Multiple Forms of Arginase Differentially Expressed from a Single Locus in *Neurospora crassa*

(Received for publication, June 17, 1998, and in revised form, August 18, 1998)

Sudhir Marathe‡‡, Yeon Gyu Yu‡‡, Gloria E. Turner*, Christine Palmier, and Richard L. Weiss

From the Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1569

The *Neurospora crassa* catabolic enzyme, arginase (L-arginine amidinohydrolase, EC 3.5.3.1), exists in multiple forms. Multiple forms of arginase are found in many vertebrates, but this is the only reported example in a microbial organism. The two major forms are structurally similar with subunit sizes of 36 and 41 kDa, respectively. The larger form is produced by mycelia growing in arginine-supplemented medium. Both forms are localized in the cytosol. The structural gene for arginase, *aga*, has been cloned and sequenced; it contains a 358-codon open reading frame with three in-frame ATGs at the amino terminus. Mutagenesis of these ATGs revealed that the first ATG initiates the 41-kDa protein and the third ATG initiates the 36-kDa protein. Mutation of the second ATG has no effect on translation. Northern analysis demonstrated that a 1.4-kilobase (kb) transcript is synthesized in minimal medium and both a 1.4- and 1.7-kb transcript are produced in arginine-supplemented medium. Primer extension identified the 5′ ends of each transcript and demonstrated that the first and third ATG of the open reading frame are the initial AUGs of the 1.7- and 1.4-kb mRNA, respectively. The results suggest that a basal promoter produces the 1.4-kb transcript and an arginine “activated” promoter is responsible for the 1.7-kb transcript. Tandem promoters are rare in eukaryotic organisms, and they often regulate developmental or tissue-specific gene expression. The possibility that arginase has a role in differentiation in *N. crassa* is being investigated.

The existence of multiple forms of arginase in eukaryotes suggests a complex regulatory role for this enzyme in the metabolism, development, and maintenance of these organisms. The mammalian liver arginase is well characterized (1–5); however, little is known about the role of extrahepatic arginases found in erythrocytes, mammary gland, kidney, brain, salivary gland, gastrointestinal tract, and lens tissue. The liver protein is immunologically distinct from extrahepatic arginases with the exception of the rat submaxillary gland arginase (3, 6). The hepatic enzyme functions as an essential component of the urea cycle, catalyzing the hydrolysis of arginine to ornithine and urea. The existence of a urea cycle is thought to keep ammonia levels low, preventing toxicity in ureotelic organisms by excreting excess nitrogen as urea, a neutral water-soluble molecule. It has been argued that the urea cycle maintains pH homeostasis by the coordinate removal of bicarbonate and ammonium ions (7, 8).

In *Neurospora crassa* and other microbial organisms that do not have a urea cycle, the products of arginase, ornithine, and urea, allow utilization of arginine as a nitrogen source. Ornithine can serve as a precursor for proline and glutamate as well as polyamine synthesis, whereas urea can provide NH₃ via urease (9). It has been hypothesized that the extrahepatic mammalian arginases could provide precursors for proline, glutamate, and polyamines. Recent experiments with the newly cloned human extrahepatic arginase (AII) suggest a role in down-regulation of nitric oxide synthesis (10). The cloning of three nonhepatic arginases in *Xenopus laevis*, and the demonstration of their differential expression during metamorphosis, is further evidence for a non-urea cycle role for arginase (11). However, the precise function of these extrahepatic arginases has not yet been defined. The fact that several prokaryotic organisms have arginase implies that the hepatic enzyme arose from a protein with a function or functions different from the mammalian urea cycle.

When arginase was purified from *N. crassa*, it migrated as a 36-kDa peptide during sodium dodecyl sulfate-polyacrylamide gel electrophoresis (PAGE). Western blot analysis with polyclonal antibodies raised against the 36-kDa peptide recognized two major immunoreactive forms, 41 and 36 kDa, in a wild-type strain (12, 13). The larger form was produced by mycelia growing in arginine-supplemented medium. Peptide mapping of the two *N. crassa* arginase proteins yielded similar peptide maps, suggesting that the two species have similar amino acid sequences. Since both forms were absent in an *aga* mutant strain which is unable to utilize arginine as the sole nitrogen source, it was suggested that this locus may either regulate or encode both species of arginase (12).

Arginase is found throughout the primary kingdoms of life. It has been postulated that all arginases are derived from a universal common ancestor before the divergence into archaea, eubacteria, and eukarya (14). Therefore, understanding the metabolic roles of two major arginases from *N. crassa* could provide insight into this evolutionary branch point where the primordial arginase evolved into an essential component of the urea cycle.

We describe the cloning and characterization of the structural gene for arginase and the relationship between the two

---

*This work was supported by United States Public Health Service Grant GM47631 and National Science Foundation Grant DCR91-19151 (to R. L. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡‡ Present address: Korea Institute of Science and Technology, P. O. Box 131, Cheongryang, Seoul, Korea.

* Present address: Dept. of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095-1569. Tel.: 310-206-0946; Fax: 310-206-4038; E-mail: turner@chem.ucla.edu.

1 The abbreviations used are: PAGE, polyacrylamide gel electrophoresis; kb, kilobase(s); VM, Vogel’s minimal medium; RFLP, restriction fragment length polymorphism; CAPS, 3-(cyclohexylamino)propanesulfonic acid; bp, base pair(s); PCR, polymerase chain reaction; ORF, open reading frame.
forms of arginase and the *aga* locus. We show that the *aga* locus produces two transcripts, 1.4- and 1.7-kb, from tandem promoters which are translated into two proteins: 36 and 41 kDa, respectively. Expression is regulated by arginine through mechanisms not yet understood. Both forms of arginase appear to be localized exclusively in the cytosol. Interestingly, the hepatic and nonhepatic arginases in mammals are encoded by separate genes, whereas the two *N. crassa* arginases are differentially expressed from a single locus. Possible roles for multiple forms of arginase in *N. crassa* are discussed.

