vivo Based on Evolutionary Conservation—We investigated how well residues identified as important for AAP-mediated regulation in the in vitro experiments correlated with their phylogenetic conservation by conducting an exhaustive comparative sequence analysis. For this purpose a total of 120 AAP sequences were compiled (Fig. 6 and supplemental Fig. 2) including 44 previously described sequences (8). AAP sequences were identified in six fungal phyla: Ascomycota, Basidiomycota, Zygomycota, Blastocladiomycota, Chytridiomycota, and Neocallimastigomycota. A phylogenetic tree of relatedness of these fungi based solely on AAP sequences was highly consistent with current understanding of the evolution of fungi at the level of phyla and subphyla (31, 32). Only two residues, Asp-12 and Tyr-13, are absolutely conserved in all AAPs; the remainder showed varying degrees of conservation. The strict conservation of Asp-12 is consistent with data from the current and previous studies showing its importance for all aspects of AAP-mediated regulation. Similarly, the in vitro results presented here for Tyr-13 show that mutating it to Ala or Pro completely abolished regulation by Arg. The only other residues showing conservation across all taxonomic groups with identifiable AAPs are Lys-14 and Trp-19. The position corresponding to N. crassa Lys-14 is Leu or Ile in all but one of the compiled sequences (in which...
AAP Sequence Requirements

it is Phe). The position equivalent to \textit{N. crassa} Trp-19 is Trp in 100 of the 120 AAPs examined and is either Phe or Tyr in the others; thus, it is always an aromatic residue. These conservation data are also consistent with the \textit{in vitro} data reported here. Five other positions corresponding to \textit{N. crassa} Ser-10, Ser-15, Asp-16, Leu-18 and Arg-20 are highly conserved in that they contain similar if not identical residues in at least 4 of the 6 fungal phyla for which AAP sequences were identified. Alanine and proline scanning indicated that mutations near the N termini and C termini, but although the N-terminal region generally did not. The initiation codon of the AAP in this yeast, at 18 amino acids long, is also the shortest of them all. Based on evolutionary data, it may not have full function because the other sequenced species of the genus \textit{Pichia}, for example \textit{Pichia stipitis} and \textit{Pichia guilliermondii}, lack an AAP-encoding uORF 5’ of the coding sequence of their genes specifying the small subunit of Arg-specific carbamoyl phosphate synthetase. \textit{Trichophyton tonsurans} presents another curious deviation. In this species the AAP is fused directly in-frame with the main open reading frame. Fusions of the AAP are known to be functional \textit{in vitro} (11, 17) and \textit{in vivo} (19).

Arginine Analogs and Arg-containing Short Peptides Provide Insights into the Structural Requirements for Regulatory Function—Because previous studies showed that the level of aminoacylated Arg-tRNA is not important for regulation (16), Arg or a metabolite of Arg must be important. Many Arg analogs have been previously tested for their effects on regulation on the wild-type AAP, such as arginine methyl ester, which conferred regulation, and precursors such as citrulline that did not confer regulation (24). In a pilot study we tested in parallel the Arg-related compounds D-arginine, arginine methyl ester, citrulline, agmatine, argininamide, and L-NG-monomethyl arginine with each of the Ala- and Pro-scanning mutants from Fig. 2 as well as with the wild-type and D12N AAP controls. The results indicated that Arg-methyl ester was the only analog that showed regulation similar to Arg when an equal amount of each RNA were translated in the reaction mixtures supplemented with 500 \textmu M concentrations of each of the Arg analogs (data not shown). This indicated that changing the primary amino acid sequence of the AAP did not alter the responses to these molecules that were structurally related to Arg, and these studies were not pursued further. The results from these and previous studies suggest that the chiral center, guanidino group, and primary amino group of arginine are most important for regulatory function as changing these groups abolished regulation.

Changes at the carboxyl group of Arg appeared to have little effect on regulatory function. This led us to investigate the effects of short peptides containing an Arg residue such as RGD, RGDS, GRGD, and GRGDS (purified peptides from Bachem and Calbiochem). The prediction was that only peptides containing an N-terminal Arg would confer regulation.

