Biochem. J. (1992) 284, 861–867 (Printed in Great Britain)

Overproduction in *Escherichia coli* of the dehydroquinate synthase domain of the *Aspergillus nidulans* pentafunctional AROM protein

Johannes P. T. W. VAN DEN HOMBERGH,*‡ Jonathan D. MOORE,* Ian G. CHARLES† and Alastair R. HAWKINS*§

* Department of Biochemistry and Genetics, Catherine Cookson Building, Medical School, University of Newcastle upon Tyne, NE2 4HH, U.K., and † Department of Cell Biology, Wellcome Foundation Ltd, Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS, U.K.

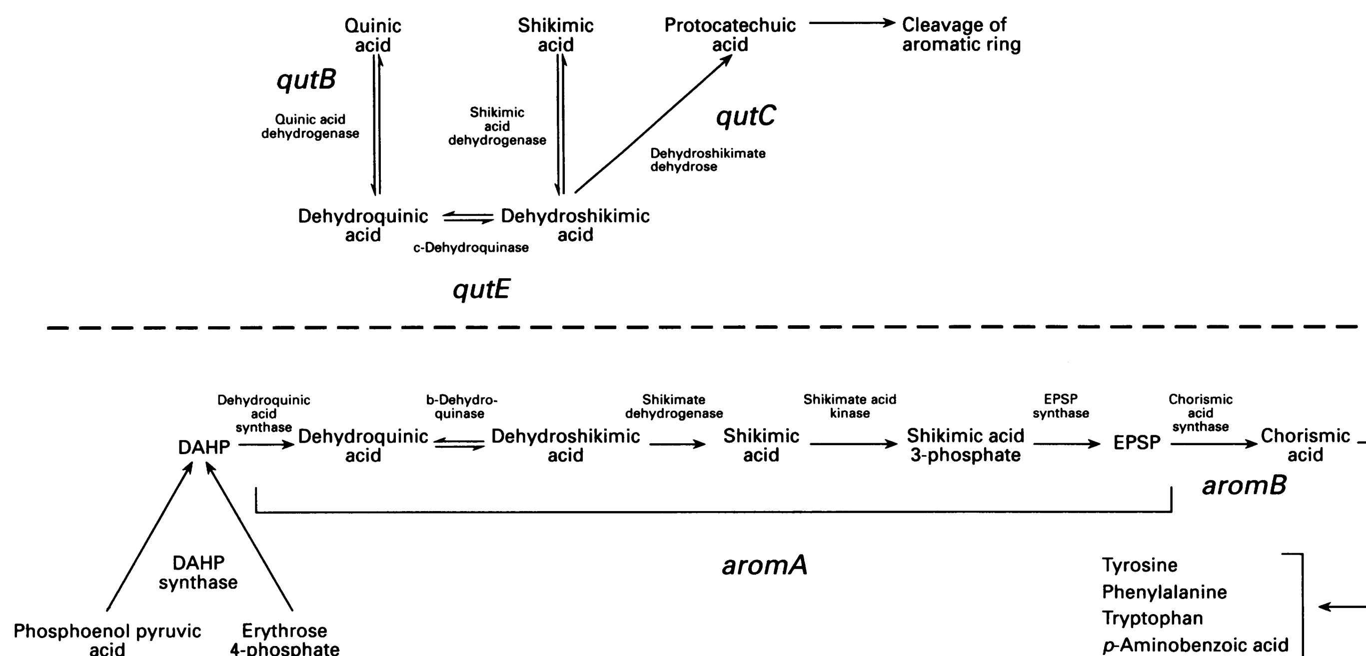
The pentafunctional AROM protein of *Aspergillus nidulans* is encoded by the complex *aromA* locus and catalyses steps 2–6 in the synthesis of chorismate, the common precursor for the aromatic amino acids and *p*-aminobenzoic acid. DNA sequences encoding the 3-dehydroquinate synthase (DHQ synthase) and 3-dehydroquinase domains of the AROM protein have been amplified with the inclusion of a translational stop codon at the *C*-terminus by PCR technology. These amplified fragments of DNA have been subcloned into the prokaryotic expression vector pKK233-2 and expressed in *Escherichia coli*. As a result, the DHQ synthase domain is overproduced in *E. coli*, forming 30 % of total cell protein, and can be purified to greater than 80 % homogeneity by a simple two-step protocol. The 3-dehydroquinase domain is produced at a specific activity 8-fold greater than the corresponding activity encoded by the *aromA* gene in *A. nidulans*. The *qutB* gene of *A. nidulans* encoding quinate dehydrogenase has similarly been subjected to PCR amplification and expression in *E. coli*. The quinate dehydrogenase is not overproduced, but is active in *E. coli* as a shikimate dehydrogenase, as the presence of the *qutB* gene allows the growth of an *E. coli* mutant strain lacking shikimate dehydrogenase on minimal medium lacking aromatic-amino-acid supplementation.