**EXPERIMENTAL PROCEDURES**

*N. crassa* and *Escherichia coli* Strains and Growth Conditions—A standard *N. crassa* wild-type, 74-OR23-1A (LA1), and arginaseless mutants *aga*, allele UM913 (LA185), and *aga* his-3, alleles UM913 and 1–234-1438 (LA187), were used for all experiments unless otherwise noted. *N. crassa* strains FGSC 4411–4430 were used for RFLP mapping and were obtained from the Fungal Genetics Stock Center, University of Kansas Medical Center, Kansas City. The *E. coli* strains DH5α and TG1 were used for plasmid amplification, and RZ1032 and CJ236 were used for preparing uracil-containing plasmid DNA. When replicating plasmids in *E. coli*, strains were grown in Luria-Bertani medium supplemented with 100 μg/ml ampicillin and 200 μg/ml formic acid containing 0.5 M NaOAc (pH 5.2), 1 mM EDTA, 7% SDS at 65 °C for 2–4 h. Hybridization with the random-primed 32P-labeled 2.4-kb *Apal-Per* fragment from pAAP (Fig. 1B) was carried out overnight (24). Membranes from Southern analyses were washed twice at 65 °C with 0.1 × SSC, 0.5% SDS for 15 min and twice at 65 °C with 0.2 × SSC, 0.5% SDS for 15 min before autoradiography. Membranes from Northern analyses were washed twice for 10 min at room temperature in 2 × SSC, 0.5% SDS and once for 10 min at room temperature and 3 times for 5–10 min at 65 °C with 0.1 × SSC, 0.5% SDS.

**DNase I Digestion Fragment Analysis**—Polymerase chain reaction (PCR) mapping—DNase I digestion fragment mapping was performed as described by Metzenberg et al. (25). Genomic DNAs from Oak Ridge and Mau-riceville parental strains (FGSC 4411 and FGSC 4416) were digested with various restriction enzymes and subjected to Southern analysis using a 1.5-kb *SalI-Per* fragment as a probe (Fig. 1B). This fragment contains most of the coding region and an additional 0.35-kb of the 3′-untranslated sequences. A polymorphism was detected for *HindIII* restriction fragment lengths. Southern analysis was then performed using HindIII-digested genomic DNA from the parental strains and 18 of their progeny.

**DNA Sequence Analysis**—Restriction fragments within pPAG (Fig. 1B) were subcloned and their sequences were determined manually by the dideoxy method (26) using Sequenase Version 2.0 (U. S. Biochemical Corp.). The plasmid, pAAP, which contains the *aga* coding region and 500 bases of the 5′ upstream region, was sequenced on both strands (Fig. 2). Nucleotide sequence analysis and protein sequence comparisons were performed with the aid of version 7 of the Genetics Computer Group programs (Madison, WI) on a VAX computer. Data base searches were performed using the BLAST network services at the National Center for Biotechnology Information (27). The 2388-bp nucleotide sequence of the *Apal-Per* fragment containing the arginase gene has been listed in GenBank/EMBL data bases under accession number L20687.

**Arginase of Neurospora**—Arginase was purified from the wild-type strain by a modification of a previously described procedure (12). The modifications were limited to the type of resins used in column chromatography: (a) salt gradient chromatography was performed using a Q-Sepharose column (Pharmacia Biotech Inc.); (b) chromatofocusing was performed using a PBE 94 (Pharmacia) column; and (c) hydroxyapatite column chromatography was performed on a 0.75 10 cm hydroxyapatite (Bio-Rad) column.

**Chemical Cleavage of Arginase for Microsequencing**—Arginase protein was cleaved with cyanogen bromide (CNBr), and the resulting peptides were fractionated and prepared for microsequencing (17). Approximately 100 μg of arginase protein was incubated in 200 μl of 70% formic acid containing 200 μg of CNBr at room temperature under nitrogen for 1 h away from any light source. After 24 h, the reaction was diluted 10-fold with water and lyophilized. The resulting peptides were fractionated by SDS-PAGE as described by Laemmli (18) and electrotransferred to polyvinylidene difluoride membranes in CAPS buffer (10 mM CAPS, 10% methanol, pH 11) at 0.15 A for 4 h. The amino-terminal sequences from 5- and 12-kDa peptide fragments were determined at the UCLA Protein Sequencing Facility.

**Polymerase Chain Reaction and DNA Clones**—Degenerate primers representing the amino acid sequences obtained from the CNBr-cleaved arginase peptides were synthesized (see Fig. 1A). Polymerase chain reactions performed on *N. crassa* genomic DNA as a template (strain LA1), consisted of 3 cycles (92 °C for 60 s, 50 °C for 60 s, 72 °C for 30 s) followed by 27 cycles (92 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s). The 261-bp PCR product was cloned and identified as a portion of the *arg* gene based on analysis of the nucleotide sequence. The PCR product was used as a probe to identify a 2.4-kb *Apal-Per* DNA fragment from a genomic DNA cosmid library pMOcosX (19). All subcloning was carried out by methods described by Maniatis et al. (15).