RGD and RGDS showed similar effects as Arg in causing Arg-specific negative regulation through the wild-type arg-2 AAP but not the D12N AAP in \textit{N. crassa} extracts as determined by luciferase assays (data not shown). Examination of the regulatory response at the molecular level by mapping the positions of ribosomes on the transcript by the toeprint assay
showed that in the wild-type AAP construct (Fig. 7A), increasing the concentration of Arg from 10 to 100–500 μM in *N. crassa* cell-free translation reactions resulted in an increased toeprint at the uORF termination codon, which is diagnostic for Arg-specific regulation (Fig. 7A, compare lane 1 to lanes 2–6). Adding the tripeptide RGD at concentrations of 100–500 μM had an effect similar to Arg (Fig. 7A, compare lanes 2–6 to lanes 7–11) on ribosome stalling at the uORF termination codon. However, the addition of 100–500 μM GRGDS to translation reactions containing the wild-type AAP did not result in an enhanced toeprint signal at the termination codon (Fig. 7A, compare lane 1 to lanes 12–16). GRG tripeptide also did not confer regulation.3 In the D12N AAP construct (Fig. 7B), similarly increasing concentrations of Arg, RGD, and GRGDS did not result in an increased toeprint signal at the uORF termination codon (Fig. 7B), consistent with the inability of the D12N AAP to confer regulation.

Was the effect of RGD a consequence of hydrolysis of the tripeptide to release Arg during the course of the reaction to Arg? We added the aminopeptidase inhibitors bestatin and amastatin (33, 34) to reaction mixtures; these additions did not change the regulatory results. Furthermore, analyses of the kinetics of regulation and the effects of adding different concentrations of Arg or RGD to *N. crassa* extracts showed that both Arg and RGD acted at similar rates on the wild-type AAP-LUC fusion. For example, we performed pulse-chase experiments and detected a similar stalling response to the addition of either 300 μM Arg or 300 μM RGD soon after initiating translation. We programmed extracts with mRNA specifying a fusion protein (Fig. 8A) containing two wild-type AAP domains (pL31, Ref. 17). We used 300 μM instead of 2 mM concentrations to demonstrate the similar sensitivity of stalling to Arg and RGD at relatively low concentrations. The results (Fig. 8B) show that, in contrast to translation reactions containing 10 μM Arg, adding 300 μM RGD causes the ap-

---

3 J. Wei and M. Sachs, unpublished information.
pearance of transient species representing translational stalling at the N-terminal AAP (N) and the internal AAP (I) with kinetics similar to those as a consequence of adding 300 μM Arg. These data indicate that RGD can replace Arg in vitro to stall ribosomes.

**DISCUSSION**

We systematically tested the function of residues within the *N. crassa* AAP for their roles in conferring Arg-specific translational regulation by scanning mutagenesis. The role of individual residues was assessed when the AAP was positioned as a regulatory uORF. Regulatory capacity was determined by analysis of the AAP capacity to modulate production of firefly luciferase (Fig. 2 and supplemental Table 3). Regulation by the AAP was verified to be mediated by ribosome stalling at the uORF termination codon by primer extension inhibition (to-eprint) assays (Fig. 3). In complementary studies, the minimal contiguous functional domain of the AAP was determined by positioning AAP sequences internally within a larger polypeptide. The effects of progressive substitutions of N-terminal and C-terminal AAP residues were examined by pulse-chase analyses of polypeptide synthesis to determine whether removal of the residues caused ribosome stalling. We found that the residues 9–20 of the *N. crassa* AAP comprised the minimal domain sufficient to confer regulatory function. These studies represent the most complete analyses to date to elucidate the roles of specific residues in a eukaryotic nascent peptide to elicit ribosome stalling in response to a small molecule.

Finally, we determined that the tripeptide RGD could function similarly to the amino acid arginine to elicit AAP-mediated ribosome stalling in vitro. These studies indicate that the site at which Arg effects ribosome stalling when the AAP is in the ribosome tunnel remains accessible despite bulky additions to this regulatory amino acid’s carboxyl group.

Scanning mutagenesis showed that many of the most highly evolutionarily-conserved residues in this fungal uORF-encoded peptide (Fig. 6 and Ref. 8) cannot be changed to Ala or Pro without completely eliminating regulatory function. Thus, Asp-12, Tyr-13, Leu-14, and Trp-19 cannot tolerate these substitutions, indicating the importance of their side chains for function. However, conservative substitutions at some of these sites (D12E, W19Y, W19F) showed reduced regulation (Figs. 2 and 3), confirming the importance of the specificity of the residue (e.g., an aromatic residue at the position of Trp-19 was essential).

Although AAP function was not entirely wild-type with substitutions with Ala or Pro at many other positions within the AAP, neither was function eliminated. Surprisingly, the AAP retained at least partial function when the highly conserved Asp-16 was substituted with either Ala or Pro (but not Asn). This indicates that Asp-16 does not provide an essential functional site, for example, through necessary side-chain interactions with Arg and/or the ribosomal tunnel. Overall, the residues that were most tolerant of substitution were generally outside of the minimal domain (Thr-9—Arg-20) that by itself conferred function; these N- and C-terminal regions that are tolerant to substitutions are also generally less conserved among AAPS from different species (Fig. 6 and Ref. 8).

