THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 266, No. 9, Issue of March 25, pp. 5854-5859, 1991
Printed in U.S.A.

The *Neurospora crassa* Carotenoid Biosynthetic Gene (Albino 3) Reveals Highly Conserved Regions among Prenyltransferases*

(Received for publication, September 14, 1990)

Alessandra Carattoli‡, Nicoletta Romano‡, Paola Ballario§, Giorgio Morelli¶, and Giuseppe Macino‡‖

From the ‡Dipartimento di Biopatologia Umana, Sezione di Biologia Cellulare Policlinico Umberto I, Università di Roma La Sapienza 00161, ¶Unita' di Nutrizione Sperimentale, Istituto Nazionale della Nutrizione Via Ardeatina 546 00178, and the §Dipartimento di Genetica e Biologia Molecolare, Centro di Studio per gli Acidi Nucleici, Università di Roma La Sapienza, Piazzale Aldo Moro 5 00185, Roma, Italy

In the filamentous fungus *Neurospora crassa* the biosynthesis of carotenoids is regulated by blue light. Here we report the characterization of the albino-3 (al-3) gene of *N. crassa*, which encodes the carotenoid biosynthetic enzyme geranylgeranyl-pyrophosphate synthetase. This is the first geranylgeranyl-pyrophosphate synthetase gene isolated. Nucleotide sequence comparison of al-3 genomic and cDNA clones revealed that the al-3 gene is not interrupted by introns. Transcription of the al-3 gene has been examined in dark-grown and light-induced mycelia. The analysis revealed that the al-3 gene is not expressed in the dark and that its transcription is induced by blue light (Nelson, M. A., Morelli, G., Carattoli, A., Romano, N., and Macino, G. (1989) *Mol. Cell. Biol.* 9, 1271–1276). The al-3 gene encodes a polypeptide of 428 amino acids. Comparison of the deduced amino acid sequence of al-3 with the sequences of prenyltransferases of other species, from bacteria to humans, showed three highly conserved homologous regions. These homologous regions may be involved in the formation of the catalytic site of the prenyltransferases.

Carotenoids are synthetized by bacteria, plants, fungi, and algae (2). While their primary functions are in photoprotection and as accessory pigments in photosynthesis, carotenoids also serve as precursors for vitamin A biosynthesis in animals and for abscissic acid biosynthesis in plants. In *Neurospora crassa* the biosynthesis of carotenoids is regulated by blue light in the mycelium but is constitutive in the asexual spores (3–5). The photoinduction of carotenogenesis in the mycelia requires the *de novo* synthesis of at least three enzymes which have been shown to be the products of the albino (al) genes (5). Three al mutants have been characterized in *N. crassa*, each of which is defective in one step of carotenogenesis. Albino 3 (al-3) mutants are defective in GGPP[1] synthetase (5), while albino 2 and albino 1 mutants are defective in phytoene synthetase (6) and phytoene dehydrogenase (7), respectively. In previous work we isolated the gene encoding GGPP synthetase from *N. crassa* by complementation of the al-3 mutant; expression studies showed that the transcription of the al-3 gene is controlled by light regulation (1).

The pathway of carotenoid biosynthesis shares some steps with the biosynthetic pathways of other isoprenoid compounds. GGPP synthetase and several other enzymes of these complex pathways are members of the prenyltransferase family. Prenyltransferases catalyze the transfer of an isoprenoid diphosphate to another isoprenoid diphosphate or to a nonisoprenoid compound through a 1′-4 condensation reaction to produce various prenyl compounds that are precursors of such diverse products as steroids, carotenoids, chlorophylls, heme a, prenylated proteins and tRNAs, glycosyl carrier lipids, plant hormones, and the side chains of quinones (8–10). Prenyltransferases produce a wide range of products, from the simple dimer geranyl pyrophosphate to the complex structure of rubber which is thousands of monomers long. GGPP synthetase catalyzes the *trans* addition of three molecules of IPP onto DMAPP to form geranylgeranyl pyrophosphate.

Here we present the sequence of the al-3 gene and its deduced amino acid sequence. The al-3 protein sequence is compared with those of other known prenyltransferases.

MATERIALS AND METHODS AND RESULTS[2]

DISCUSSION

We have determined the nucleotide sequence of the al-3 gene, which encodes the carotenoid biosynthetic enzyme GGPP synthetase. It is known that blue light induces the biosynthesis of carotenoids in *N. crassa* mycelia and that the activity of GGPP synthetase increases after light treatment (8). The al-3 gene encodes an mRNA of 1683 nucleotides, which is colinear with the al-3 gene, as revealed by cDNA sequence and S1 nuclease mapping analysis. We analyzed the expression of the al-3 gene and found that its mRNA is not present in dark-grown mycelia but is induced by blue light after a short pulse of illumination.

* This work was supported by grants from the Ministero Agricoltura e Foreste, Piano Nazionale Tecnologie Applicate alle Piante, and Instituto Pasteur-Fondazione Cenci Bolognetti. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) X53979.