**Isolation of *N. crassa* Genomic DNA and Total RNA** *N. crassa* genomic DNA was isolated as described previously (20). Total RNA was extracted from approximately 1 g (wet weight) of mycelia as follows: mycelia, 4 g of glass beads, 3.5 ml of extraction buffer (50 mM NaCl, 5 mM EDTA, 1% SDS, 50 mM Tris, pH 7.5) and 3.5 ml of phenol/chloroform/isoamyl alcohol (25:24:1) were vortexed for 3–4 min in a 30 ml Corex tube and centrifuged at 5000 × g for 10 min; the aqueous phase was extracted 3 times in this manner until the interface after centrifugation was clear of any debris. Nucleic acids were precipitated with 0.3 M NaOAc (pH 5.2) and 2 volumes of ethanol at 0 °C. Poly(A)*RNA was isolated using a previously reported protocol (21).

**Southern and Northern Blotting**—DNAs were fractionated by gel electrophoresis in a 0.8% agarose gel and transferred to a nylon membrane (MSI Magna Graph) by capillary transfer using 10 × SSC (15). Total RNA was denatured with glyoxal, separated on a 1% agarose gel (22), and capillary transferred to a nylon membrane using 20 × SSC.

The membranes were UV cross-linked (23) and baked in a vacuum oven at 80 °C for 2 h. The baked membranes were prehybridized in 0.5 M NaHPO4 (pH 7.2), 1 mM EDTA, 7% SDS at 65 °C for 2–4 h. Hybridization with the random-primed 32P-labeled 2.4-kb *Apal-Per* fragment from pAAP (Fig. 1B) was carried out overnight (24). Membranes from Southern analyses were washed twice at 65 °C with 0.1 × SSC, 0.5% SDS for 15 min and twice at 65 °C with 0.2 × SSC, 0.5% SDS for 15 min before autoradiography. Membranes from Northern analyses were washed twice for 10 min at room temperature in 2 × SSC, 0.5% SDS and once for 10 min at room temperature and 3 times for 5–10 min at 65 °C with 0.1 × SSC, 0.5% SDS.

**Mutations that Selectively Inhibit Arginase Activity**—Mutations that selectively modify the amino-terminal amino acid sequences of the arginase gene (see Fig. 2) were introduced as described by Kunkel et al. (31) and Carter (32). The template for mutagenesis was the 880-bp insert of pAAP (Fig. 1B), which was cloned into pBluescript KS I (Stratagene) or M133BM20 RF (Boehringer Mannheim). Mutagenesis with PBS II KS clones was carried out according to Kunkel et al. (31) but the M133BM20 RF mutagenesis was
modified as follows: 70 mM Tris-Cl (pH 7.5) was used in the kinase reaction buffer: polyethylene glycol 8000/ammonium acetate was substituted for polyethylene glycol 8000/NaCl; 200 mM Tris-Cl (pH 7.5), 2 mM MgCl₂, 50 mM NaCl was used for annealing, and the extension buffer was altered (17.5 mM Tris-Cl, pH 7.5, 3.75 mM MgCl₂, 21.5 mM dithiothreitol, 0.75 mM ATP, 0.4 mM dNTPs). The nucleotide sequences of the oligonucleotides used to alter the first, second, and third ATG codons were: OMA1, 5'-GAGGGGAAATTTGACAGAAAGA-3'; OMA2, 5'-CCTGTTTAGTTGTTGATTTGCT-3'; and OMA3, 5'-TTGGGAGAGCTGGCAAGGC-3'.

Mutant DNA Constructs—Constructs were made with the 880-bp fragment from pAAPM1 with the altered M1, M2, M3, or M2/M3 which targeted the altered DNA to the his-3 locus or allowed for random genomic integration. The random integration constructs were done as follows: ApaI linearized pAAP (Fig. 1B) was subjected to SauI partial digestion. The 4.4-kb fragment containing the wild-type arginase gene lacking the 880-bp ApaI-SalI fragment; the latter was replaced with the mutated 880-bp ApaI-SalI DNAs yielding pAAPM1, pAAPM2, and pAAPM2/M3. Constructs for targeting to the his-3 locus were made using M13 clones. The wild-type 600-bp BamHI-SalI fragment was removed from pAAP (Fig. 1B) using a partial double digest with BamHI and SauI and replaced with the BamHI-SalI fragment containing the M1* mutation. The M3 mutation was cloned by ApaI-SalI replacement. The resulting constructs, pAAPM1* and pAAPM3, were cloned into a his-3 targeting vector pRAUW122 (33) as follows: the 2.1-kb BamHI-PstI fragment in pAAP (Fig. 1B) was replaced with the BamHI-PstI fragment from pAAPM1*; the resulting plasmid pAAPM3* was cloned into the EcoRI-XbaI sites of pRAUW122 resulting in pAAPM3*122. The M3 mutation was cloned by cloning the ApaI-PstI fragment from pAAPM3 into the ApaI-PstI sites of pRAUW122.

Transformation—Conidial protoplasts from N. crassa aga mutant strain LA185 were co-transformed with 1 μg of DNA from each mutant construct made for random integration and 0.1 μg of the E. coli hygromycin phosphotransferase gene (hph) from plasmid pMP6. Transformation was done as described by Vollmer and Yanofsky (34). Selection for the hph marker was made on plates containing 150 μg/ml hygromycin in the bottom agar. Conidial protoplasts from mutant strain LA787, aga his-3, were transformed using 1.0 μg of pAAPM1*122 DNA; this construct carries the distal portion of the wild-type his-3 gene which is targeted to the his-3 locus. Recombination at the his-3 locus restores prototrophy. Transformation of LA787 with pAAPM3122 used the same mode of selection, but was performed by electroporation. Fresh conidia with 2.0 μg of DNA according to Vann (35) with modifications by Ivey et al. (36). Hygromycin-resistant transformants and positive prototrophs were assayed for arginase activity. All positives were purified to homokaryons by three cycles of single conidial isolation. Transformants Tm1, Tm2, Tm3, and Tm2/Tm3 were used for characterization experiments and will be described under “Results.”