What similarities and differences are observed with the AAP and other nascent peptides that stall ribosomes? The bacterial TnaC and SecM stalling peptides and the cytomegalovirus uORF2 gp48 stalling peptide represent a group of nascent peptides that require a Pro codon to be in the peptidyltransferase center for stalling to occur (35–38). In prokaryotes, Pro in the ribosomal P-site can slow peptidyltransferase activity even in the absence of stalling sequences in the nascent peptide (39, 40). Changing the last codon of the AAP to Pro (A24P) did not increase regulation but reduced it; thus, the mode of action by which the AAP stalls ribosomes is not aided by Pro at the peptidyltransferase center.

The evolutionary conservation of the length of the AAP and the data from primer extension inhibition studies with the AAP localized at internal regions within larger coding regions strongly indicate that there is a relatively narrow window of opportunity within the ribosome for the nascent AAP to exert regulatory function (8, 17). Analyses of other nascent peptides that exert regulatory control by stalling also show such a narrow window (4, 5). The simplest interpretation of these data is that the regulatory segment within the AAP must interact with the translation machinery at specific sites in that machinery to exert control. In this regard it is interesting to consider that only at AAP Ser-10 did substitution with Pro have a complete knock-out effect on function, whereas substitution with Ala had little effect; at most other sites, Pro substitution simply had a more substantial, but qualitatively similar, effect on regulation. This region of the AAP, which bears similarities to TnaC and SecM, would be predicted to interact with the ribosome near eukaryotic ribosomal protein L17 (the homolog of bacterial L22), based on studies of the stalling peptides from bacteria.

The results of testing the function of the AAP as an internal domain show that 12 residues of the AAP (Thr-9—Arg-20) had the capacity to cause ribosome stalling within two heterologous contexts; that is, surrounded by luciferase residues or surrounded by multiple Ala residues. This is the most highly conserved region of the AAP and was the region that was overall the least tolerant of substitution. With one exception, all single-residue substitutions that abrogated regulation were within this region. Outside of this region, only the adjacent A21P mutation could abrogate regulation. These complementary experiments (scanning mutagenesis of the intact AAP and defining the functional core of the AAP by its capacity to confer stalling in a new context) are, thus, in good agreement and identify the key functional elements of the AAP.

The conservation of these key functional elements of the AAP among many different evolutionary branches of the fungi, as determined from analyses of predicted AAP sequences, indicate the importance of the AAP to this group of organisms. The conservation of the AAP in six fungal phyla indicates that the AAP has been maintained for approximately a billion years (41).

More generally, that a relatively short stretch of residues in a nascent peptide such as the AAP can be demonstrated to transiently stall translation elongation in response to a small molecule raises the possibility that related translational con-
trol mechanisms employing short coding sequences could be widely used for regulatory purposes.

Acknowledgments—We thank Dr. Peng Fang for assistance and Randolph Addison, Cheng Wu, and Jiajie Wei for helpful discussions and comments on the manuscript.