‖ To whom correspondence should be addressed. Tel.: 6-445-2806; Fax: 6-446-2891.

[1] The abbreviations used are: GGPP, geranylgeranyl pyrophosphate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; FGSC, Fungal Genetic Stock Center; PIPES, 1,4 Piperazinediethanesulfonic acid; FPP, farnesyl pyrophosphate.

[2] Portions of this paper (including "Materials and Methods," "Results," Table 1, and Figs. 1–4) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

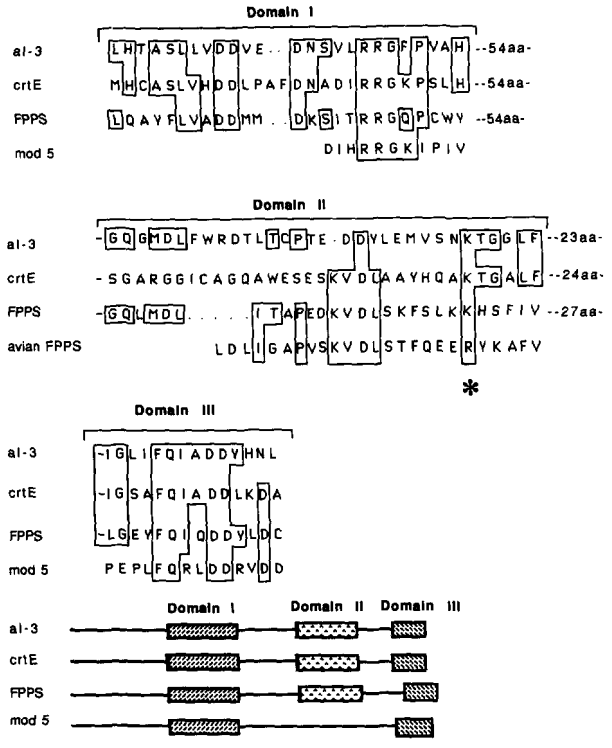

FIG. 5. **Alignment of the deduced al-3 protein sequence with the *S. cerevisiae* FPPS and *Rhodobacter capsulatus* crtE proteins.** Domains I, II, and III show amino acid alignment of the al-3, FPPS, and crtE proteins. Conserved amino acid residues present in at least two of the three sequences are *boxed*. The distance between domains is indicated. Domains I and III of the *S. cerevisiae* mod5 protein and domain II of avian liver FPPS are also shown. The *asterisk* indicates the avian FPPS arginine residue discussed under "Results." The three domains of the al-3, FPPS, crtE, and mod5 proteins and their relative positions within the genes are shown diagrammatically at the bottom (not drawn to scale).

The polypeptide encoded by the al-3 gene has a molecular mass of 47,876 daltons, is weakly basic and hydrophilic, and does not possess any hydrophobic membrane-spanning regions. This is in agreement with the finding that the GGPPS activity, isolated from various sources, is present in the soluble fraction of cellular extracts (8, 36, 37). GGPP synthetase is a prenyltransferase that catalyzes the 1′-4 condensation of dimethylallyl pyrophosphate with three isopentenyl pyrophosphates. We therefore compared the al-3 polypeptide with other known prenyltransferases. Comparison with the FPP synthetase from human (29), rat (30), and *Saccharomyces cerevisiae* (28) showed significant homologies in three different regions of these proteins. The relative positions of the homologous regions were the same in all the proteins analyzed. Furthermore, the three domains were localized in the proteins at very similar distances. These facts suggest that the homologous regions may be involved in the formation of the catalytic site of the prenyltransferases. Comparison of domains I and III showed the presence of the motif DD*XX*D. These aspartate residues could be responsible for the binding of the cations Mg^{2+} or Mn^{2+}, shown to be important for the catalytic activities of the prenyltransferases (38, 39). The analysis of the conserved amino acids found in the three domains suggested that in all three domains not only the aspartate residues but also the positively charged amino acids may be important for enzyme activity. One of the major biological functions of arginine residues is to interact with phosphorylated metabolites. Lysine residues may also serve this function and indeed are known to be important in a number of enzymes acting on phosphorylated substrates (40, 41). Furthermore, it has been demonstrated that the arginine-specific reagent hydroxyphenylglyoxal is a powerful inhibitor of prenyltransferases (38, 42).

Domain II is homologous to the region proposed to be the active site of avian FPP synthetase. Brems *et al.* (32) identified this region by a site-directed photoaffinity label on the purified enzyme and proposed that the arginine residue (whose position is indicated in Fig. 5 with an *asterisk*), conserved also in human and rat FPP synthetase (not shown) and substituted by lysine in all the other known prenyl transferases, could be responsible for the binding of the pyrophosphate group. These authors also showed evidence of the involvement of 2 arginine residues in the function of prenyltransferases.

In view of the model that we propose for the function of these conserved domains, it is interesting that the yeast prenyltransferase (involved in tRNA modification) has some homology, involving arginyl residues, with the first domain and a stronger homology with the third domain, in which all 3 aspartate residues are conserved.