Subcellular Fractionation—Extracts were fractionated from LA1 grown in VMA and VA medium at 30 °C for 20 h. Crude nuclei were isolated according to Hautala et al. (37) with modifications as reported by Orbach et al. (20). The crude nuclear pellet was further purified by the procedure of Grove and Marzluf (38). Mitochondrial and vacuolar fractionation were by the method of Cramer et al. (39) with modifications reported in Zerez et al. (40).

RESULTS

Cloning the Structural Gene for Arginase—Several strategies were used to clone the structural gene for N. crassa arginase. Hybridization with the heterologous probe, CAR1 from S. cerevisiae (41), or with degenerate oligonucleotide primers designed from known arginase sequences were unsuccessful. Amino-terminal sequencing of purified arginase from strain LA1 was not possible because the NH₂ terminus is blocked. Alteration of the purification protocol to delete the heat step did not affect blocking of the NH₂ terminus. Therefore, the purified protein was used to generate peptide fragments for determining internal amino acid sequences; arginase was cleaved with cyanogen bromide (CNBr), and amino acid sequences were determined from two of the peptide fragments. The sequences showed strong similarities with two separate domains of arginase sequences from rats liver, human liver (42, 43), and S. cerevisiae (41). This oriented peptide fragments within the arginase protein.

Degenerate oligonucleotide primers, which incorporated the codon bias observed for N. crassa genes, were synthesized based on the peptide sequences. A schematic representation of the nucleotide sequence and the position of the primers is shown in Fig. 1A. A PCR using N. crassa genomic DNA as a template, amplified a unique 261-bp DNA fragment. The nucleotide sequence of the DNA fragment revealed a single open reading frame. The putative amino acid sequence was identical to the arginase peptide sequences, demonstrating that the DNA fragment was derived from the arginase structural gene. The PCR product was then used to probe a N. crassa genomic DNA library, pMOcosX (19). The probe hybridized to cosmid G18:H4 (plate G18, column H, number 4). Restriction and nucleotide sequence analysis of portions of the cosmid revealed that it contained the coding region for arginase but little of the upstream sequence. Therefore, an 1.8-kb EcoRI-PstI DNA fragment encoding arginase from cosmid G18:H4 was used to identify a second cosmid, X13:G10, from the same genomic cosmid DNA library. Southern analysis of cosmid X13:G10 identified a 5.5-kb PstI fragment which contained the desired gene. The sequence of this fragment showed that it contained the complete coding region for arginase and approximately 3.3 kb of the 5’ upstream region. This fragment was subcloned into pBluescript II KS to generate plasmid pAPA1 (Fig. 1B). A 2.4-kb ApaI-PstI fragment was further subcloned from pAPA1 to generate pAAPM. This fragment was sequenced on both strands (Fig. 2).

RFLP Mapping of the Arginase Gene—N. crassa chromosomes have been marked on the basis of the segregation patterns of RFLPs for cloned genomic DNA sequences. These patterns have been recorded for a set of 18 progeny derived from a cross of Oak Ridge and Mauriceville parental strains (25). The 261-bp PCR product was used as a probe in mapping the arginase gene. DNAs from two parent strains were digested with different restriction enzymes. The hybridization patterns for several of Oak Ridge restriction digestes were different from Mauriceville digestes identifying arginase gene-specific RFLPs. Since the polymorphism was most pronounced in HindIII digests, genomic DNA from the two parental strains and 18 progeny from a mating of the two strains was digested with HindIII and subjected to Southern analysis. The segregation pattern of the RFLP for the cloned gene matched exactly the pattern of the un-10 locus (see Table I) located on the right arm of linkage group VII. The aga locus maps to the right arm of linkage group VII, close to the un-10 locus (2 centimorgans from wc-1, 24 centimorgans from arg-10). This places the arginase gene at or near the aga locus on chromosome VII.

Arginase Is Encoded by the aga Locus—Previous immunological studies revealed that expression of both major immunoreactive species of arginase is affected by aga mutations; in strain LA185, both species are absent (12, 44). These observations were suggestive of a relationship between arginase and the aga locus. Since the aga locus and the arginase gene map to the same location on chromosome VII, it seemed plausible that they are the same. To confirm this conclusion, strain LA185 was transformed with various regions of pAGA and transformants were tested for their ability to utilize arginine as the sole nitrogen source; the 2.4-kb ApaI-PstI fragment from pAAP (see Fig. 1B) was sufficient to complement the aga mutation. This
observation, combined with the results from RFLP mapping, indicates that the aga locus is the structural gene for arginase.

**Nucleotide Sequence of the Arginase Gene**—The nucleotide sequence of the 2,389-bp ApaI-PstI fragment of pAGA, determined for both DNA strands, is shown in Fig. 2. It has a 358-codon open reading frame (ORF) with three in-frame ATGs at positions 111, 1170, and 11100 bp marked M 1, M 2, and M 3, respectively. The ORF is interrupted by two segments of 72 and 73 base pairs which have the characteristics of *N. crassa* introns. These introns were confirmed by cloning and sequencing a cDNA from a Zap library. The codon usage shows a bias which is typical of highly expressed *N. crassa* genes: pyrimidines are strongly preferred to purines in third codon positions, cytosine is preferred over thymine, and guanine is strongly preferred over adenine.