INTRODUCTION

Steps 2–6 in the shikimate pathway, which lead to the production of chorismate, the common precursor for the aromatic amino acids and *p*-aminobenzoic acid, are catalysed by a pentafunctional enzyme, the AROM protein, which is encoded by the complex *aromA* locus in *Aspergillus nidulans* (Kinghorn & Hawkins, 1982; Charles *et al.*, 1985, 1986). The *aromA* locus has been isolated by molecular cloning and shown to consist of a single open reading frame of 4812 nucleotides, whose inferred protein sequence has strong similarity to the five equivalent monofunctional *Escherichia coli* enzymes (Charles *et al.*, 1986). On the basis of this clear similarity the *aromA* locus is proposed to have evolved by the fusion of five loci to encode a single pentafunctional protein (Hawkins, 1987). Two of the intermediates metabolized by the AROM protein, 3-dehydroquinate (DHQ) and dehydroshikimate (DHS) are also present in the pathway catabolizing quinate to protocatechuate (see Scheme 1; Grant *et al.*, 1988; Beri *et al.*, 1987; Whittington *et al.*, 1987; Lamb *et al.*, 1990). DHQ and DHS are interconverted by two quite separate 3-dehydroquinase enzyme activities, one encoded by the AROM protein and one encoded by the single function *qutE* gene, but these enzymes have no significant protein sequence similarity, suggesting they arose by convergent evolution (Charles *et al.*, 1985; Da Silva *et al.*, 1986). Recent data have shown the pools of DHQ and DHS in the two pathways are exchangeable and that these metabolites leak from the AROM protein at a rate comparable with the rate of flux through the shikimate pathway (Lamb *et al.*, 1991, 1992). The AROM protein does, however, share a complex evolutionary history with other enzymes and proteins in the quinic acid pathway in that (*a*) the quinate dehydrogenase encoded by the *qutB* gene shows significant similarity with the shikimate dehydrogenase domain of the AROM protein (Hawkins *et al.*, 1988), suggesting that it may have shikimate dehydrogenase activity *in vivo*, (*b*) the repressor protein encoded by the *qutR* gene shows strong sequence similarity throughout its length with the three *C*-terminal domains of the AROM protein encoding shikimate kinase, 3-dehydroquinase and shikimate dehydrogenase (Anton *et al.*, 1987; Hawkins *et al.*, 1992), and (*c*) Garbe *et al.* (1991) have pointed out that the central portion of the AROM dehydroquinate synthase domain shows some sequence similarity with the 3-dehydroquinase enzyme encoded by the *qutE* gene, a possible case of retro-evolution (Horowitz, 1965). In order to probe the structure of the AROM protein and test the hypothesis that the *aromA* gene evolved by multiple gene fusions, a programme of site-specific mutagenesis was previously undertaken (Hawkins & Smith, 1991). The AROM protein has been shown to fall in two approximately equal halves, with the *N* terminal half consisting of DHQ synthase and 3-phosphoshikimate 1-carboxyvinyltransferase (EPSP synthase), and the *C*-terminal half comprising shikimate kinase, 3-dehydroquinase and shikimate dehydrogenase. The DHQ synthase and 3-dehydroquinase domains of the AROM protein have been shown to fold and function as monofunctional enzymes when the appropriate DNA sequence is expressed in *E. coli*, but the EPSP synthase domain only appears to be enzymically active when produced as at least a bifunctional protein with DHQ synthase (Hawkins & Smith, 1991). The 3-dehydroquinase encoded by the *qutE* gene has been purified to homogeneity from an

Abbreviations used: DHQ, 3-dehydroquinate; DHS, dehydroshikimate; *qut*, wild-type allele at a quinic acid-utilization-gene locus; *aromA*, the wild-type allele encoding the pentafunctional AROM protein; AROM, the pentafunctional AROM protein; DHQ synthase, 3-dehydroquinate synthase (EC 4.6.1.3); EPSP synthase, 3-phosphoshikimate 1-carboxyvinyltransferase; IPTG, isopropyl thiogalactoside.

‡ Present address: Department of Genetics, Wageningen Agricultural University, DreijenLaan 2, 6703 HA, Wageningen, The Netherlands.
§ To whom correspondence should be addressed.



Scheme 1. Genes, enzymes and metabolites comprising the shikimate and quinate pathways in *A. nidulans*

Abbreviations: DAHP, 3-deoxy-D-arabinoheptulosonic acid 3-phosphate; b-, biosynthetic; c-, catabolic.

overexpressing *E. coli* strain (Hawkins & Smith, 1991), and preliminary trials have led to the production of small protein crystals (N. Isaacs, J. D. Moore & A. R. Hawkins, unpublished work). A major aim of our overall research programme is to analyse the structure and function relationships between the enzymes of the shikimate and quinate pathways and further probe the domain structure of the AROM protein. Towards this goal, we have (1) overexpressed DNA sequences from the *aromA* gene of *A. nidulans* encoding DHQ synthase in *E. coli*, resulting in a large overproduction of the DHQ synthase domain, and shown that it can be purified to > 80% homogeneity by a simple two-step protocol; (2) produced a monofunctional 3-dehydroquinase domain in *E. coli* encoded by the corresponding DNA sequence from the *A. nidulans aromA* gene. The enzyme activity produced in *E. coli* is 8-fold higher than the 3-dehydroquinase activity produced in wild-type strains of *A. nidulans*; (3) shown that the *qutB* gene encoding quinate dehydrogenase encodes a protein that has a shikimate dehydrogenase activity *in vivo*.

MATERIALS AND METHODS

The origins of the *E. coli* strains, recombinant lambda phage, recombinant plasmids, and methods for their propagation have been described previously (Willetts *et al.*, 1969; Blattner *et al.*, 1970; Hawkins *et al.*, 1985; Hawkins & Smith, 1991). Chemicals were of AnalaR or greater purity and obtained from local suppliers. Molecular-biology reagents and the high-expression vector pKK233-2 were obtained from Pharmacia or BRL–Gibco, and their routine use followed the manufacturer's recommended conditions. Routine recombinant DNA techniques followed the protocols of Maniatis *et al.* (1982), and dideoxy nucleoside DNA sequencing and the use of buffer gradient gels were as previously described (Sanger *et al.*, 1977; Biggin *et al.*, 1983). Specific 17-mer oligonucleotides were synthesized according to Matthes *et al.* (1984), with modifications to the wash cycle as described by Sproat & Gait (1985), but longer oligonucleotides were synthesized by means of an Applied Biosystems 381A DNA synthesizer.

