Characterization of the 3-dehydroquinase domain of the pentafunctional AROM protein, and the quinate dehydrogenase from *Aspergillus nidulans*, and the overproduction of the type II 3-dehydroquinase from *Neurospora crassa*

Alastair R. HAWKINS,* Jonathan D. MOORE and A. Monisola ADEOKUN

Department of Biochemistry and Genetics, Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K.

The AROM protein of *Aspergillus nidulans* is a multidomain pentafunctional polypeptide that is active as a dimer and catalyses steps 2–6 in the prechorismate section of the shikimate pathway. The three C-terminal domains (including the type I 3-dehydroquinase) of the AROM protein are homologous with the *qutR*-encoded QUTR protein that represses transcription of the eight genes comprising the quinic acid utilization (*qut*) gene cluster, and the two N-terminal domains are homologous with the *qutA*-encoded QUTA protein that transcribes the *qut* genes. As part of a larger research programme designed to compare the structures of the three proteins and to probe the domain structure and interaction within each protein, we have overproduced and purified the 3-dehydroquinase domain of the AROM protein. Additionally we have overproduced and purified the *qutB*-

encoded quinate dehydrogenase and overproduced the *qa-2* encoded type II 3-dehydroquinase of *Neurospora crassa*. We report that the AROM 3-dehydroquinase domain has a monomeric native state, with an apparent k_{cat}/K_m ratio that is approx. 160-fold lower than the value for the native *N. crassa* AROM protein. The AROM protein 3-dehydroquinase domain is sensitive to inactivation by borohydride in the presence of the substrate 3-dehydroquinase, confirming that it is a typical type I 3-dehydroquinase. The purified quinate dehydrogenase is bifunctional, being able to metabolize shikimate as a substrate. The apparent K_m values for quinate (450 μ M), shikimate (1.7 mM) and NAD⁺ (150 μ M) are all similar to values reported for the *qa-3*-encoded enzyme from *N. crassa*.

INTRODUCTION

The pentafunctional AROM protein catalyses steps 2–6 in the prechorismate shikimate pathway converting 3-deoxy-D-*arabino*-heptulosonic acid-7-phosphate into 5-enolpyruvyl shikimate-3-phosphate. The AROM protein is present in fungi, yeast and *Euglena* and is encoded by a single gene, but, by contrast, the five genes encoding the equivalent enzymic activities tend to be dispersed in prokaryotes and plants (Bentley, 1990). There is no evidence for aggregation of these enzymes in prokaryotes, but plants possess a bifunctional protein consisting of the 3-dehydroquinase (DHQase; EC 4.2.1.10) and shikimate dehydrogenase (EC 1.1.1.25) activities equivalent to the two C-terminal domains of the AROM protein (Mousdale et al., 1987) indicating in this instance the presence of four genes. The two intermediates 3-dehydroquinase (DHQ) and dehydroshikimate (DHS) that are interconverted by the 3-dehydroquinase activity (EC 4.2.1.10) associated with the AROM protein are also intermediates in a related pathway responsible for the catabolism of quinate to protocatechuate (PCA) (Giles et al., 1967). The quinate pathway metabolites DHQ and DHS are interconverted by a second DHQase isoenzyme whose corresponding gene is subject to regulation at the level of transcriptional control by a distinct genetic mechanism (Giles et al., 1985; Grant et al., 1988). In *Aspergillus nidulans* the genes encoding the shikimate and quinate pathway DHQases have been isolated by molecular cloning and their nucleotide sequences determined (Charles et al., 1986; Da Silva et al., 1986). Comparison of the deduced protein sequences showed no similarity at the nucleotide or amino acid sequence

level, leading to the proposal that the genes encoding the two enzymes were phylogenetically distinct, potentially identifying a case of convergent evolution (Hawkins, 1987; Duncan et al., 1987). Subsequent work has confirmed this prediction and introduced a new classification dividing DHQases into two distinct classes (Kleanthous et al., 1992). Type I DHQases have monomer molecular masses of around 27 kDa, are metabolically active as dimers, and catalyse a *cis* abstraction of water from DHQ using an active-site lysine to form a Schiff base as part of the reaction mechanism (Schneider et al., 1993). Type II DHQases have subunit molecular masses in the range 16–18 kDa, are metabolically active as dodecamers and catalyse a *trans* abstraction of water from DHQ without using an active-site lysine to form a Schiff base as part of the reaction mechanism. Type I DHQases are almost exclusively biosynthetic in nature, whereas type II DHQases are commonly active in both the catabolic quinate pathway and the biosynthetic shikimate pathway (White et al., 1990; Garbe et al., 1991). Recently however a dual-function type II DHQase active in both the quinate and shikimate pathways has been characterized in *Amycolatopsis methanolica* (Euverink et al., 1992).

The AROM protein of *Neurospora crassa* has been the most extensively studied example of this protein where a number of interesting features, including channelling and catalytic facilitation, have been ascribed based on *in vitro* biochemical data (Gaertner et al., 1970; Welch and Gaertner, 1976). This work was subsequently challenged when the enzyme preparations were found to be proteolytically damaged and deficient in zinc atoms (Lambert et al., 1985). Subsequently the AROM protein has

Abbreviations used: DHQ, 3-dehydroquinase; DHS, dehydroshikimate; PCA, protocatechuate; IPTG, isopropyl thiogalactoside; DHQase, 3-dehydroquinase; GST, glutathione S-transferase; i.m.a.c., immobilized-metal-affinity chromatography.