The predicted amino acid sequence exhibits high identity to known arginase protein sequences. The highest identity, 72%, is with arginase from the human fungal pathogen *Coccidioides immitis*; in addition, comparison of the two introns of *N. crassa* with the two found in *C. immitis* indicates that they are located in identical positions within the respective genes and are very close in size (45). A PILEUP analysis shown in Fig. 3 illustrates a comparison of arginase protein sequences from vertebrates to microorganisms. There are several regions of 6–10 amino acids which are almost fully conserved among the arginases, indicative of common structural features. One striking feature of the *N. crassa* arginase is that the first 33 amino acids fail to show any homology with arginase from any other organism.

**Arginase Gene Expression**—Transformation and RFLP mapping indicated that aga is the structural gene for arginase. To determine the origin of the two protein species, we analyzed arginase gene expression at the transcriptional and translational levels. A wild-type strain of *N. crassa* (LA1) was grown in VM, VMA, and VA (see “Experimental Procedures”). Mycelia were harvested after 6 h of growth and subjected to arginase activity assays and Northern and Western analyses. Arginase activity increased approximately 2-fold in VMA and approximately 4-fold in VA medium (data not shown). Western analysis detected a single 36-kDa protein in VM-mycelia (Fig. 4A, lane 1). An additional 41-kDa protein was detected in VMA and VA-mycelia (Fig. 4A, lanes 2 and 3). The level of the 36-kDa protein was elevated in VMA-mycelia, and the level of both proteins was highest in VA-mycelia. Densitometry revealed that the amount of the 36-kDa protein was 10 times greater than the 41-kDa protein in VM (0.29:03), 1.5 times in VMA (0.79:0.52), and twice as great in VA (2.38:1.14).
A probe consisting of the 2.4-kb insert from pAAP detected a single 1.4-kb transcript in VM-mycelia (Fig. 4B, lane 1). An additional 1.7-kb transcript was detected in VMA and VA-mycelia (Fig. 4B, lanes 2 and 3). The level of the 1.4-kb mRNA was elevated in VMA-mycelia, but the level of both transcripts was highest in VA-mycelia. These observations indicate that transcription of both mRNAs is induced by arginine. The Northern and Western analyses were consistent with a one-to-one correlation between the 1.4-kb transcript and 36-kDa protein and the 1.7-kb transcript and 41-kDa protein and support the view that the 36- and 41-kDa peptides are translated from the 1.4- and 1.7-kb transcripts, respectively.

### Mapping the 5' End of Arginase mRNAs

To determine the transcriptional start sites of the two mRNA molecules, Poly(A) RNA extracted from VM and VMA-grown cultures was used in primer extension experiments. A pair of transcription start sites were detected at nucleotide positions 187 and 188 relative to the first in-frame methionine using RNA from both VMA and VM-grown cultures (Fig. 5A, lanes 1 and 2). The segregation pattern of the arginase gene-specific RFLP for a HindIII digest is shown. The RFLP pattern is marked for the parental strains, Oak Ridge (FGSC number 4411) and Mauriceville (FGSC number 4416), and 18 progeny. The segregation pattern matches with that for the un-10 locus which is located on the right arm of chromosome VII.

### Table I

| Strain | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|--------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| un-10  | O  | M  | M  | M  | O  | (M) | O  | M  | O  | M  | O  | O  | O  | O  | O  | O  | O  | O  | O  |
| Arginase| O  | M  | M  | M  | O  | (M) | O  | M  | O  | M  | O  | O  | O  | O  | O  | O  | O  | O  | O  |

---

**FIG. 2. The 2389-bp nucleotide sequence of pAAP.** The 358-codon open reading frame is interrupted by 72 bp (Intron-I) and 73 bp (Intron-II). The arrows at positions +236, +87, and +88 mark the transcription start sites which are in **boldface**. The three in-frame ATG codons marked in **bold** as M1, M2, and M3 are at positions +1, +70, and +100, respectively. Putative cis regulatory elements are underlined in **bold**, the GAGA sites have double underlines with the polyadenylation signal being underlined. GenBank/EMBL accession number is L20687.

The segregation pattern of the arginase gene-specific RFLP for a HindIII digest is shown. The RFLP pattern is marked for the parental strains, Oak Ridge (FGSC number 4411) and Mauriceville (FGSC number 4416), and 18 progeny. The segregation pattern matches with that for the un-10 locus which is located on the right arm of chromosome VII.

---

A probe consisting of the 2.4-kb insert from pAAP detected a single 1.4-kb transcript in VM-mycelia (Fig. 4B, lane 1). An additional 1.7-kb transcript was detected in VMA and VA-mycelia (Fig. 4B, lanes 2 and 3). The level of the 1.4-kb mRNA was elevated in VM-mycelia, but the level of both transcripts was highest in VA-mycelia. These observations indicate that transcription of both mRNAs is induced by arginine. The Northern and Western analyses were consistent with a one-to-one correlation between the 1.4-kb transcript and 36-kDa protein and the 1.7-kb transcript and 41-kDa protein and support the view that the 36- and 41-kDa peptides are translated from the 1.4- and 1.7-kb transcripts, respectively.