3-Dehydroquinase and quinate dehydrogenase enzyme assays were as previously described (Hawkins *et al.*, 1984; Coggins *et al.*, 1987), and SDS/PAGE was performed by the method of Laemmli (1970), with a 5%-acrylamide stacking gel and a 7.5% separating gel.

PCR amplification

PCR reactions were carried out using a Gene Amp Kit from Perkin–Elmer/Cetus, the manufacturer's recommended general reaction conditions, in combination with a Biometra Trio Thermoblock, being followed. Individual reaction conditions were modified according to the following protocol: (1) for 30-mer oligonucleotides containing up to six mismatched bases, initial DNA denaturation was at 94 °C for 1.5 min, and annealing was at 50 °C for 1 min; extension took place at 72 °C for 1 min, and subsequent denaturation was at 94 °C for 1 min; the cycle was repeated 24 times; and (2) for 17-mer oligonucleotides with single mismatched bases: initial denaturation was at 94 °C for 1 min and annealing at 42.5 °C for 3 min; extension was at 42.5 °C for 12 min and subsequent denaturation at 94 °C for 1 min; the cycle was repeated 24 times.

In each case approx. 2 ng of template DNA was used in conjunction with appropriate primers at an approximate concentration of 0.15 μM each. The reactions were carried out in a final volume of 100 μl, and the efficiency of amplification was monitored by electophoresing 10 μl of the final reaction mixture through an 0.8%-agarose gel and staining with ethidium bromide. In each case, a very strong signal of the predicted size was observed, with no additional bands or any evidence of degradation as monitored by a complete lack of 'smearing' of the final PCR product. After electrophoretic analysis, the remaining 90% of the PCR product was subject to phenol extraction, ethanol precipitation, and, after drying, digestion with the restriction endonucleases *Nco*I and *Hin*dIII. Digestion with the two restriction enzymes was possible, since the oligonucleotides used

for the PCR reaction were designed to provide a sequence of at least 7 bp beyond the enzyme recognition site.

Isopropyl thiogalactoside (IPTG)-induced gene overexpression was initially screened for by inoculating individual colonies containing recombinant pKK233-2 or unmodified pKK233-2 (control) with a toothpick into 2 ml of Luria broth containing ampicillin (100 $\mu g \cdot ml^{-1}$) and IPTG (0.2 $mg \cdot ml^{-1}$). Cells were then grown with vigorous agitation for 12–14 h, harvested by centrifugation, and lysed in 100 μl of SDS/PAGE loading buffer by heating in a boiling-water bath for 5 min. Cells equivalent to an A_{550} of 0.1 were subjected to SDS/PAGE with appropriate markers, and the proteins were stained with Coomassie Brilliant Blue. Plasmids from overexpressing strains detected by this screening procedure were then transformed into strain JM105, which contains the *LacI*^{q} mutation, allowing more controlled IPTG-inducible gene overexpression. Strains of JM105 containing recombinant pKK233-2 or unmodified (control) plasmids were initially grown in Luria broth supplemented with ampicillin (100 $\mu g \cdot ml^{-1}$) at 37 °C, and then used to seed a second culture of drug-supplemented Luria broth to an A_{550} of 0.05. This culture was then aerated vigorously at 37 °C until an A_{550} of 0.2 was achieved, at which time IPTG (0.2 $mg \cdot ml^{-1}$) was added and the culture incubated for a further 8 h; cells equivalent to an A_{550} of 0.1 were taken at zero time and at 7 h intervals after that, and were subjected to SDS/PAGE as described above.

Sonicated cell-free extracts of *E. coli* were prepared for enzyme assay as described previously (Hawkins & Smith, 1991), and $(NH_4)_2SO_4$ precipitations of such extracts were achieved by thoroughly mixing appropriate volumes of extract and a 100%-saturated (at 4 °C) $(NH_4)_2SO_4$ solution and incubating the mixture at 4 °C for 1 h. The cells were sonicated in a buffer consisting of 50 mM-potassium phosphate, 1 mM-dithiothreitol and 1 mM-phenylmethanesulphonyl fluoride (buffer A).

RESULTS AND DISCUSSION

PCR amplification of the DNA sequences encoding the AROM protein DHQ synthase and 3-dehydroquinase domains and the *qutB* encoded quinate/shikimate dehydrogenase

The approximate *C*-terminus of the DHQ synthase domain lies within the DNA co-ordinates 1135–1236 reported in Charles *et al.* (1986) and is based on inferred protein sequence homology and expression studies with discrete portions of the *aromA* DNA sequence (Hawkins, 1987; Hawkins & Smith, 1991). Specific 30-mer oligonucleotides were designed to insert TGA and TAA translational stop signals at DNA co-ordinates 1180 and 1303, and *Hin*dIII-restriction-endonuclease recognition sites at co-ordinates 1182 and 1305; the translation stop signal at co-ordinate 1180 is in the position predicted to be the end of the DHQ synthase domain, and the translational stop signal at co-ordinate 1303 is within the EPSP synthase domain. A specific 30-mer primer introducing an *Nco*I site was used in conjunction with plasmid pKK13 (see Fig. 1) to introduce an in-frame translational start codon at the 5′ end of the DHQ synthase domain. The 5′ end of the 3-dehydroquinase domain has previously been located to co-ordinate 3102 (Charles *et al.*, 1986) and the 3′ end to approx. 3851. Specific 30-mer oligonucleotides were designed in conjunction with plasmid pKK28 (see Fig. 1) to (1) introduce an *Nco*I site containing an in-frame translational start codon at co-ordinate 3098 in the 5′ end of the 3-dehydroquinase domain, (2) introduce the translational stop signals TAA at co-ordinate 3859, TAG at co-ordinate 3870, and TGA at co-ordinates 3926, 3992 and 4073, and (3) introduce recognition sequences for the restriction endonuclease *Hin*dIII at co-ordinates 3864, 3874, 3928, 3994 and 4075 (see Hawkins *et al.*, 1988).