We have found that the crtE gene product of *Rhodobacter capsulatus* also has an impressive homology with all three conserved domains. All the prenyltransferases considered here perform a 1′-4 condensation reaction between an allylic and an homoallylic substrate. CrtE has been indicated to be a phytoene synthetase (31) and therefore it does not belong to the prenyltransferase family. Due to these homologies, we propose that the crtE enzyme may have a 1′-4 condensation activity.

These similarities and homologies among prenyltransferases suggest that the genes may have a phylogenetic relation. It is conceivable that other prenyltransferases could share the same regions of homology with those considered here. The prediction that the conserved amino acids, in the three homologous regions, play an essential role in the enzyme functions, will be tested by site-directed mutagenesis experiments on the al-3 gene, making use of the albine 3 mutant as the recipient strain for the transformation of mutated sequences.

Acknowledgments—We thank Enza Ilardi for providing us with the *N. crassa* cDNA clone bank. We are grateful to Uwe Bertsch and Eugenio Pontieri for assistance in the screening of the cDNA bank. We thank Tullio Aversa for technical support. We thank Tom Schmidhauser, Mary Anne Nelson, and Gloria Coruzzi for invaluable advice and suggestions.

Addendum—During the preparation of the manuscript, we learned from Dr. P. Edwards[3] that *S. cerevisiae* hexaprenyl pyrophosphate synthetase has amino acid sequence homology with domains I and III.

REFERENCES

1. Nelson, M. A., Morelli, G., Carattoli, A., Romano, N., and Macino, G. (1989) *Mol. Cell. Biol.* **9,** 1271–1276
2. Goodwin, T. W. (1980) in *The Biochemistry of Carotenoids* (Chapman, and Hall, eds) Vol 1, Plants, New York
3. De Fabo, E. C., Harding, R. W., and Shropshire, W. (1976) *Plant Physiol.* **57,** 440–445
4. Harding, R. W., and Shropshire, W., Jr. (1980) *Annu. Rev. Plant Physiol.* **31,** 217–238
5. Harding, R. W., and Turner, R. V. (1981) *Plant Physiol.* **68,** 745–749
6. Kushwaha, S. C., Kates, M., Renaud, R. L., and Subden, R. E. (1978) *Lipids* **13,** 352–355
7. Goldie, A. H., and Subden, R. E. (1973) *Biochem. Genet.* **10,** 275–284

[3] P. Edwards, Division of Cardiology, University of California at Los Angeles, personal communication.

8. Brown, M. S., and Golstein, J. L. (1980) *J. Lipid Res.* **21,** 505–517
9. Anderegg, R. J., Betz, R., Carr, S. A., Crabb, J. W., and Duntze, W. (1988) *J. Biol. Chem.* **263,** 18236–18240
10. Glomset, J. A., Gelb, M. H., and Farnsworth, C. C. (1990) *Trends. Biochem. Sci.* **15,** 139–142
11. Davis, R. H., and De Serres, F. J. (1970) *Biochemistry* **18,** 5294–5299
12. Schweizer, M., Case, M. E., Dykstra, C. C., Giles, N. H., and Kushner, S. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 5086–5090
13. Kuiper, M. T. R., Akins, R. A., Holtrop, M., de Vries, H., and Lambowitz, A. M. (1988) *J. Biol. Chem.* **263,** 2840–2847
14. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. Vogel, H. J. (1964) *Am. Nat.* **98,** 435–446
16. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299
17. Glišin, V., Crkvenjakov, R., and Byus, C. (1974) *Biochemistry* **13,** 2633–2637
18. Aviv, H., and Leder, P. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69,** 1408–1412
19. Lehrach, H., Diamond, D., Wozney, J. M., and Boedtker, H. (1977) *Biochemistry* **16,** 4743–4751
20. Cobianchi, F., and Wilson, S. H. (1987) in *Methods Enzymol.* **152,** 94–97
21. Sanger, F., Coulson, A. R., Barrel, B. G., Smith, A. J. H., and Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178
22. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395
23. Ballance, D. J. (1986) *Yeast* **2,** 229–236
24. Orbach, M. J., Porro, E. B., and Yanofsky, C. (1986) *Mol. Cell. Biol.* **6,** 2452–2461
25. Legerton, T. L., Yanofsky, C. (1985) *Gene (Amst.)* **39,** 129–140
26. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132
27. Roberts, A. N., and Yanofsky, C. (1989) *Nucleic Acids Res.* **17,** 197–214
28. Anderson, M. S., Yarger, J. G., Burck, C. L., and Poulter, C. D. (1989) *J. Biol. Chem.* **264,** 19176–19184
29. Sheares, B. T., White, S. S., Molowa, D. T., Chan, K., Ding, V. D.-H., Kroon, P. A., Bostedor, R. G., and Karkas, J. D. (1989) *Biochemistry* **28,** 8129–8135
30. Clarke, C. F., Tanaka, R. D., Svenson, K., Wamsley, M., Fogelman, A. M., and Edwards, P. A. (1987) *Mol. Cell. Biol.* **7,** 3138–3146
31. Armstrong, G. A., Alberti, M., Leach, F., and Hearst, J. E. (1989) *Mol. Gen. Genet.* **216,** 254–268
32. Brems, D. N., Bruenger, E., and Rilling, H. C. (1981) *Biochemistry* **20,** 3711–3718
33. Najarian, D., Dihanich, M. E., Martin, N. C., and Hopper, A. K. (1987) *Mol. Cell. Biol.* **7,** 185–191
34. Anderson, M. S., Muehlbacher, M., Street, I. P., Proffitt, J., and Poulter, C. D. (1989) *J. Biol. Chem.* **264,** 19169–19175
35. Powell, G. K., and Morris, R. O. (1986) *Nucleic Acids Res.* **14,** 2555–2565
36. Brinkhaus, F. L., and Rilling, H. C. (1988) *Arch. Biochem. Biophys.* **266,** 607–612
37. Dogbo, O., and Camara, B. (1987) *Biochim. Biophys. Acta* **920,** 140–148
38. Dogbo, O., Laferriere, A., d'Harlingue, A., and Camara, B. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 7054–7058
39. King, H. L., Jr., and Rilling, H. C. (1977) *Biochemistry* **16,** 3815–3819
40. Riordan, J. F., McElvany, K. D., and Borders, C. L., Jr. (1977) *Science* **195,** 884–886
41. Xia, Z., and Storm, D. R. (1990) *J. Biol. Chem.* **265,** 6517–6520
42. Barnard, G. F., and Popjak, G. (1980) *Biochim. Biophys. Acta* **617,** 169–182