* To whom correspondence should be sent.

been studied in *A. nidulans* by *in vivo* methods based on the principles of metabolic control analysis, where it has been shown that the AROM protein is extensively leaky (Lamb et al., 1992). In addition it has been shown that flux of the common metabolites DHQ and DHS can be directed within and between the two pathways by up-modulation of particular enzyme concentrations, allowing increased flux into the shikimate pathway (Lamb et al., 1991).

The three C-terminal domains of the AROM polypeptide (including the DHQase domain) shows sequence identity with the QUTR transcription repressor protein encoded by the *qutR* gene in *A. nidulans*. The QUTR protein represses transcription of the eight genes comprising the quinic acid utilization (*qut*) gene cluster, and is the first documented case of a eukaryotic repressor protein having sequence similarity to that of metabolic enzymes (Anton et al., 1987; Hawkins et al., 1992). Most recently, however, it has been reported that the two N-terminal domains of the AROM protein show sequence identity with the QUTA transcriptional activator protein encoding by the *qutA* gene. This is the first case where the genesis and evolutionary history of a pair of interacting eukaryotic transcriptional regulatory proteins has been established (Hawkins et al., 1993).

As part of a continuing programme of research designed to compare and probe the domain structure and function within the AROM and QUTR repressor proteins of *A. nidulans*, we have overproduced the DHQase domain of the AROM polypeptide in *Escherichia coli* and characterized the purified protein. We are also pursuing a study of metabolic flux within and between the quinate pathways and to this end we report the overproduction, purification and characterization of the *qutB*-encoded quinate dehydrogenase (EC 1.1.1.24). The purpose for this exercise is to allow us to determine the elasticity control coefficients associated with this enzyme and use these in calculations to determine the flux control coefficients of the three quinate-pathway enzymes (Lamb et al., 1991, 1992). The crystallization of the type I 3-dehydroquinase from *Salmonella typhi* has been reported (Boys et al., 1992), and we report here the overproduction of a type II DHQase (EC 4.2.1.10) from *Neurospora crassa*. This enzyme will be used in crystallization trials with type II DHQases from *A. nidulans* and *M. tuberculosis* as practical experience demonstrates that trials with a range of related enzymes increases the chances of success. Successful crystallization will facilitate a structural comparison of these two classes of enzymes.

MATERIALS AND METHODS

Plasmids, recombinant lambda and strains

The origins of the recombinant phage lambda and plasmids other than those constructed as part of the research described in this manuscript have been described previously (Charles et al., 1986; Moore and Hawkins, 1993). The *E. coli* strains used in this work have been described previously, as have methods for their growth and propagation (van den Hombergh et al., 1992; Moore and Hawkins, 1993; Studier et al., 1990).

Materials

Chemicals and solvents other than DHQ were of AnalaR or greater purity and purchased from local suppliers. DHQ was made from quinate and purified by chromatography as previously described (Grant et al., 1988; Grewe and Haendler, 1966). General molecular-biology reagents were from Pharmacia, or Gibco/BRL, *Taq* polymerase was from BCL Ltd., plasmid pGEX-2T was from Pharmacia and plasmid pRSETB was from

Invitrogen. Activated Sepharose 4B and chelating Sepharose FF were from Pharmacia, production of GSH-Sepharose 4B followed the manufacturer's instructions accompanying the plasmid pGEX-2T. GSH was from Pharmacia, quinate and NAD⁺ were from Sigma, and specific 30-mer oligonucleotides were purchased from J. Keyte, Department of Biochemistry, University of Nottingham, Nottingham, U.K.

Methods

Use of restriction endonucleases followed manufacturers' recommendations and amplification of target DNA sequences by means of the PCR was as previously described (van den Hombergh et al., 1992; Moore and Hawkins, 1993). All other routine molecular-biology protocols followed individual manufacturer's recommendations or were as described previously (Maniatis et al., 1982). SDS/PAGE was carried out according to the method of Laemmli (1970), using a 5%-acrylamide stacking gel and 10%-acrylamide separating gel, and proteins were made visible by staining with Coomassie Brilliant Blue. Small-scale (2 ml) growth and induction of fusion-protein synthesis in *E. coli* with IPTG was as previously described (Hawkins and Smith 1991; van den Hombergh et al., 1992). For large-scale growth and isopropyl thiogalactoside (IPTG) induction (100 ml and over), cells were grown at 30 °C with shaking to mid-exponential phase, when IPTG was added to a final concentration of 0.2 mg·ml⁻¹ and incubation continued for a further 2.5 h (pGEX-2T) or 3–4 h (pRSETB). *E. coli* cells were harvested by centrifugation and stored frozen at –18 °C until needed. Cells were lysed in appropriate buffers (see the Results section) by sonication while being incubated in a salt/ice bath at 0 °C; lysis was monitored by the appearance of a straw-coloured opalescent change in the resuspended cell suspension. Cell-free lysates were prepared by centrifugation at 22000 *g* for 20 min and decanting the soluble protein from the cell debris and insoluble proteins. Affinity chromatography was carried out in glass-fibre-plugged Pasteur pipettes using a 1 ml bed volume equilibrated with the appropriate buffer (see the Results section). The GSH-Sepharose 4B column matrix was directly usable; however, the chelating Sepharose matrix was charged by passing 10 ml of a saturated solution (at room temperature) of ZnSO₄ (in deionized water) through it, followed by a 20 ml wash with deionized water before equilibration with the appropriate buffer. DHQase, quinate dehydrogenase and shikimate dehydrogenase assays were carried out at 37 °C as previously described (van den Hombergh et al., 1992). *K_m* determinations for DHQase activity were carried out in 50 mM potassium phosphate buffer, pH 7.2, at 37 °C with DHQ concentrations ranging from 532 to 22.6 μM. *K_m* determinations for the substrates quinate, shikimate and NAD⁺ utilized by the *qutB*-encoded quinate dehydrogenase enzyme were carried out in 50 mM Tris/HCl, pH 8.6, at 37 °C with the concentrations of substrate used varying between 10 and 0.05 mM. The kinetic constants were obtained by computer analysis from Hanes plots ([S]/*v* against [S]) using linear regression. In the case of the *qutB*-encoded quinate/shikimate dehydrogenase, quadruplicate measurements were taken; in the case of the DHQase enzymes, duplicate measurements were taken, and all replicates were within 5%. The estimates of the apparent *K_m* values of the *qutB*-encoded quinate/shikimate dehydrogenase for quinate and shikimate were derived from experiments where the concentration of NAD⁺ was fixed at 2.5 mM. For the estimation of the *K_m* for NAD⁺ the quinate concentrations was fixed at 10 mM.