Two specific 17-mer oligonucleotides were used in conjunction with recombinant lambda phage λQ1 to (1) introduce an *Nco*I site containing an inframe translational start codon at the 5′ end of the *qutB* gene, and (2) introduce a recognition sequence for the restriction endonuclease *Hin*dIII 25 nucleotides downstream of the natural translational stop signal. Phage λQ1 is a recombinant bacteriophage containing the *A. nidulans qut-A*, *-B*, *-D*, *-E* and *-G* genes.

In each case the PCR reaction was carried out and monitored as described in the Materials and methods section, and led to the production of a high yield of a single species of DNA with the desired co-ordinates.

Subcloning and expression of the PCR products

After digestion with the restriction endonuclease *Nco*I and *Hin*dIII, the PCR products were ligated in-frame to the suitably digested *E. coli* high-expression vector pKK233-2, which contains the powerful hybrid *trc* promoter. Ligation mixes containing the AROM protein DHQ synthase domains were used to transform *E. coli* strain GLW38 (*aroB*$^-$; lacking DHQ synthase); ligation mixtures containing the AROM protein 3-dehydroquinase domain were used to transform *E. coli* strain SK3430 (*aroD*$^-$; lacking 3-dehydroquinase), and ligation mixtures containing the *qutB* gene encoding quinate dehydrogenase were used to transform *E. coli* strain GLW35 (*aroE*$^-$; lacking shikimate dehydrogenase). Transformants were selected by resistance to ampicillin on complete medium and subsequently replica-plated on to drug-free minimal medium with and without supplementation with aromatic amino acids, and incubated at 30 °C for 24 h.

The plasmids encoding DHQ synthase were designated pKK-40 and -41; those encoding 3-dehydroquinase pKK-35,-35 and -46–48 and those encoding quinate dehydrogenase, pKK42. Within any group of plasmids the plasmid with the highest number has the longest PCR-generated DNA fragment. The cloning of all eight PCR products was successful, and in the case of plasmids containing the DHQ synthase and 3-dehydroquinase domains, allowed the growth of *aroB*$^-$ and *aroD*$^-$ mutants of *E. coli* on unsupplemented minimal medium. The plasmids containing the *qutB* gene (designated pKK42), however, did not allow the growth of *aroE* mutants (lacking shikimate dehydrogenase) on unsupplemented minimal medium after incubation at 30 °C for 24 h. Representative colonies were then replated in a patch on to drug-supplemented complete medium and replica-plated on to supplemented and non-supplemented minimal medium. After incubation at 30 °C for 72 h, small colonies began to grow within the patch of cells on non-supplemented medium and continued to grow for a further 48 h. At this point cells from one of the colonies growing on the non-supplemented minimal medium were purified, their plasmid (designated pKK43) rescued, and used to re-transform the original *aroE*$^-$ mutant strain of *E. coli*. The transformants containing plasmid pKK43 were all capable of growth on non-supplemented minimal medium. We interpret the difference in growth response in otherwise identical strains of *E. coli* GLW55 (*aroE*$^-$; lacking shikimate dehydrogenase), containing pKK-42 or -43, to mean that pKK43 is a mutant derivative of pKK42. A very similar phenomenon was recently observed with plasmids encoding the AROM 3-dehydroquinase domain, where the difference between two plasmids containing the same AROM DNA sequence, one of which allowed growth of an *aroD*$^-$ (lacking 3-dehydroquinase) mutant of *E. coli* on non-supplemented minimal medium, was due to a change in the *trc* promoter DNA (Hawkins & Smith, 1991).