Supplemental material to :

The *Neurospora crassa* carotenoid biosynthetic gene (albino 3) reveals highly conserved regions among prenyltransferases.

by

Alessandra Carattoli, Nicoletta Romano, Paola Ballario, Giorgio Morelli and Giuseppe Macino.

MATERIALS AND METHODS

N. crassa strains and plasmids: The following *Neurospora crassa* wild-type and mutant strains (al, albino; qa, quinate; aro, aromatic cluster gene) were obtained from the FGSC[1] (University of Kansas, Kansas City, Kansas): Oak Ridge wild-type 74-OR23A (FGSC[1] no. 987), al-3 (RP100; FGSC[1] no. 2082) and qa-2 aro-9 (M246 Y325M6; FGSC[1] no. 4407). The al-3 qa-2 aro-9 (RP100 M246 Y325M6) strain was constructed by standard methods(11).
The plasmid clone bank containing inserts of *N. crassa* DNA in the pRAL1 vector was provided by Alan M. Lambowitz (Ohio State University). The pRAL1 vector contains the qa-2[+] gene that serves as the selectable marker in transformations of *N. crassa* and the chloramphenicol resistance gene for selection in *Escherichia coli*. The pNC39 plasmid was isolated from this bank and pNC39ΔNruI was made by the circularization of pNC39 restricted with NruI. The *N.crassa* light-induced cDNA library was provided by Enza Ilardi, Istituto Nazionale della Nutrizione, Rome.

Transformation: Preparation of al-3 spheroplasts and their transformation with pRAL1, pNC39 and pNC39ΔNruI were as described by Schweizer et al. (12). Transformants were plated onto selection plates after regeneration in liquid medium (13). Transformation of *E.coli* was carried out as described in Maniatis et al. (14).

Media and culture conditions: For photoinduction experiments, conidia were inoculated at $2x10^5$ cells per ml in 100 ml Vogel minimal medium (15) plus 2% sucrose and shaken in the dark at 30°C for 18 h. Selected cultures were illuminated directly in the flask by applying 250 $\mu W/cm^2$ in the blue region for 20 min with two Sylvania GRO-LUX F 18W-GRO lamps. Light-induced and dark-grown mycelia were collected for total RNA preparation by filtration onto filter paper disks and immediately frozen in liquid nitrogen.

Colony hybridization: The cDNA library was screened by colony hybridization. Filters were prepared and hybridized as described by Maniatis et al. (14).
The ^{32}P labeled double stranded probe of the al-3 gene was prepared by nick-translation (14).

Preparation of RNA: Total RNA was extracted from frozen mycelia powdered in a Waring blendor under liquid nitrogen. Powdered mycelia were treated with guanidinium thiocyanate as a protein denaturant (16) and RNA was collected by centrifugation through CsCl cushions (17). Polyadenylated RNA was isolated by chromatography on oligo(dT)-cellulose (18).

Northern (RNA) hybridizations: RNA was denatured, electrophoresed on agarose gels containing formaldehyde (19), and transferred onto Hybond-N membranes. The ^{32}P-labelled single stranded probes were prepared by primer extension on M13 templates in the presence of ^{32}P labeled ATP (20), followed by digestion with the appropriate restriction enzymes and strand separation on a 6% acrylamide/bis gel. Filters were hybridized at 48°C in 5X SSPE (0.9 M NaCl, 0.05 M sodium phosphate (pH 7.7), 0.05 M EDTA), 5X Denhardt solution (0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinyl pyrrolidone), 50% formamide, and 0.1% sodium dodecyl sulfate. Filters were washed at 55°C in 0.1% SSPE-0.1%SDS for 15 min.