Borohydride reduction of the type I DHQase enzyme was carried out according to the following protocol. Two 100 μl

aliquots of enzyme (10 units) in 25 mM Tris, pH 7.5, were placed in 1.5 ml polypropylene Eppendorf tubes, and one was made 360 μ M with respect to DHQ. After thoroughly mixing, 2 μ l of a 20 mg $\cdot$ ml⁻¹ solution of NaBH₄ in 40 mM NaOH was added to each sample and the solutions mixed by gentle pipetting. The addition of the borohydride solution caused some frothing of the two solutions, which were then spun in a bench-top centrifuge to check for the presence of denatured protein, but none was observed. The total activities of the two enzyme solutions were measured by direct enzyme assay and compared with one another and with the starting value before the addition of borohydride.

RESULTS AND DISCUSSION

Overproduction of the type I DHQase domain of the pentafunctional AROM polypeptide in *E. coli*

In order to characterize the individual domains comprising the dimeric AROM protein, previous attempts have been made to overproduce the DHQase activity in *E. coli*; however, only an extremely modest success was achieved, namely an 8-fold elevation over the levels found in *A. nidulans* (Kinghorn and Hawkins, 1982; van den Hombergh et al., 1992). In order to attempt to elevate the level of overproduction and at the same time introduce a single-step affinity purification, the DNA sequence encoding the enzyme was amplified by the PCR and subcloned into the *E. coli* expression vectors pGEX-2T and pRSETB (see Table 1). The pGEX-2T plasmid utilizes the powerful IPTG-inducible *tac* promoter and produces fusion proteins with the 27 kDa glutathione S-transferase (GST) protein from *Schistosoma japonicum* fused to the N-terminal end of the target experimental protein under study. Soluble fusion proteins can then be purified from cell-free bacterial lysates under non-denaturing conditions by affinity chromatography using GSH immobilized on Sepharose 4B. The pRSETB plasmid utilizes the powerful T7 promoter and IPTG-inducible T7 polymerase to produce fusion proteins with a 4.47 kDa metal-chelating domain fused to the N-terminal end of the target experimental protein

under study. Soluble fusion proteins can then be purified from cell-free bacterial lysates under non-denaturing conditions using immobilized-metal-affinity chromatography (i.m.a.c.). The use of these fusion vectors had a second equally important purpose as they allowed us to detect and assay heterologous domain interactions through measurement of kinetic parameters with the different fusion proteins and isolated DHQase domain.

Recombinant plasmids of both types were rescued from ligation mixtures by transformation of *E. coli* strain SK3430 and selecting for ampicillin resistance and the appropriate constructs identified by restriction-endonuclease-digestion analysis. Strain SK3430 is an *aroD*⁻ mutant lacking the *E. coli* type I DHQase and, as *E. coli* has only a single (type I) DHQase, such strains have a consequent auxotrophic requirement for aromatic amino acids. In the case of the GST-DHQase fusion it was therefore possible to carry out a preliminary screen for enzyme activity by replica-plating bacteria containing pGEX-2T recombinant plasmids on to minimal medium lacking aromatic amino acids as previously described (Hawkins and Smith, 1991), to see if the fusion protein could complement the *aroD*⁻ lesion. The appropriate recombinant plasmids were able to complement the *aroD*⁻ mutation, allowing growth on non-supplemented minimal medium, and four separate constructs designated pXF1-4 were taken for further analysis to determine whether a fusion protein was being produced in substantial amounts. Isolates of *E. coli* strain SK3430 transformed with the four experimental plasmids pXF1-4 and control pGEX-2T alone were screened by small-scale SDS/PAGE for protein production as described in the Materials and methods section. All four experimental strains were found to substantially overproduce a protein of the size expected for the correct fusion protein (55.8 kDa) and were regrown and IPTG-induced in bulk cultures (100 ml) at 30 °C. The control strain containing pGEX-2T produced the GST protein of molecular mass 27 kDa and lacked the fusion protein, showing that the IPTG induction worked and that production of the fusion protein was specific to the control strains. Cell-free lysates of these bulk cultures of *E. coli* strain SK3430 carrying