The results of the subcloning and expression studies indicate

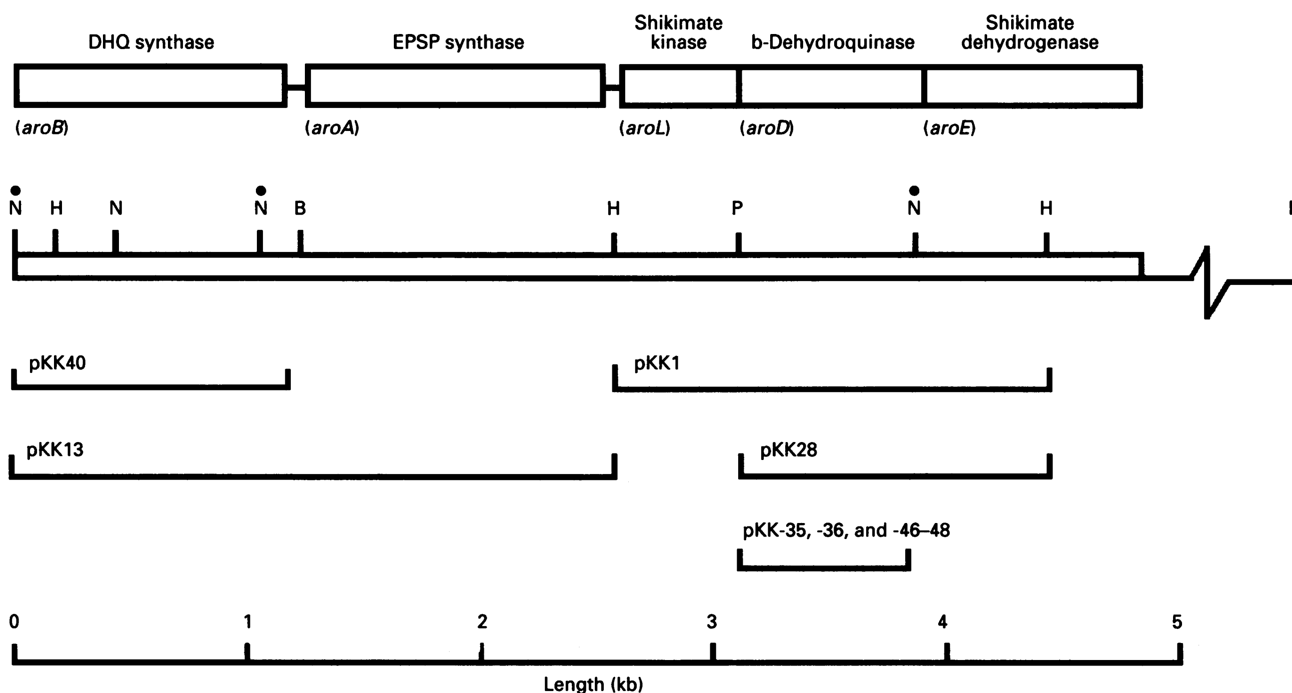


Fig. 1. Construction of recombinant plasmids

The *aromA* locus, AROM protein and associated enzyme activities are shown drawn to scale, with the equivalent designations for the monofunctional *E. coli* enzymes shown in parentheses. Restriction-endonuclease recognition sites for *Bam*HI (B), *Hind*III (H), *Nco*I (N) and *Pst*I (P) are shown. '●' over the first *Nco*I site indicates that this site was incorporated into the 30-mer oligonucleotide used to generate the PCR-amplified DHQ synthase domain. pKK13 was the template for the PCR amplification of the DHQ synthase domain and pKK28 the template for the PCR amplification of the 3-dehydroquinase domain. pKK28 was derived from pKK1 by digesting with *Pst*I and self-ligation after exposure of the cut ends of the DNA to T4 DNA polymerase. The *aromA* encoding DNA in plasmids pKK-1, -13, -28, -35, -36, -40, -43 and -46–48 were all ligated into suitably digested pKK233-2 plasmid DNA (Hawkins & Smith, 1991).

that each of the PCR products subcloned into the pKK233-2 expression vector encoded protein products that were able to fold and function enzymically in *E. coli*. Further, these results show that the *qutB*-encoded quinate dehydrogenase does have the shikimate dehydrogenase activity *in vivo* that was suggested by the similarity in amino acid sequence of the enzyme with the equivalent shikimate dehydrogenase domain in the AROM protein.

Measurement *in vitro* and purification of plasmid-encoded proteins

Having established that the subcloned PCR products encoded proteins that could fold and function in *E. coli*, we screened appropriate strains for IPTG-inducible overexpression by SDS/PAGE and enzyme assay as described in the Materials and methods section.

DHQ synthase domain. Three *aroB* mutant strains of *E. coli* containing recombinant plasmids supporting growth on unsupplemented minimal medium were screened for overproduction of the *arom* DHQ synthase domain. One strain, containing a plasmid we designate pKK40 (containing a translational stop codon at co-ordinate 1180), produced substantial amounts of protein that migrated in the position predicted for the AROM DHQ synthase domain (43 kDa). Plasmid pKK40 was then transformed into strain GKW38 ($aroB^-$), and an induction experiment was carried out as described in the Materials and methods section. Fig. 2(*a*) shows a photograph of the final Coomassie Brilliant Blue-stained gel, inspection of which confirms that the AROM DHQ synthase domain is being overproduced to the point that it constitutes around 30 % of the cell protein; this is the first report of the successful overproduction of an AROM protein domain. The AROM-containing DNA insert in plasmid pKK40 was subcloned into appropriate M13 vectors and subject to a DNA sequence analysis, which showed that the PCR product was of the correct co-ordinates and that the expected nucleotide changes introducing the translational stop codon had been successfully accomplished. Four $aroB^-$ mutant strains of *E. coli* containing recombinant plasmids encoding the DHQ synthase domain with a translational stop codon at co-ordinate 1303 (designated pKK41) were screened in an identical manner, but no overproducing strains were found.

A 10-litre culture of GLW38 harbouring plasmid pKK40 was subject to an 8 h induction in the presence of IPTG as described in the Materials and methods section, and the induced cells were harvested by centrifugation. The cells were resuspended in 400 ml of 50 mM-potassium phosphate, pH 6.6, containing 1.4 mM-β-mercaptoethanol and 10 μM-$ZnSO_4$ (buffer B). The cells were disrupted by sonicating them for 5 min bursts at an amplitude of 12 μm, with 2 min of cooling between each of six bursts of sonication, and were stirred on ice throughout the entire period. After sonication the cell suspension was clarified by centrifugation at 9000 rev./min ($r_{av.}$ 8.6 cm) in a Beckman JA14 rotor at 4 °C for 20 min. The clarified supernatant at 4 °C was adjusted to 30 % saturation with, and, after the mixture had been stirred at 4 °C for 1 h, the precipitate was collected by centrifugation at 9000 rev./min at 4 °C in a JA14 rotor for 30 min. After dialysis against buffer B, aliquots of the precipitate and supernatant were subjected to SDS/PAGE; most of the AROM DHQ synthase was found in the precipitated fraction. The 30 %-satd.-$(NH_4)_2SO_4$ precipitate of protein was then dissolved in a minimum volume of buffer B and subjected to ascending gel-filtration chromatography through a Sephacryl S300HR column of dimensions 75 cm × 2.5 cm, 5 ml fractions being collected at a flow rate of 1.2 ml·min^{-1}. Proteins in the