REFERENCES

1. Sachs, M. S., and Geballe, A. P. (2006) Genes Dev. 20, 915–921
2. Morris, D. R., and Geballe, A. P. (2000) Mol. Cell. Biol. 20, 8635–8642
3. Hood, M. M., Neafsey, D. E., Galagan, J., and Sachs, M. S. (2009) Annu. Rev. Microbiol. 63, 385–409
4. Ramu, H., Mankin, A., and Vazquez-Laslop, N. (2009) Mol. Microbiol. 71, 811–824
5. Ito, K., Chiba, S., and Pogliano, K. (2010) Biochem. Biophys. Res. Commun. 393, 1–5
6. Davis, R. H. (1986) Microbiol. Rev. 50, 280–313
7. Radford, A. (2004) Adv. Genet. 52, 165–207
8. Hood, H. M., Spevak, C. C., and Sachs, M. S. (2007) Fungal Genet. Biol. 44, 93–104
9. Hong, J., Salo, W. L., Lusty, C. J., and Anderson, P. M. (1994) J. Mol. Biol. 243, 131–140
10. van den Hoff, M. J., Jonker, A., Beintema, J. J., and Lamers, W. H. (1995) J. Mol. Evol. 41, 813–832
11. Fang, P., Wang, Z., and Sachs, M. S. (2000) J. Biol. Chem. 275, 26710–26719
12. Luo, Z., Freitag, M., and Sachs, M. S. (1995) Mol. Cell. Biol. 15, 5235–5245
13. Freitag, M., Dighde, N., and Sachs, M. S. (1996) Genetics 142, 117–127
14. Gaba, A., Jacobson, A., and Sachs, M. S. (2005) Mol. Cell 20, 449–460
15. Wang, Z., Fang, P., and Sachs, M. S. (1998) Mol. Cell. Biol. 18, 7528–7536
16. Wang, Z., Gaba, A., and Sachs, M. S. (1999) J. Biol. Chem. 274, 37565–37574
17. Fang, P., Spevak, C. C., Wu, C., and Sachs, M. S. (2004) Proc. Natl. Acad. Sci. U.S.A. 101, 4059–4064
18. Gaba, A., Wang, Z., Krishnamoorthy, T., Hinnebusch, A. G., and Sachs, M. S. (2001) EMBO J. 20, 6453–6463
19. Delbecq, P., Werner, M., Feller, A., Filipkowski, R. K., Messenguy, F., and Piérad, A. (1994) Mol. Cell. Biol. 14, 2378–2390
20. Werner, M., Feller, A., Messenguy, F., and Piérad, A. (1987) Cell 49, 805–813
21. Luo, Z., and Sachs, M. S. (1996) J. Bacteriol. 178, 2172–2177
22. Delbecq, P., Calvo, O., Filipkowski, R. K., Piérad, A., and Messenguy, F. (2000) Curr. Genet. 38, 105–112
23. Geballe, A. P., and Sachs, M. S. (2000) in Translational Control of Gene Expression (Sonenberg, N., Hershey, J. W. B., and Mathews, M. B., eds) pp. 595–614, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
24. Wang, Z., and Sachs, M. S. (1997) J. Biol. Chem. 272, 255–261
25. Sarkar, G., and Sommer, S. S. (1990) Biotechniques 8, 404–407
26. Fang, P., Wu, C., and Sachs, M. S. (2002) Fungal Genet. Biol. 36, 167–175
27. Wang, Z., and Sachs, M. S. (1997) Mol. Cell. Biol. 17, 4904–4913
28. Spevak, C. C., Park, E. H., Geballe, A. P., Pelletier, J., and Sachs, M. S. (2006) Biochem. Biophys. Res. Commun. 350, 834–841
29. Sachs, M. S., Wang, Z., Gaba, A., Fang, P., Belk, J., Ganesan, R., Amrani, N., and Jacobson, A. (2002) Methods 26, 105–114
30. Ivanov, I. P., Loughran, G., and Atkins, J. F. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 10079–10084
31. James, T. Y., Kauff, F., Schoch, C. L., Matheny, P. B., Hofstetter, V., Cox, C. J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H. T., Rauhut, A., Reeb, V., Arnold, A. E., Amtuff, A., Stajich, J. E., Hosaka, K., Sung, G. H., Johnson, D., O’Rourke, B., Crockett, M., Binder, M., Curtis, J. M., Slot, J. C., Wang, Z., Wilson, A. W., Schüssler, A., Longcore, J. E., O’Donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P. M., Powell, M. J., Taylor, J. W., White, M. M., Griffith, G. W., Davies, D. R., Humber, R. A., Morton, J. B., Sugiyama, J., Rossman, A. Y., Rogers, J. D., Pfister, D. H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R. A., Kohlmeier, J., Volkmann-Kohlmeier, B., Spotts, R. A., Serdani, M., Crous, P. W., Hughes, K. W., Matsurua, K., Langer, E., Langer, G., Unterreiner, W. A., Lücking, R., Büdel, B., Geiser, D. M., Aptom, A., Diedrich, P., Schmitt, I., Schultz, M., Yah, R., Hibbett, D. S., Lutzoni, F., McLaughlin, D. J., Spatafora, J. W., and Vilgalys, R. (2006) Nature 443, 818–822
32. McLaughlin, D. J., Hibbett, D. S., Lutzoni, F., Spatafora, J. W., and Vilgalys, R. (2009) Trends Microbiol. 17, 488–497
33. Taylor, A. (1993) FASEB J. 7, 290–298
34. Yasothornsrikul, S., Toneff, T., Hwang, S. R., and Hook, V. Y. (1998) J. Neurochem. 70, 153–163
35. Cao, J., and Geballe, A. P. (1996) Mol. Cell. Biol. 16, 7109–7114
36. Cruz-Vera, L., and Yanofsky, C. (2008) J. Bacteriol. 190, 4791–4797
37. Garza-Sánchez, F., Janssen, B. D., and Hayes, C. S. (2006) J. Bacteriol. 281, 34258–34268
38. Muto, H., Nakatogawa, H., and Ito, K. (2006) Mol. Cell Biol. 22, 545–552
39. Muto, H., and Ito, K. (2008) Biochem. Biophys. Res. Commun. 366, 1043–1047
40. Wohlgemuth, L., Brenner, S., Beringer, M., and Rodnina, M. V. (2008) J. Biol. Chem. 283, 32229–32235
41. Berbee, M. L., and Taylor, J. W. (2010) Fungal Biol. Rev. 24, 1–16