DNA sequencing: The HindIII-SalI, SalI-EcoRI, EcoRI-EcoRI and BamHI-AccI fragments obtained from pNC39ΔNruI were sucloned in M13mp18 and in M13mp19. All DNA sequences were determined by primer extension on M13 clones using the dideoxy chain termination method (21). Synthetic oligodeoxynucleotides were used to obtain full sequence of both strands. The cDNA clone was sequenced on both strands by the same method . No differences were found between the genomic and cDNA sequences.

S1 nuclease mapping : To perform the 5' S1 nuclease mapping experiments, uniformly-labeled single-stranded DNA probes were prepared from a pNC39ΔNruI SalI- HindIII fragment cloned in M13mp18, using as primer an internal oligonucleotide, indicated in.Fig. 2. S1 nuclease mapping experiments of the 3' end of the al-3 mRNA were performed using a pNC39ΔNruI EcoRI- AccI restriction fragment that was 3' end-labelled with ^{32}P using the fill-in technique. Probes were hybridized with 20-30 µg of total RNA or with 2 µg of poly(A)[+] RNA in 0.4M NaCl, 20mM PIPES[1] pH 6.4, 2 mM EDTA at 42°C for 3 hours. S1 nuclease digestions were carried out in S1 digestion mix (0.3M NaCl, 30mM NaAcetate pH 4.6, 1mM $ZnSO_4$) containing 300 U/ml of S1 nuclease. When poly(A)+ RNA was used, 20 µg of denatured heterologous DNA were added to the S1 digestion mixtures. The digestion products were analyzed by 6% acrylamide/bis-urea gel electrophoresis.

Computer methods: Comparative sequence analyses were performed with programs from the University of Wisconsin Genetics Computer Group (22).

RESULTS

Localization of the al-3 gene in pNC39

We had previously cloned the GGPP[1] synthetase gene by transformation of an al-3 qa-2 aro-9 mutant of *N. crassa* with a genomic library in the pRAL1 plasmid (see Materials and Methods) (1). From this library we had isolated the plasmid pNC39, with a 7.0 kbp genomic DNA insert containing the intact al-3 gene. The 62% of the transformants containing this plasmid showed dark orange pigmented conidia, in contrast to the albino phenotype of the al-3 parental strain (table 1). To determine the minimal region of the insert necessary for complementation, we transformed the al-3 qa-2 aro-9 strain with several plasmids containing various subclones of the pNC39 insert. The pNC39ΔNruI plasmid , with an insert of 3.0 kbp (obtained by deletion of the 5.0 kbp NruI fragment of the pNC39 plasmid (Fig.1)) was still able to complement the al-3 mutation showing 54 % wild type transformants (table 1). Thus pNC39ΔNruI plasmid was chosen for further studies.

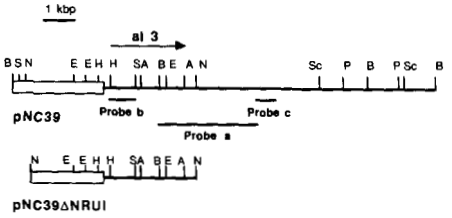

Fig. 1: Restriction maps of the al-3 containing and related plasmids. *Neurospora crassa* sequences are indicated by horizontal lines; boxes represent vector sequences. Abbreviations for restriction sites: **B**: BamHI, **E**: EcoRI, **H**: HindIII, **S**: SalI, **Sc**: SacI, **N** : NruI; **P**: PstI; **A**: AccI. Probes a, b and c indicate probes used in Northern blot experiments, discussed in the text. The *Neurospora* and vector sequences are drawn to scale (indicated by a 1 kbp bar).

Plasmids	Phenotypes of qa+ transformants wt %	al-3 %
pRAL1	0	100
pNC39	62	38
pNC39Δ Nru I	54	46

Table 1: Percentage of wild type and albino 3 phenotypes obtained by al-3 qa-2 aro-9 strain transformation with different plasmids.
The pNC39 plasmid contains a 7.0 kbp *N. crassa* genomic insert cloned in the pRAL 1 vector. The pNC39ΔNruI plasmid contains a 3.0 kbp DNA insert obtained by deletion of the 5.0 kbp NruI fragment of plasmid pNC39.

DNA sequence of the al-3 gene and its predicted gene product

The HindIII-SalI, SalI-EcoRI, EcoRI-EcoRI and BamHI-AccI fragments from pNC39ΔNruI were subcloned into M13mp18 and M13mp19 for nucleotide sequencing. Each construct was used as the template in sequencing reactions using the universal primer (21). The complete sequence of the clones was obtained using synthetic primers corresponding to the ends of previously sequenced DNA. Both strands were sequenced with overlaps to resolve any ambiguities.The DNA sequence of the al-3 coding and flanking regions is shown in Figure 2. This DNA segment corresponds to the 2.0 kbp HindIII-AccI fragment of Fig.1. Within the sequence there is only one sizable reading frame, which encodes a protein of 428 amino acids. This reading frame is translated above the nucleotide sequence in Fig. 2. The codon usage of the putative gene product follows the general rules for *N. crassa* genes, which includes a strong bias towards the use of codons terminating in C (42%) or G (33%) (23,24).