Table 1 Recombinant plasmids generated in the course of this work, and typical specific activities in cell-free lysates of strains harbouring the individual plasmids

Enzyme specific activity is expressed as μ mol of product produced/min per mg of protein at 37 °C. Abbreviations: NA, not applicable; MC, metal-chelating domain. The oligonucleotide primers for the PCR amplification of DNA were 30-mers containing singly either *Bam*HI or *Eco*RI recognition sites for pGEX-2T subcloning, and *Nco*I or *Hind*III recognition sites for pRSETB, pKK233-2 and pTrc99a subcloning. In each case the first of the pair of restriction enzymes was for the primer at the 5' end of the target gene and the second at the 3' end of the target gene. The substrate for AROM DHQase was DNA from recombinant lambda ARH 8.1 (Charles et al., 1986). The PCR product for the pGEX-2T cloning was from co-ordinates 3106-3870 (see Charles et al., 1986; Hawkins et al., 1988) with an additional two amino acids, glycine and serine, at the N-terminus in place of the original glutamine and valine due to the introduction of a *Bam*HI site necessary for subcloning. When the GST/DHQase fusion protein is cleaved with thrombin, these two altered amino acids remain on the N-terminus of the DHQase domain. The primer for the 3' end of the AROM protein DHQase sequences to be amplified had an artificially introduced in-frame translational stop codon. Plasmid pRF1 was constructed by subcloning an *Nco*I/*Hind*III restriction fragment from plasmid pNUFC79 (Moore and Hawkins, 1993) into *Nco*I/*Hind*III-digested pRSETB. This plasmid contains *aromA* DNA from co-ordinates 3098-3870, but with the N-terminal leucine and glutamine residues replaced by methionine and alanine as a consequence of introducing an *Nco*I site necessary for cloning. The original cloning procedure used to generate pNUFC79 generated an in-frame artificial stop codon at codon 1291. The substrate for the PCR of the *qutB* gene was DNA from recombinant lambda λ Q1 (Hawkins et al., 1985) and the substrate for the PCR of the *qa-2* gene was plasmid pVK88 (Schweizer et al., 1981).

Plasmid no.	Vector	Affinity domain	Promoter	Enzyme specific activity		<i>E. coli</i> strain for complementation	<i>E. coli</i> strain for overproduction
				DHQase	Quinate dehydrogenase		
pXF1	pGEX-2T	GST	<i>tac</i>	1.48	NA	SK3430	SK3430
pXF2	pGEX-2T	GST	<i>tac</i>	0.97	NA	SK3430	SK3430
pXF3	pGEX-2T	GST	<i>tac</i>	1.65	NA	SK3430	SK3430
pXF4	pGEX-2T	GST	<i>tac</i>	2.47	NA	SK3430	SK3430
pRF1	pRSETB	MC	T7	0.41	NA	—	BL21(DE3)
pTR100	pTrc99a	NA	<i>trc</i>	NA	0.00	GLW55	GLW55
pRF2	pRSETB	MC	T7	NA	11.24	—	BL21(DE3)
pKK60	pKK233-2	NA	<i>trc</i>	5.3	NA	SK3430	SK3430
pTR101	pTrc99a	NA	<i>trc</i>	5.3	NA	SK3430	SK3430

experimental or control (pGEX2T) plasmids were assayed for type I DHQase activity, and all four experimental strains were found to substantially overproduce a soluble GST–DHQase fusion protein (see Table 1), to about 2×10^4 -fold the level found in the original expression experiments (Kinghorn and Hawkins, 1982). *E. coli* strain SK3430 does not express the T7 polymerase, so it is not possible to carry out the complementation test using recombinant pRSETB plasmids. In this case correctly constructed pRSETB recombinant plasmids were rescued from ligation mixtures in *E. coli* strain SK3430 and retransformed into strain BL21(DE3) (which has an IPTG-inducible T7 polymerase gene on a prophage) and screened by small-scale SDS/PAGE for recombinant-protein production. Fusion-protein overproduction was observed, but it was extremely variable (results not shown); one substantial overproducer (designated pRF1) was then grown in bulk (100 ml) at 30 °C, and cell-free lysates (in 50 mM potassium phosphate, pH 7.2) assayed for type I DHQase activity. Soluble DHQase activity was found in the cell-free lysate, but the levels were substantially lower than for the strains producing the GST–DHQase fusion proteins (see Table 1).