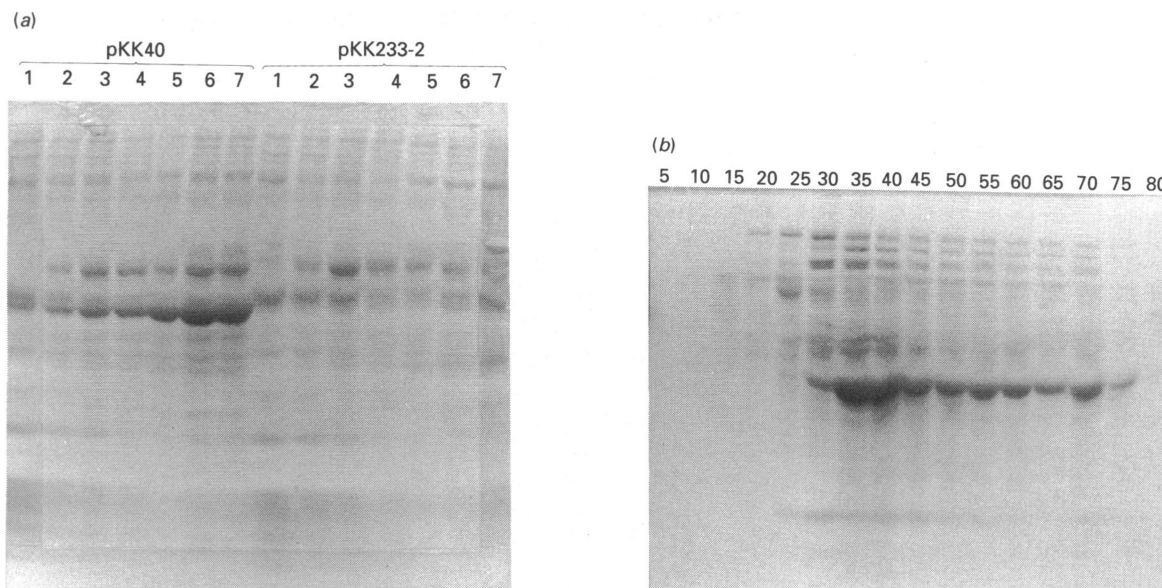


Fig. 2. Induction and partial purification of the AROM protein DHO synthase domain

(*a*) *E. coli* strain JM105 harbouring either plasmid pKK40 (experimental) or plasmid pKK233-2 (control) was subject to a 7 h-time-course induction as described in the Materials and methods section. Total cell protein from samples taken at 1–7 h is shown after separation by SDS/PAGE. (*b*) The gel shows the protein contained within fractions from the elution profile of a Sephacryl S300HR column after application of the 0–30 % $(NH_4)_2SO_4$ protein precipitate in buffer B from a clarified sonicated extract of *E. coli* strain GLW38 (*aroB*⁻, lacking DHQ synthase) harbouring plasmid pKK40. SDS/PAGE was standardized by electrophoresing protein molecular-mass markers comprising phosphorylase *b* (91 kDa), BSA (67 kDa), alcohol dehydrogenase (41 kDa) and chymotrypsinogen (25.1 kDa). The DHQ synthase domain migrated in a position consistent with its predicted molecular mass of 43 kDa.

elution profile were located by A_{280} measurements and analysed by subjecting 25 μl samples from every fifth fraction to SDS/PAGE. Fig. 2(*b*) shows the elution profile from the Sephacryl column for fractions 5–80, inspection of which shows that fractions 45–70 contain the AROM DHQ synthase domain at approx. 80 % purity. As the necessary substrates and reagents for the DHQ synthase assay are not commercially available, we were unable to measure the enzyme specific activity. Although we are unable to confirm by direct enzyme assay that the overproduced protein is the DHQ synthase domain, we believe that this is the case for the following reasons: (1) *aroB*⁻ (lacking DHQ synthase) mutant strains of *E. coli* fail to grow on minimal medium lacking aromatic-amino-acid supplements; (2) plasmid pKK40 allows *aroB*⁻ *E. coli* mutants to grow on minimal medium lacking aromatic-amino-acid supplements, whereas the unmodified expression vector, pKK233-2, alone does not; (3) plasmid pKK40 has been shown by directed sequence analysis to encode the nucleotide sequence of the *aromA* gene specifying DHQ synthase; and (4) plasmid pKK40, when harboured by the *lacI*^{q} strain JM105, specifies an IPTG-inducible protein of the molecular mass predicted for the DHQ synthase domain.