There are two possible initiation codons at the beginning of the open reading frame. The *bona fide* methionine codon of the al-3 reading frame most likely occurs at position 162 (Fig. 2). The nucleotides that precede this codon (cgtcaccATGG) are believed to be important for correct translation and are found in the same position in many *N. crassa* genes (25). The putative al-3 polypeptide has a molecular weight of about 47.8 Kd. The predicted product is weakly basic and hydrophilic. The hydropathy profile of Kyte and Doolittle (26) suggested that there are no hydrophobic membrane- spanning regions (not shown).

The coding portion of the gene has a G+C content of 57%, while the 5' and 3' non-translated sequences have G+C contents of 45% and 42%, respectively. This is in agreement with what has been found in other *N. crassa* genes (27). The 5' non-transcribed region contains a putative TATA box about 30 nucleotides upstream of the 5' transcriptional initiation site. A search for short sequences homologous to published promoter sequences of other *N.crassa* genes failed to turn up any significant homologies. A potential polyadenylation recognition element (AATGAA) is present 50 bp upstream of the 3' end of the mRNA.

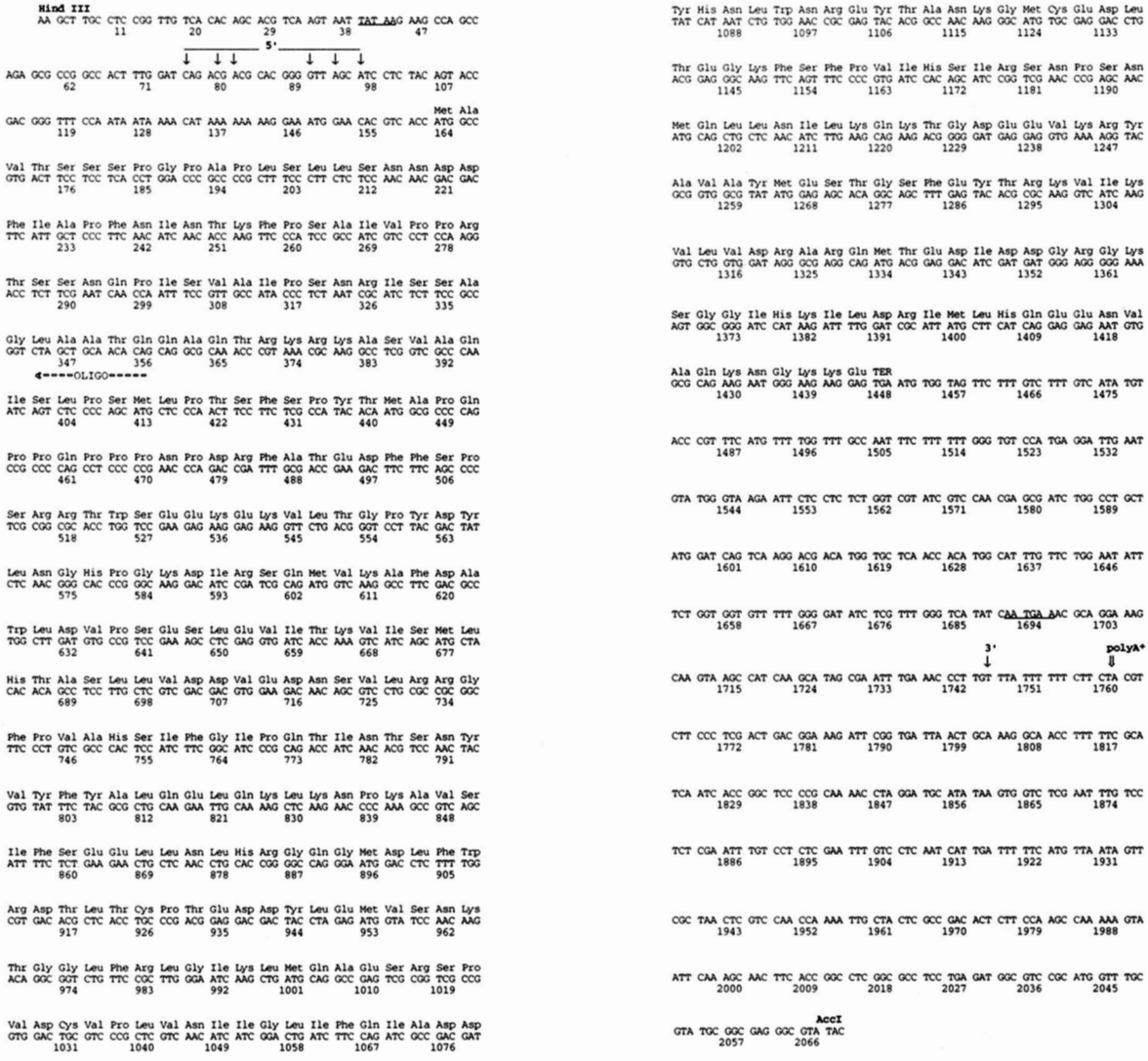

Fig.2: DNA sequence of the al-3 gene and its predicted gene product. Nucleotide sequence of the 2.0 kbp HindIII-AccI fragment of Fig. 2, containing the coding region of the al-3 gene. The putative TATA sequence and polyadenylation recognition signal are underlined. The arrows indicate the 5' and 3' ends of the mRNA as determined by S1 nuclease mapping. The open arrow shows the mRNA 3' end determined by cDNA sequence. The sequence of the oligonucleotide used for labeling the probe for 5' end studies is also shown.

Transcript analysis

In a previous paper (1) we demonstrated that the al-3 mRNA level increases about 15 fold in mycelia, after a short period of illumination, compared to its low level in the dark. Here we present results that confirm the photoinduction of the al-3 mRNA, but on the contrary its level in the dark is undetectable by Northern blotting analysis. This difference is due to the fact that the previous Northern blots were probed with a DNA fragment that contained the al-3 gene plus another gene. This gene is closely linked to al-3 and encodes a messenger RNA of almost identical size (about 1700 nucleotides); the linked gene is transcribed in the opposite direction and expressed in both light and dark regimes [4]. Fig. 4A shows that the longer probe a (Fig. 1) that had been used in previous work (1) detected both mRNAs, that of al-3 plus the mRNA of the closely-linked but unrelated gene. Fig. 3 demonstrated that when a probe was used which contains only the al-3 gene (probe b), the mRNA for al-3 was detected only in light-grown mycelia and was undetectable in dark-grown mycelia. Probe c detected only the mRNA of the linked gene; this mRNA comigrated with the al-3 mRNA, but was not photoregulated. These results have been confirmed with more sensitive S1 nuclease experiments (Fig. 4). Probes 1 and 2 were totally degraded when total RNA from dark-grown mycelia was used, while they were completely protected using total or poly(A)+ RNA from light-induced mycelia (see below).