Purification of the DHQase fusion proteins

In order to purify the GST–DHQase fusion protein, approx. 0.5 g of cells (wet weight) were sonicated in 15 ml 10 mM potassium phosphate pH 7.2, 1 mM dithiothreitol (buffer 1) and, after preparation of cell-free extracts, the supernatant was assayed for DHQase activity. Initial attempts at purification used the manufacturer's recommended method of batch purification. The cell-free supernatant was made 0.15 M with respect to NaCl and then incubated with GSH-substituted Sepharose 4B at room temperature with gentle agitation for 15 min. This typically led to 60% of the DHQase activity adhering to the Sepharose beads. After washing with 50 ml of 20 mM potassium phosphate (pH 7.2)/0.15 M NaCl (buffer 2) and recovering the Sepharose beads by centrifugation, the fusion protein was eluted from the Sepharose 4B by resuspending the column matrix in 3 vol. of 50 mM Tris, pH 8.0, containing 5 mM GSH (buffer 3). The resuspended Sepharose was incubated at room temperature with gentle agitation, and the release of the fusion protein monitored with time by assay of Sepharose-free supernatant. Enzyme activity appeared within 2 min of adding the elution buffer, but the 3-DHQase enzyme was then rapidly and quantitatively denatured within 8 min to produce a white flocculent insoluble precipitate. This quantitative denaturation occurred each time batch purification was attempted and was a general phenomenon encountered with each of the different over-producing strains tested; this method of purification was therefore abandoned.

As an alternative purification procedure, columns of GSH-substituted Sepharose 4B were poured in glass Pasteur pipettes as described in the Materials and methods section and used to purify fusion proteins encoded by plasmids pXF3 and pXF4. The columns were equilibrated with buffer 2, and cell-free supernatants were applied to the column in buffer 1. Assay of the column flow-through fractions showed that typically 75% of the DHQase activity bound to the column. The columns were then washed with 10 bed vol. of buffer 2 and finally eluted with 4 bed vol. of 50 mM Tris buffer, pH 8.0, containing 5 mM GSH (buffer 3), 0.3 ml fractions being collected. Soluble activity was typically recovered in fractions 2–7, with a peak in fractions 3–5. Assay of the combined fractions showed that typically 85% of the bound activity was released in a stable soluble form, amounting to around 60% recovery of the original starting DHQase activity.

Figure 1 Representative SDS/PAGE gels showing the induction and purification of the AROM DHQase fusion proteins

(a) Boiled extracts of *E. coli* strain SK3430 induced with IPTG and containing: 1, plasmid pXF1; 2, plasmid pXF2; 3, plasmid pXF3; 4, plasmid pXF4 and 5, plasmid pGEX-2T. Lanes 6, 7 and 8 contain approx. 10, 20 and 40 µg of the affinity-purified GST–DHQase protein. These lanes were deliberately overloaded to detect the presence of any low concentration contaminating proteins. (b) Boiled extracts of *E. coli* BL21 (DE3) containing: 1, plasmid pRSETB IPTG-induced; 2, plasmid pRF1 uninduced; and 3, pRF1 IPTG induced. Lanes 4, 5 and 6 contain approx. 0.5, 1 and 2 µg of the affinity-purified metal-chelating domain–DHQase fusion protein. In each case standard molecular-mass markers of phosphorylase *b* (molecular mass 97 kDa), BSA (molecular mass 66 kDa), ovalbumin (molecular mass 45 kDa), carbonic anhydrase (molecular mass 29 kDa) and β -lactoglobulin (molecular mass 18.4 kDa) were co-electrophoresed in the unnumbered lanes to calibrate the gel. Each of the lanes containing whole-cell extracts contained approx. 15 µg of protein.

Figure 1(a) shows a representative SDS/PAGE gel showing induction of the fusion protein encoded by plasmid pXF3 and the state of purity (> 98%) after the single-step affinity purification. Typical purifications started at a specific activity of around 1.6 units per mg and yielded pure enzyme of approx. 12 units/mg. In one instance the DHQase domain was purified from the GST domain after digestion with thrombin, which cuts in a short linker region between the two domains. Digestion was carried out for 3 h at room temperature in buffer 3 containing 2.5 mM CaCl_2 , with 1 unit of thrombin per mg of fusion protein. Digestion was monitored by SDS/PAGE and DHQase enzyme assay to assess stability of the isolated DHQase domain, but no loss of DHQase activity could be detected over the digestion time course, demonstrating that the DHQase domain was stable when isolated. After digestion to completion the mixture of GST and DHQase domains was dialysed overnight against two changes of 25 mM Tris, pH 8.0, to remove the GSH. As a more rapid alternative to dialysis, spinning the digested protein successively

through four 1 ml Sephadex columns (poured in 1 ml plastic syringes) equilibrated with 25 mM Tris, pH 8.0, was found to be equally effective at removing the GSH. Following dialysis, or spun chromatography, the DHQase activity was recovered free of contaminating GST protein by re-chromatographing the mixture of proteins on a GSH-substituted Sepharose 4B column, the GST protein sticking to the column and the DHQase activity appearing in the unbound column flow-through.

The protein containing the DHQase domain fused to the metal-chelating domain was purified by i.m.a.c. Cell-free lysates in 50 mM potassium phosphate (pH 7.2)/0.5 M NaCl (buffer 4) were assayed for DHQase activity and chromatographed on a chelating Sepharose column that had previously been charged with Zn^{2+} (see the Materials and methods section) and equilibrated with buffer 4 following a 10 bed vol. wash with deionized water to remove unbound Zn^{2+} . Assay of the column flow-through showed that typically 95% of the soluble DHQase activity bound to the column, which was subsequently washed with 10 bed vol. of buffer 2. Preliminary experiments established that the DHQase activity was tightly bound to the column and that the great majority of the contaminating protein could be eluted from the column by washing with 0.2 M glycine in buffer 4. In order to elute the DHQase activity, therefore, a single-step elution of 2.0 M glycine in buffer 4 was applied to the column which had 1 bed vol. of buffer 4 as a reservoir on top of the Sepharose. This effectively gave a short increasing gradient of glycine concentration from 0 to 2.0 M, which eluted contaminating proteins before the 0.2 M glycine step eluted the DHQase activity. Figure 1(b) is a representative SDS/PAGE gel showing induction of the fusion protein and the state of purity (> 90%) after the single-step affinity purification.