**Mapping the 5' End of Arginase mRNAs**—Northern analysis demonstrated differential synthesis of two arginase transcripts. To determine the transcriptional start sites of the two mRNA molecules, Poly(A) RNA extracted from VM and VA-mycelia was used in primer extension experiments. A pair of transcription start sites were detected at nucleotide positions +87 and +88 relative to the first in-frame methionine using RNA from both VM and VM-grown cultures (Fig. 5A, lanes 1.
Fig. 3. Comparison of the arginase protein sequences from human liver, Hs AI (43), human extrahepatic, Hs AII (10), rat liver, Rn (42), Xenopus liver, Xl (60), Xenopus extrahepatic, Xl (11), N. crassa, Nc (this study, Fig. 2), C. immitis, Ci (45), S. cerevisiae, Sc (41), Schizosaccharomyces pombe, Sp (61), and the Agrobacterium Ti plasmid, Ab (62). The comparison was assembled by PILEUP, a GCG package program. Macboxshade depicts all identical amino acids in white on black, while similar amino acids are black on gray. Nonmatches are black on white.
and 2). A second major transcription start site was detected 324 bp upstream at position -236 with RNA from the VMA-mycelia (Fig. 5B, lane 1), but not with RNA from VM-mycelia (Fig. 5B, lane 2). The inability to detect a start site at -236 with RNA from VM-mycelia argues that arginine is required for expression of the 1.7-kb transcript. The 0.3-kb difference in size of the two observed mRNAs is comparable to the 324 nucleotides separating the two transcription start sites. These observations mark position -236 as the start site for the 1.7-kb mRNA and positions -187 and -188 as start sites for the 1.4-kb transcript (see Fig. 2). The 1.7-kb mRNA initiates upstream of the three in-frame ATGs, whereas the 1.4-kb mRNA initiates between the second and third ATG codons.

Site-directed Mutagenesis—These observations are consistent with a model in which the 36-kDa protein is produced from the first ATG codon of the 1.4-kb transcript (M3 of the arginase open reading frame) while the 41-kDa protein is initiated at the first ATG codon in the 1.7-kb transcript (M1 of the arginase open reading frame). To determine the initiator ATGs for the two proteins, a series of arginase gene constructs was generated carrying mutations at one or more of the first three in-frame ATG codons. Mutations at M1 and M2 changed the ATG codons to ATA (Met to Ile), and the mutation at M3 changed the ATG codon to CTG (Met to Leu). A fourth construct, which altered both M2 and M3, contained two additional changes upstream of M1 at nucleotides -41 and -46.

The altered ATG constructs were cloned into either pBS II KS or the his-3 targeting vector pRAUW122 as described under “Experimental Procedures” and used to transform the aga mutant strain, LA185, or the aga his-3 strain, LA787. Since the outcome of producing only one form of arginase could not be predicted, strain LA185 was co-transformed with the E. coli hygromycin phosphotransferase gene (hph) and selection was for hygromycin resistance. Transformation of LA787 selects for recombination at the his-3 locus which is detected by restoration of prototrophy. The recipient strains, LA185 and LA787, have no detectable arginase protein by Western analysis (Fig. 6, lanes 7 and 8) and an almost undetectable transcript (Fig. 7 lane 2), in addition there is no arginase enzymatic activity (data not shown). The purified transformants were checked for the aga- phenotype by testing their growth in liquid medium containing arginine and assaying for arginase activity. Arginase-deficient strains cannot grow in arginine-supplemented medium due to a polyamine deficiency (44). Transformants are identified by the nature of the alteration: Tm1 has the altered
his-3 grown on VMH or Vogel's minimal medium (VM) or Vogel's minimal medium supplemented with 1 mM arginine (VMA). The **aga his-3** strains required 1 mM histidine (VMH) and 100 μg/ml putrescine when grown with 1 mM arginine (VMHAp). Tm3 grown on VM (lane 1), Tm3 grown on VMA (lane 2), Tm2 grown on VM (lane 3), Tm2 grown on VMA (lane 4), Tm1 grown on VMA (lane 5), Tm1 grown on VM (lane 6), **aga his-3** grown on VMH (lane 7), **aga his-3** grown on VMHAp (lane 8), wild-type grown on VMA (lane 9), and wild-type grown on VM (lane 10). The numbers on the right represent the sizes of the proteins in kilodaltons. Protein amount loaded varied between 32 μg for wild-type, Tm1 and Tm3 to 42 μg for Tm2 and **aga his-3**.

M1, Tm2 has altered M2, Tm3 has mutated M3, and Tm2/Tm3 has altered M2 and M3 as well as nucleotide substitutions at -46 and -41. Tm1 and Tm3 are **his-3** targeted transformants. To characterize the effect of the mutations at the transcriptional and translational level, the transformants were grown in VM, VMA, and VA liquid medium and subjected to Northern and Western analysis. Tm1, Tm3, and Tm2/Tm3 strains failed to grow in VA medium. These transformants produce only one form of arginase which has catalytic activity in vitro. The inability to grow on arginine as a sole nitrogen source with one form in vivo raises a number of interesting questions regarding the regulation and interaction of the two forms.

Western analysis determined that Tm3, which has the third in-frame ATG mutated, produced only the 41-kDa protein when grown in VM; no arginase protein was seen in VM (Fig. 6, lanes 1 and 2). This indicates that arginase expression is regulated normally when integrated at the **his-3** locus (compare lanes 1 and 2 versus 9 and 10). The Tm1 transformant produced only the 36-kDa protein (Fig. 6, lanes 5 and 6). The transformant with mutated M2 synthesized the 36-kDa form in VM and both proteins in VMA (Fig. 6, lanes 3 and 4). Two species were absent in the **aga his-3** strain (Fig. 6, lanes 7 and 8). These observations demonstrate that ATGs at +1 (M1) and +100 (M3) serve as translation initiation codons for the 41- and 36-kDa proteins, respectively.