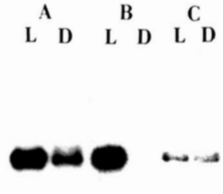

Fig.3: Northern hybridizations. Total mRNA was prepared from light-induced (L) and from dark-grown (D) mycelia. Samples (4 µg) of each RNA were separated on a 1.2% agarose-6% formaldeyde gel, transferred to Hybond-N membrane, and hybridized with probes a, b and c (Fig. 2). Results are shown, respectively, in panels A, B and C

4- A. Carattoli unpublished results

Isolation of cDNA clones of al-3

To confirm that the al-3 gene contains no intervening sequences, we isolated cDNA clones from a library made from light-induced mycelia. Using a 700 bp HindIII-SalI ^{32}P radiolabeled fragment of the al-3 gene (probe b in Fig. 1), 10 positive clones were identified by screening 60,000 colonies of the library. The longest positive cDNA clone was selected for sequence analysis. The complete sequence of the al-3 cDNA clone revealed that the cDNA was shorter, at the 5' end, by about 80 nucleotides than the putative mRNA determined by S1 nuclease mapping (see below). The sequence analysis of the cDNA showed colinearity with the genomic sequence, indicating that there are no introns in this gene.

S1 nuclease mapping

To determine the 5' initiation and 3' termination sites of the al-3 mRNA, S1 protection experiments were performed (Fig. 4). The 5' end of the al-3 mRNA was mapped as shown in Fig. 4A. A DNA probe spanning the 5' end of the al-3 gene (probe 1, 358 bp) was uniformly labeled, hybridized to RNA and treated with S1 nuclease. Fig. 4A shows that S1 nuclease-protected products obtained with total RNA (lane 2) or poly(A)+ RNA (lane 3), corresponded to two groups of closely-spaced bands (from 265 to 283 nucleotides in length). These multiple bands most likely represent multiple transcription initiation sites. This is consistent with the finding that, in fungi, multiple initiation sites are quite common (27).

Fig. 4B shows the 3' end mapping of the al-3 mRNA. Probe 2 (520 bp) was protected from S1 nuclease digestion, producing a protected fragment of 200 nucleotides. This result suggested that the 3' end of the al-3 mRNA was at about nucleotide 1745 (Fig. 2). The cDNA sequence revealed, instead, that the 3' end of the al-3 mRNA was 15 nucleotides further downstream (nucleotide 1759, Fig. 2). This discrepancy could be explained by the fact that the region immediately downstream from nucleotide 1745 is very T-rich. This peculiar sequence has a low hybrid stability, and therefore may produce shorter hybrid molecules in S1 nuclease protection experiments. The length of the al-3 mRNA, as revealed by cDNA sequence and S1 nuclease experiments, was 1683 nucleotides.