Overproduction and purification of the *A. nidulans* *qutB*-encoded quinate/shikimate dehydrogenase

As part of a larger programme of research designed to determine the structures of the enzymes comprising the shikimate and quinate pathways, we have previously reported negative attempts to overproduce the *A. nidulans* *qutB*-encoded quinate/shikimate dehydrogenase in *E. coli* (van den Hombergh et al., 1992). This enzyme is active in the catabolic quinate utilization pathway and is related to, but distinct from, the shikimate dehydrogenase activity of the shikimate pathway AROM protein. *E. coli* has only one shikimate dehydrogenase (encoded by the *aroE* gene), and this is active in the shikimate pathway. In order to attempt overproduction of the enzyme in *E. coli* the *qutB* protein-coding DNA sequence was amplified by the PCR and subcloned into the expression vector pTrc99a (which contains the powerful IPTG-inducible *trc* promoter) and the fusion protein expression vector pRSETB. We chose to use two vector systems, as the pTrc99a vector frequently failed to drive overproduction of desired proteins from cloned DNA fragments, whereas in our hands the pRSETB system had a higher success rate. Appropriate constructs and screening for enzyme overproduction by SDS/PAGE was as described in the Materials and methods section and the preceding part of this section. Recombinant pTrc99a plasmids were screened for their ability to complement *E. coli* strain GLW55 (which is an *aroE*⁻ mutant lacking shikimate dehydrogenase) and allow growth on minimal medium lacking aromatic amino acids as described in the preceding part of this section. The recombinant pTrc99a plasmids (designated pTR100) were able to complement the *aroE*⁻ mutation allowing growth on non-supplemented minimal medium, but no protein overproduction could be detected by small-scale SDS/PAGE (results not shown) or direct enzyme assay of cell free lysates (see Table 1). As *E. coli*

Figure 2 Representative SDS/PAGE gels showing the induction and purification of the *qutB*-encoded quinate dehydrogenase fusion protein, and the induction of the *qa-2*-encoded type II DHQase

(a) Boiled extracts of *E. coli* containing: 1, pRSETB, induced with IPTG; 2, pRF2 uninduced; and 3, pRF2 induced with IPTG. Lanes 4, 5 and 6 contain approx. 1, 2 and 4 μg of the affinity-purified metal-chelating domain-quinate dehydrogenase fusion protein. (b) Boiled extracts of *E. coli* induced with IPTG and containing: 1, plasmid pTrc99a; 2, plasmid pTR101; 3, plasmid pKK233-2; and 4, plasmid pKK60. The molecular-mass standards electrophoresed in the unnumbered lanes are described in the legend to Figure 1, and each of the lanes showing whole-cell extracts contained approx. 15 μg of protein.

strain GLW55 (*aroE*⁻) does not express the T7 polymerase, the *in vivo* complementation test could not be carried out using the pRSETB recombinant plasmids. Appropriate recombinant pRSETB plasmids were screened in *E. coli* strain BL21 (DE3) for their ability to direct fusion-protein overproduction using small-scale SDS/PAGE, and in this case substantial overproduction was observed. Several strains substantially overproducing a fusion protein of the expected molecular mass were subsequently grown in bulk (100 ml) at 30 °C and induced with IPTG. Assay of a cell-free lysate of one of these strains harbouring a plasmid designated pRF2 showed that substantial amounts of soluble quinate dehydrogenase were produced (see Table 1). The metal-chelating domain-quinate dehydrogenase fusion protein was purified by i.m.a.c. exactly as described for the DHQase fusion protein described in the preceding part of this section. Figure 2(a) is a representative SDS/PAGE gel showing induction of the fusion protein and the state of purity (> 98%) after the single-step affinity purification. For reasons that are not immediately

Table 2 Probable subunit structure and apparent K_m values of the proteins purified in the course of this work

Abbreviation used: MC, metal-chelating domain. Unless otherwise indicated, K_m measurements were carried out at 37 °C. K_m determinations were carried out on GST–DHQase fusion proteins encoded by plasmids pXF3 and 4 and on the purified DHQase domain above after cleavage from the fusion protein with thrombin. The values in square brackets next to the K_m values for quinate and shikimate are the apparent K_m values reported for the *qa-3*-encoded *N. crassa* enzyme (Barea and Giles, 1978).