RNAs isolated from mycelia grown in VM and VMA for 6 h were probed with the 2.4-kb insert of pAAP. The 1.4-kb transcript was detected in VM RNA for wild-type, Tm1, Tm2, and Tm3, and both transcripts were detected in these strains grown in VMA (data not shown).

Characterization of Tm2/Tm3 provided some suggestive information regarding the basal promoter. Western analysis revealed that this transformant makes only the 41-kDa protein (data not shown) which is consistent with the Tm3 result. Northern analysis revealed a surprising result: RNA isolated from mycelia grown in VM produced only the 1.7-kb transcript (Fig. 7, lane 5); this is in contrast to the Tm2 and Tm3 strains that make both transcripts in VMA (Fig. 7, lane 4, and data not shown). This result suggests that nucleotides at -41 and -46, altered in Tm2/Tm3, are necessary for transcription of the 1.4-kb message.

**Enzyme Localization**—Extracts from wild-type mycelia grown in both VMA and VA were fractionated to mitochondrial, vacuolar, nuclear, and cytoplasmic fractions. Western blot analysis detected arginase protein in the cytosolic and crude nuclei fractions. However, purified nuclei did not contain any cross-reactive material, suggesting that our crude nuclear fraction was contaminated with cytoplasmic proteins (data not shown). The failure to detect either form of arginase in organelles is consistent with the lack of arginine enzymatic activity in the mitochondria and vacuole (63).

**DISCUSSION**

This is the first report of multiple forms of arginase in a microbial organism. The larger protein is produced from a larger mRNA transcript in response to exogenous arginine, and translation begins at the first methionine in the open reading frame. The smaller protein is translated from a shorter mRNA in which the first ATG codon encodes the third methionine in the arginase open reading frame. Both species of arginase are localized in the cytosol of mycelia grown in submerged culture. In *N. crassa* cyclophilin A-binding protein, cyclophilin, is made as a 24- and a 20-kDa protein and both forms are encoded by a single gene. This is also the case for the **cyt-20** gene which encodes both cytosolic and mitochondrial valyl-tRNA synthetases. Lack of direct protein work with the **cyt-20** gene product precludes comparison with arginase and cyclophilin (54). The two forms of cyclophilin were shown to be in separate cellular compartments the cytosol and the mitochondria (46).

Several nuclear genes in *S. cerevisiae* encode two protein species: **LEU4** (47), **FUM1** (48), **VAS1** (49), **TRM1** (50), and **HTS1** (51, 52). In these examples, the two forms of the protein are associated with different locations, cytosol and mitochondria. The genes have ORFs with at least two in-frame ATG codons near their 5' ends and produce two transcripts: a shorter transcript which includes only the downstream ATG and produces the cytoplasmic form, and a longer transcript, originating upstream of the ORF, producing a larger protein which is targeted to the mitochondria. In the case of **LEU4** the mitochondrial form is targeted but not processed. The *S. cerevisiae* **SUC2** gene encodes two forms of invertase, an intracellular and a secreted form. The longer secreted invertase is glycosylated and regulated by glucose, the shorter form is not modified and its transcript is constitutive (53). Many of these proteins are translated from separate ATGs but are processed into similar proteins once they reach their subcellular location. The *N. crassa* arginase proteins do not appear to be altered, both forms are found in the same cellular compartment and they both respond to the inducer arginine.

The biological significance of the synthesis of two species of arginase in *N. crassa* is not yet clear. Both forms have enzymatic activity. Since the 41-kDa species is produced specifically in response to exogenous arginine, it may have a role in arginine utilization. The 33-amino acid sequence at the NH2 ter-
minus of the 41-kDa protein, 18 of which are hydrophobic, does not show any homology to arginase from other organisms. One possibility is that the extended NH₂ terminus of the 41-kDa protein targets it to a subcellular organelle during differentiation. We have only examined vegetative growth conditions; perhaps early in development or during nitrogen limitation the arginine-induced form will be found in an organelle. Another possibility is that the 41-kDa protein interacts with the 36-kDa protein in order to achieve changes in enzymatic activity. This is possible given the result that both forms increase with elevated amounts of arginine. Alternatively, the two proteins may differ in catalytic properties in ways appropriate for the conditions leading to their production.

Preliminary studies suggest that product inhibition kinetics differ for the two forms. The most compelling argument for cooperativity of the 36- and 41-kDa proteins is the observation that mutant strains producing a single form fail to grow on arginine as the sole nitrogen source. Arginase may require a combination of the two subunits for maximal activity, with many possible combinations. One example would be a hexamer forming a dimer of trimers. Alternatively, maximal activity may only occur when there is interaction between the two homogeneous hexamers. We cannot rule out the possibility that the transformant strains may lack necessary regulatory elements that are required for maximal arginase expression when arginine is the sole nitrogen source. The constructs include ~800 nucleotides upstream of the first ATG, but important enhancer sequences are sometimes found several kilobases upstream of the initiator ATG in eukaryotic genes.