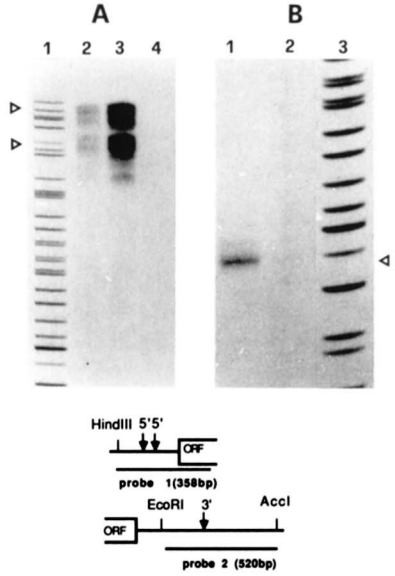

Fig.4: S1 nuclease mapping. S1 nuclease mapping of the 5' end (probe 1, panel A) and 3' end (probe 2, panel B) of al-3 mRNA. L indicates RNA extracted from light-induced mycelia and D RNA extracted from dark-grown mycelia. Probes 1 and 2 are shown below the panels.
Panel A: lanel, sequence ladder; lane 2, probe 1 plus 20 µg of total L RNA, S1-treated; lane 3, probe 1 plus 2 µg of poly(A)+ L RNA, S1-treated; lane 4, probe 1 plus 20 µg of total D RNA, S1-treated
Panel B: lane 1,probe 2 plus 30 µg of total L RNA, S1-treated; lane 2, probe 2 plus 30 µg of total D RNA, S1-treated; lane 3, sequence ladder.Arrow heads in panels A and B indicate the fragments protected from S1 nuclease digestion.

Homology with other prenyltransferases

Comparison of the al-3 sequence with the FPP[1] synthetase from *Saccharomyces cerevisiae* (28) showed a significant homology in three different regions of these two proteins. These regions are also highly conserved between the FPP[1] synthetase from human (29) and rat (30) (data not shown). This sequence homology is also evident in the sequence of the crtE gene, which is part of the carotenoid biosynthetic cluster from *Rhodobacter capsulatus* (31), and which contains the same three homologous domains. The relative positions of the homologous domains are the same in these three proteins. Fig. 5 shows the alignment of the three domains of al-3 with the yeast FPP[1] synthetase and *R. capsulatus* crtE proteins.

When al-3 was compared with FPP[1] synthetase and crtE, the first region of homology (domain I in Fig. 5) showed 42% and 53% amino acid identity, or 54% and 65% amino acid similarity, respectively. A closer examination revealed a consensus sequence (LXXDDXXDXXXRRGXP) of highly-charged amino acids. The second region (domain II) showed a weaker homology among the three proteins, but it is still possible to consider it as a conserved region if the al-3 sequence is seen as a mosaic of amino acids of the other two proteins, with the left portion being homologous to FPP[1] synthetase, and the right portion to crtE. It is interesting to note that this domain has been indicated to be the active site of prenyltransferases by work of Brems et al.(32). The third region (domain III) had 60% to 70% homology, among the three related proteins, over a span of 13 amino acids. The consensus sequence of this region (FQIXDDXXD/N) is also highly charged, as is that of the first domain. It is noteworthy that, at the distance of 12 amino acids from domain III, there is an invariant lysine residue .

The three domains are localized in the proteins with very similar spacing, with 54-59 residues between domain I and domain II, 23-37 residues beetwen domains II and III. Furthermore, the three proteins share other similarities. Domains I and III are flanked by short hydrophobic regions of about 20 amino acids. It is interesting to note that these are the only regions with hydrophobic character in these proteins . Domain II is immersed in a region of strong acidic character. The carboxy termini of both GGPP[1] synthetase and FPP[1] synthetase are highly charged.

There are other enzymes that make use of DMAPP[1] to prenylate non-isoprene substrates whose genes have been isolated. One of these is the *S. cerevisiae* mod5 gene, encoding the tRNA-modifying enzyme Δ2-isopentenyl pyrophosphate:tRNA isopentenyl transferase (33). Comparison of the al-3, FPP[1] synthetase and crtE amino acid sequences with that of mod5 revealed only a very limited homology to the conserved domains I and III. In domain I, the only residues conserved are four contiguous amino acids (RRGK) of the consensus sequence. In the third domain, the homology includes the entire consensus sequence (FQXXDDXXD).This limited homology should be considered significant, due to the conservation of amino acids whose role could be central in these domains (see Discussion). It is noteworthy that the distance between the mod5 I and III domains is very similar to the respective distances in al-3 and crtE.

Other enzymes utilizing IPP[1] or DMAPP[1] as substrates, such as IPP- DMAPP[1] isomerase from *S. cerevisiae* (34) and cytokinin prenyltransferases from *Agrobacterium tumefaciens* and *Pseudomonas savastanoi* (35), do not share any homology with the prenyl transferases discussed in this paper.