3-Dehydroquinase domain. Six strains each of *E. coli* containing recombinant pKK233-2 plasmids containing the 3-dehydroquinase-encoding sequence ending at co-ordinates 3859 (pKK35) or 3870 (pKK36), 3928 (pKK46), 3994 (pKK47), or 4075 (pKK48) were screened for IPTG-inducible protein overproduction. None of the strains screened showed any evidence of overproduction of the AROM 3-dehydroquinase domain (results not shown). *E. coli* strains harbouring plasmid pKK233-2 and each of the five recombinant plasmids (pKK-35, -36, -46–48) were subjected to an 8 h induction, and cell-free extracts were prepared, as described in the Materials and methods section. Preliminary 3-dehydroquinase assays demonstrated that a substantial non-linear non-dose-dependent activity was sometimes present in all cell-free extracts, obscuring the accurate measurement of enzyme specific activity. In order to obviate this problem, 0–30 %- and 30–50 %-satd.-$(NH_4)_2SO_4$ protein precipitates were recovered from the extracts and redissolved in buffer A. Most of the 3-dehydroquinase activity was found within the 30–50 %-satd.-$(NH_4)_2SO_4$ precipitate, and, after this partial purification, was found to give linear dose-dependent enzyme assays; however the negative control (*aroD*⁻ containing pKK233-2) still had a non-specific background activity. The data presented in Table 1 are from an induction experiment in which the 3-dehydroquinase levels were determined in *E. coli* strain SK3430 (*aroD*⁻, lacking 3-dehydroquinase) containing the plasmids pKK-35, -36, -46, -47 and -48 (experimental) or 233-2 (negative control). In this particular case the background non-specific activity was very low. Previously a recombinant pKK233-2 plasmid, designated pCLB, containing the AROM 3-dehydroquinase domain up to co-ordinate 3859, but in addition carrying a 'non-specific' tail of protein sequence HGCSQAWLFMRMREDFQPDTP derived by translating plasmid sequences, was described (Hawkins & Smith, 1991). Plasmid pCLB confers on the *aroD*⁻ (lacking 3-dehydroquinase) mutant strains of *E. coli* the ability to grow on unsupplemented minimal medium, but its presence does not lead to the production of enzyme activity that can be measured *in vitro*. Plasmid pKK35 specifies the same AROM 3-dehydroquinase domain, but lacking the plasmid-encoded peptide 'tail'; similarly plasmids pKK-36, and -46–48 specify 3-dehydroquinase domains without a 'tail' derived by translating plasmid DNA sequences, but which are 4, 19, 41 or 48 amino acids (of the correct AROM sequence) longer than the domain encoded by pKK35. We believe that the simplest interpretation of these data (see Table 1) is that we have, in effect, 'titrated' the end of the *C*-terminus of a monofunctional domain specifying 3-dehydroquinase to a position at or around amino acid number 1290 (see a corrected *aromA* sequence in Hawkins *et al.*, 1988). The specific activity of the 3-dehydroquinase encoded by pKK36 is approx. 8-fold higher than the typical 3-dehydroquinase specific

Table 1. 3-Dehydroquinase specific activity in *E. coli* strain SK3430 (*aroD*⁻; lacking 3-dehydroquinase) containing various plasmids encoding the *A. nidulans* AROM 3-dehydroquinase

Activity is shown in units/mg of protein and as a value relative to the non-specific background rate present in SK3430 containing the control plasmid unmodified pKK233-2. The *C*-terminal end of the enzyme encoded by pKK35 is designated '0' and the numbers in the column headed '*C*-Terminal extension' refer to the numbers of extra amino acids beyond this point specified at the *C*-terminus by each of the other four plasmids. Abbreviation: NA, not applicable.

Plasmid	3-Dehydroquinase activity (units/mg)	Relative activity	*C*-Terminal extension
KK35	0.0209	9.95	0
KK36	0.0471	22.43	4
KK46	0.0038	1.81	19
KK47	0.0036	1.71	41
KK48	0.0029	1.38	68
KK233-2	0.0021	1.0	NA

activity associated with the native AROM protein in *A. nidulans* (Lamb *et al.*, 1991*b*), indicating that we have achieved a modest level of overexpression. The 3-dehydroquinase activity encoded by plasmids pKK-35 and -36 does not decrease after incubation of a crude cell-free extract at 4 °C for 26 h, demonstrating that the enzyme activity persists long enough to allow purification from *E. coli*.

The amino acid sequence of the *E. coli* 3-dehydroquinase enzyme has recently been revised, with 24 amino acid differences from the original sequence and, in addition, it contains an extra 12 residues at the *C*-terminus (Duncan *et al.*, 1986; Chaudhuri *et al.*, 1991). When this revised sequence is aligned with the amino acid sequence of the *A. nidulans* AROM 3-dehydroquinase, the *C*-terminal amino acid of the bacterial sequence aligns with residue I^{1288} in the fungal sequence. Plasmid pKK36 specifies a fungal 3-dehydroquinase with a *C*-terminus at residue P^{1290}; this sequence is only two residues longer than the position of the *C*-terminus suggested by sequence alignment and specifies the AROM monofunctional 3-dehydroquinase with the highest specific activity. Taken together these observations strengthen the proposal that the *C*-terminus of the AROM 3-dehydroquinase domain is at, or close to, residue P^{1290}, and, consequently, indicates that there is no 'linker region', as previously suggested (Hawkins, 1987), between the 3-dehydroquinase and shikimate dehydrogenase domains in the native AROM protein (see Fig. 1).

The *qutB*-encoded quinate dehydrogenase. In all, 23 *E. coli aroE*⁻ (lacking shikimate dehydrogenase) strains harbouring recombinant pKK233-2 plasmids containing the *qutB* gene were screened by SDS/PAGE as described in the Materials and methods section to look for overproduction of quinate dehydrogenase. None of the strains screened showed any evidence of overproduction of quinate dehydrogenase. Two *aroE*⁻ mutant strains of *E. coli*, one containing the recombinant plasmids pKK42 (the original non-complementing plasmid) and one pKK43 (derived from pKK42 and identified by its ability to confer growth of *aroE* mutant strains of *E. coli* on non-supplemented minimal medium), were subjected to an 8 h induction in the presence of IPTG. Sonicated cell-free extracts were assayed for quinate and shikimate dehydrogenase activity, but none could be detected in either strain.