Protein	Subunit molecular mass from DNA sequence (kDa)	Native molecular mass	Probable subunit structure	Substrate	K_m
GST–DHQase fusion protein	55.8	182	Tri/Tetramer	DHQ	160 μ M (pXF3) 200 μ M (pXF4)
MC–DHQase fusion protein	32.7	34	Monomer	DHQ	170 μ M
AROM protein DHQase	28.2	29	Monomer	DHQ	250 μ M (pXF3) 270 μ M (pXF4)
GST–quinate dehydrogenase fusion protein	40.58	57.5	Monomer	Quinate	450 μ M [370 μ M]
				Shikimate	1.7 mM [1.18 mM]
				NAD ⁺	150 μ M
<i>S. typhi</i> DHQase	27.7	43.7	Dimer	DHQ	15 μ M (25 °C); 25 μ M (37 °C)

obvious, *E. coli* strain BL21(DE3) harbouring pRF2 produced the quinate dehydrogenase fusion protein at high constitutive levels; the addition of IPTG did not confer any increase in the enzyme specific activity (see Table 1).

Overproduction of the *N. crassa qa-2*-encoded type II DHQase

Again, as part of a larger programme of research designed to determine the structures of the enzymes comprising the shikimate and quinate pathways, we have previously reported attempts to overproduce the *qa-2*-encoded type II DHQase in *E. coli* (Hawkins et al., 1982). These first attempts achieved only very low levels of production and did not have the benefit of current *E. coli* expression plasmids. The protein-coding region of the *qa-2* gene was amplified by the PCR and subcloned into the *E. coli* expression vectors pKK233-2 and pTrc99a, both of which have the powerful IPTG-inducible *trc* promoter. Plasmid pTrc99a differs from pKK233-2 in that it has a bigger multi-cloning site, and has a plasmid borne *lacIⁿ* gene allowing more controlled gene expression. We have found previously that some genes direct greater levels of protein overproduction depending on which of the two plasmids are used as the vector, therefore we used both to obviate any potential strong difference in overproduction. Appropriate constructs of both plasmids were rescued from ligation mixtures in *E. coli* strain SK3430 (*aroD*[−]) and identified by their ability to complement the *aroD*[−] mutation and allow growth on non-supplemented minimal medium, as described above. Appropriate recombinant plasmids were screened in *E. coli* for their ability to direct DHQase overproduction using small-scale SDS/PAGE.

Representative strains containing either of the two types of recombinant plasmid (pKK60 and pTR101) were grown in bulk (100 ml) at 30 °C and, and IPTG induction, cell-free lysates were assayed for DHQase activity (see Table 1). Figure 2(b) is a representative SDS/PAGE gel showing induction of the type II DHQase enzyme. Inspection of Table 1 and Figure 2(b) shows that both recombinant plasmid types directed type II DHQase overproduction to the same levels.

Physical and kinetic characterization of the purified proteins

Native-molecular-mass determinations

The native molecular masses of the GST–DHQase fusion, the metal-chelating-domain–DHQase fusion, the type I DHQase

cleaved from the GST–DHQase fusion, and the metal-chelating-domain–quinate dehydrogenase fusion proteins were estimated by sieving chromatography on a column of dimensions 115 cm × 1.2 cm using Sephacryl S-300 in 50 mM potassium phosphate, pH 7.2. The column was calibrated by measuring the elution volumes of Blue Dextran (molecular mass 2×10^6 Da), alcohol dehydrogenase (molecular mass 150 kDa), BSA (molecular mass 66 kDa), carbonic anhydrase (molecular mass 29 kDa) and cytochrome *c* (molecular mass 12.4 kDa) and using these to construct a standard curve which was then used to estimate the molecular mass of each of the purified proteins. Table 2 summarizes the estimated native molecular masses and gives the most probable subunit structure for each protein based on the known subunit molecular mass from the DNA sequence. The molecular mass of the type I DHQase from *Salmonella typhi* was also determined as a control, but its elution volume was not included in the determination of the standard curve. Inspection of Table 2 leads to the following general conclusions: (1) the native molecular mass of the isolated type I DHQase domain of the AROM protein is consistent with it being monomeric, whereas the estimated molecular mass of the control type I DHQase from *S. typhi* is consistent with it being dimeric, as has been previously demonstrated (Moore et al., 1993). This is of potentially great significance, as most type I DHQases have a native dimer structure, the only exception being the enzyme from *Pisum sativum* (pea), which is monomeric and forms a bifunctional enzyme with shikimate dehydrogenase (Mousdale et al., 1987). The GST–DHQase fusion, however, gives an apparent molecular mass consistent with a higher-order structure consistent with either a trimer or tetramer, and is probably caused by an aggregation effect due to the presence of the GST domain. (2) The native molecular mass of the quinate dehydrogenase is consistent with it being monomeric, which is in agreement with the finding for the equivalent enzyme from *N. crassa*. The molecular mass estimated from both SDS/PAGE and Sephacryl chromatography is higher than that predicted from the known DNA sequences, and this anomalous migration is probably due to the presence of a metal-chelating domain on the N-terminal end of the quinate dehydrogenase enzyme.

Borohydride reduction of the type I DHQase

In order to help confirm the prediction that the DHQase of the AROM protein is a type I enzyme, borohydride reduction in the

presence or absence of the substrate DHQ was carried out as described in the Materials and methods section. The activity of the GST-DHQase fusion protein was reduced by 82% by borohydride in the presence of DHQ, but was unaffected by borohydride alone, confirming that the AROM protein DHQase is a typical type I enzyme.