Northern analysis and primer extension suggest that expression of arginase is regulated at the level of transcription. Three unique sets of repeats occur in the promoter region of the 1.4-kb transcript (Fig. 2). The most proximal set consists of three TACC repeats. TACC repeats are found in the promoter of nit-3, the structural gene for nitrate reductase (55), as well as a . N. crassa. gene homologous to the bacterial two-component histidine kinase that is thought to interact with a putative nitrogen response regulator. It is noteworthy that expression of all three genes is affected by nitrogen availability via mechanisms that are not well understood. Interestingly, the ~46 nucleotide change in Tm2/Tm3 occurs in the first set of TACC repeats and appears to prevent transcription of the 1.4-kb transcript. The mutated sequence becomes TTCC, with an ad-

**DNA sequence within the ORF, but preceding the transcription start site of the 1.4-kb mRNA, contains both a TATA box and a CCAAT box (Fig. 2), both of these cis-elements are necessary for the transcription of many eukaryotic genes (56, 57). Several GAGA sequences are found upstream of the transcription start site for the 1.7-kb transcript as well as one in the downstream promoter which is responsible for the 1.4-kb mRNA. GAGA factor-binding sites have been identified by DNase I footprinting in the proximal promoters of 10 Droso-

**REFERENCES**

1. Beruter, J., Colombe, J. P., and Bachmann, C. (1978) Biochem. J. 175, 449–454
2. Hirsch-Kolb, H., and Greenberg, D. M. (1968) J. Biol. Chem. 243, 6123–6129
3. Johnson, C. P., Grody, W. W., and Cederbaum, S. D. (1996) Comp. Biochem. Physiol. 114, 107–132
4. Tarrab, R., Rodriguez, J., Huitron, C., Palacios, R., and Soberon, G. (1974) Eur. J. Biochem. 48, 458–468
5. Viale-Breithard, F., and Orth, G. (1972) J. Biol. Chem. 247, 1227–1235
6. Glass, R. D., and Knox, W. E. (1973) J. Biol. Chem. 248, 5758–5789
7. Atkinerson, G., and Bourke, E. (1984) Trends Biochem. Sci. 9, 297–302
8. Atkinerson, C. E. (1990) Physiology. Ed. 65, 243–267
9. Davis, R. H. (1986) Microbiol. Rec. 50, 280–313
10. Gotth, T., Sonoki, T., Nasasaki, A., Terada, K., Takiguchi, M., and Mori, M. (1986) FEBS Lett. 195, 119–122
11. Patterson, D., and Shi, Y. J. (1994) J. Biol. Chem. 269, 25328–25334
12. Borkovich, K. A., and Weiss, R. L. (1987) J. Biol. Chem. 262, 7573–7586
13. Borkovich, K. A., and Weiss, R. L. (1997) J. Bacteriol. 169, 5510–5517
14. Ouzounis, C. A., and Kyprides, N. C. (1994) J. Mol. Biol. 309, 101–104
15. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
16. Vogel, H. J. (1964) Am. Nat. 98, 435–452
17. LeGendre, N., and Matsudaira, P. (1989) A Practical Guide to Protein and Peptide Purification for Microsequencing, Academic Press, New York
18. Laemmli, U. K. (1970) Nature 227, 689–698
19. Orbach, M. J., Sachs, M. S., and Yanofsky, C. (1990) J. Biol. Chem. 265, 10981–10987
20. Orbach, M. J., Puro, E. B., and Yanofsky, C. (1986) Mol. Cell. Biol. 6, 2452–2461
21. Sachs, M. S., David, M., Werner, S., and Rajbhandary, U. L. (1986) J. Biol. Chem. 261, 869–873
22. Thomas, S. P. (1985) Proc. Natl. Acad. Sci. U.S.A. 77, 5201–5205
23. Church, G. M., and Gilbert, W. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 1991–1995
24. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 126, 6–13
25. Metzenberg, R. L., Stevens, J. N., Selker, E. U., and Morzycka-Wroblewska, E. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 2067–2071
26. Scherrer, P., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
27. Altschul, S. F., Gish, W., Miller, W., and Lipman, D. J. (1995) J. Mol. Biol. 215, 403–410
28. Smith, P. M. (1976) Anal. Biochem. 72, 248–254
29. Davis, R. H., and Mora, J. (1968) J. Bacteriol. 95, 248–254
30. Archibald, R. M. (1945) J. Biol. Chem. 157, 507–518
31. Reusch, T. A., Roberts, J. D., and Zalkow, R. A. (1987) Methods Enzymol. 154, 367–382
32. Carter, P. (1987) Methods Enzymol. 154, 382–403
33. Aramayo, R., and Metzenberg, R. L. (1989) Fungal Genet. Newsl. 34, 9–13
34. Wahlner, S. J., and Yanofsky, C. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 4869–4873
35. Vann, D. C. (1995) Fungal Genet. Newsl. 42A, 53
36. Ivey, F. D., Hodge, P. N., Turner, G. E., and Borkovich, K. A. (1996) Mol. Biol. Cell. 7, 1283–1297
37. Houtala J. A., Conner, B. H., Jacobson, J. W., Patel, G. L., and Giles, N. H. (1977) J. Bacteriol. 130, 704–713
38. Grove, G., and Marzluf, G. (1981) J. Biol. Chem. 256, 463–470
39. Cramer, C. L., Ristow, J. L., Paulus, T. J., and Davis, R. H. (1983) Anal. Biochem. 129, 384–392
40. Zerze, C. R., Weiss, R. L., Franklin, C., and Bowman, B. J. (1986) J. Biol. Chem. 261, 8877–8882
41. Sumrada, R. A., and Cooper, T. G. (1986) J. Bacteriol. 160, 1078–1087
42. Kawamoto, S., Amaya, Y., Murakami, K., Tokunaga, F., Iwanaga, S.,...