In conclusion we observe that the use of PCR technology in combination with a plasmid containing an IPTG-inducible promoter has allowed us to overproduce and partially purify the DHQ synthase domain of the pentafunctional AROM protein of *A. nidulans*. Using the same technology we have demonstrated that: (*a*) we can produce an active monofunctional 3-dehydroquinase domain of the AROM protein that has sufficient activity to be detected in an '*in vitro*' assay and to persist *in vitro* long enough to allow purification; (*b*) the *C*-terminus of the AROM 3-dehydroquinase domain is located at, or very near to, residue P^{1290}, and consequently to strongly suggest that there is no 'linker region' between the AROM 3-dehydroquinase and shikimate dehydrogenase domains in the native AROM protein; (*c*) the *qutB*-encoded quinate dehydrogenase has the shikimate dehydrogenase activity *in vivo* that was suggested by the strong amino acid sequence similarity between it and the AROM shikimate dehydrogenase domain.

This research was supported by Science and Engineering Research Council (S.E.R.C.) Grant GR/E1055.5 awarded to A.R.H. J.D.M. is in receipt of an S.E.R.C. Studentship. We thank Mrs. I. Stobbs for typing the manuscript.

REFERENCES

Anton, I. A., Duncan, K. & Coggins, J. R. (1987) J. Mol. Biol. **197**, 367–371

Beri, R., Whittington, H., Roberts, C. F. & Hawkins, A. R. (1987) Nucleic Acids Res. **19**, 7991–8001

Biggin, M. D., Gibson, T. S. & Hong, C. F. (1983) Proc. Natl. Acad. Sci. U.S.A. **86**, 3963–3965

Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thomson, K., Faber, H. E., Furlong, L., Grunwald, D. J., Keifer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1970) Science **196**, 161–169

Charles, I. G., Keyte, J. W., Brammar, W. J. & Hawkins, A. R. (1985) Nucleic Acids Res. **13**, 8119–8128

Charles, I. G., Keyte, J. W., Brammar, W. J., Smith, M. & Hawkins, A. R. (1986) Nucleic Acids Res. **14**, 2201–2213

Chaudhuri, S., Duncan, K., Graham, L. D. & Coggins, J. R. (1991) Biochem. J. **275**, 1–6

Coggins, J. R., Boocock, M. R., Chaudhuri, S., Lambert, J. M., Lumsden, J., Nimmo, G. A. & Smith, D. D. S. (1987) Methods Enzymol. **142**, 325–341

Da Silva, A. J. S., Whittington, H., Clements, J., Roberts, C. F., Hawkins, A. R. (1986) Biochem. J. **240**, 481–488

Duncan, K., Chaudhuri, S., Campbell, M. S. & Coggins, J. R. (1986) Biochem. J. **238**, 475–483

Garbe, T., Servos, S., Hawkins, A. R., Young, D., Dougan, G. & Charles, I. G. (1991) Mol. Gen. Genet. **228**, 385–392

Grant, S., Roberts, C. F., Lamb, H., Stout, M. & Hawkins, A. R. (1988) J. Gen. Microbiol. **134**, 347–358

Hawkins, A. R. (1987) Curr. Genet. **11**, 491–498

Hawkins, A. R. & Smith, M. (1991) Eur. J. Biochem. **196**, 717–724

Hawkins, A. R., Da Silva, A. J. S. & Roberts, C. F. (1984) J. Gen. Microbiol. **130**, 567–574

Hawkins, A. R., Da Silva, A. J. S. & Roberts, C. F. (1985) Curr. Genet. **9**, 305–311

Hawkins, A. R., Lamb, H. K., Smith, M., Keyte, J. W. & Roberts, C. F. (1988) Mol. Gen. Genet. **214**, 224–231

Hawkins, A. R., Lamb, H. K. & Roberts, C. F. (1992) Gene **110**, 109–114

Horowitz, N. H. (1965) in Evolving genes and proteins (Bryson, B. & Vogel, H. J., eds.), pp. 15–23, Academic Press, New York

Kinghorn, J. R. & Hawkins, A. R. (1982) Mol. Gen. Genet. **186**, 145–152

Laemmli, U. K. (1970) Nature (London) **227**, 680–685

Lamb, H. K., Hawkins, A. R., Smith, M., Harvey, I. J., Brown, J., Turner, G. & Roberts, C. F. (1990) Mol. Gen. Genet. **223**, 17–23

Lamb, H. K., Bagshaw, C. R. & Hawkins, A. R. (1991) Mol. Gen. Genet. **227**, 187–196

Lamb, H. K., van den Hombergh, J. P. T. W., Newton, G. H., Moore, J. D., Roberts, C. F. & Hawkins, A. R. (1992) Biochem. J. **284**, 181–187

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Mol. Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Matthes, W. D., Zenke, W. M., Grundstrom, T., Staub, A., Winterzerith, M. & Chambon, P. (1984) EMBO J. **3**, 801–805
Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Sproat, B. S. & Gait, M. J. (1985) Nucleic Acids Res. **13**, 2959–2977
Whittington, H., Grant, S., Roberts, C. F., Lamb, H. & Hawkins, A. R. (1987) Curr. Genet. **12**, 135–139
Willetts, N. S., Clark, A. J. & Low, B. (1969) J. Bacteriol. **97**, 244–249

Received 16 July 1991/27 November 1991; accepted 4 December 1991