Determination of K_m values

The K_m values for the various substrates of the two type I DHQase fusion proteins, the isolated AROM type I DHQase domain, a control type I DHQase from *S. typhi*, and the quinate dehydrogenase fusion protein were determined as described in the Materials and methods section, and the values obtained are shown in Table 2. Inspection of Table 2 reveals the following points. (1) The K_m value for the *S. typhi* DHQase at 25 °C is essentially identical with the value previously reported, strongly suggesting that the other apparent K_m values reported in Table 2 are representative. (2) The apparent K_m values for both DHQase fusion proteins and the isolated DHQase show small differences; all, however, are substantially higher, with an apparent value of approximately 200 μ M. During the determination of K_m values for the AROM DHQase either monofunctional or as a fusion protein, a significant deviation from Michaelis-Menten kinetics was apparent at high (> 400 μ M) DHQase concentrations. At these high concentrations the activity was reproducibly enhanced in different preparations of the enzyme, possibly indicating that the enzyme adopted some higher-order structure with altered kinetic parameters under these conditions. The fact that all three K_m values are similar strongly suggests that once this DHQase domain has folded, it is stable and is not significantly influenced by the presence or absence of a heterologous protein domain attached to its N-terminus. The apparent K_m values for DHQ reported here (see Table 2) are the highest for any type I DHQase enzyme and are approx. 40-fold higher than the K_m value (5 μ M) for the equivalent enzyme from *N. crassa* when it is part of the pentafunctional AROM protein (Chaudhuri et al., 1986). The k_{cat}/K_m value (at 37 °C) for the isolated DHQase domain was calculated to be $2.5 \times 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ from the 40-fold-increased K_m and an estimate of catalytic centre activity of 5 s^{-1} . Comparison with a corresponding value of $3.9 \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$ (at 25 °C) for the DHQase activity of the intact pentafunctional AROM protein of *N. crassa* suggests that the isolated domain is a markedly less efficient catalyst. This could be due to the fact that the AROM DHQase domain, when isolated from the rest of the dimeric AROM protein, is apparently a monomer. The ideal comparison with the DHQase activity of the *A. nidulans* native AROM protein is not possible at present as this enzyme has not been purified to homogeneity.

These data show that, when folded, the isolated DHQase domain of the *A. nidulans* AROM protein is stable but has a different native subunit structure and an elevated K_m for its substrate DHQase. These observations strongly support an emerging view that the individual domains comprising the pentafunctional AROM polypeptide interact with one another at a gross tertiary level when part of the native dimeric AROM protein (Hawkins and Smith, 1991; Moore and Hawkins, 1993). The K_m values for the *qutB*-encoded enzyme substrates quinate, shikimate and NAD⁺ and all very similar to the values for the equivalent *qa-3*-encoded enzyme of *N. crassa*. The observations, when combined with the data derived from the type I DHQase

domain, strongly suggest that fusing a heterologous domain, allowing affinity purification, to the N-terminus of a target protein does not prevent folding or radically alter the affinity of the enzyme for its substrate, and does allow overproduction of proteins previously unstable in *E. coli*.

In conclusion these data show that the use of 'affinity tagging' as an aid to bulk affinity purification and probing domain interaction has an important role to play in the study of shikimate- and quinate-pathway enzymes.

REFERENCES

- Anton, I. A., Duncan, K. and Coggins, J. R. (1987) *J. Mol. Biol.* **197**, 367–371
- Barea, J. L. and Giles, N. H. (1978) *Biochim. Biophys. Acta* **524**, 1–14
- Bentley, R. (1990) *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384
- Boys, C. W. G., Bury, S. M., Sawyer, L., Moore, J. D., Charles, I. G., Hawkins, A. R., Deka, R., Kleanthous, C. and Coggins, J. R. (1992) *J. Mol. Biol.* **227**, 352–355
- Charles, I. G., Keyte, J. W., Brammar, W. J., Smith, M. and Hawkins, A. R. (1986) *Nucleic Acids Res.* **14**, 2201–2213
- Chaudhuri, S., Lambert, J. M., McColl, I. A. and Coggins, J. R. (1986) *Biochem. J.* **239**, 699–704
- Da Silva, A. J. S., Whittington, H., Clements, J., Roberts, C. F. and Hawkins, A. R. (1986) *Biochem. J.* **240**, 481–488
- Duncan, K., Edwards, R. M. and Coggins, J. R. (1987) *Biochem. J.* **246**, 375–386
- Euverink, G. J. W., Hessels, G. J., Vrijbloed, J. W., Coggins, J. R. and Dijkhuizen, L. (1992) *J. Gen. Microbiol.* **138**, 2449–2457
- Gaertner, F. H., Ericson, M. C. and De Mogs, J. A. (1970) *J. Biol. Chem.* **245**, 595–600
- Garbe, T., Servos, S., Hawkins, A. R., Dimitriadis, G., Young, D. and Charles, I. G. (1991) *Mol. Gen. Genet.* **228**, 385–392
- Giles, N. H., Partridge, C. W. H., Ahmed, S. I. and Case, M. E. (1967) *Proc. Natl. Acad. Sci. U.S.A.* **58**, 1930–1937
- Giles, N. H., Case, M. E., Baum, J., Geever, R., Huiet, L., Patel, V. B. and Tyler, B. (1985) *Microbiol. Rev.* **49**, 338–358
- Grant, S., Roberts, C. F., Lamb, H. K., Stout, M. and Hawkins, A. R. (1988) *J. Gen. Microbiol.* **134**, 347–358
- Grewe, R. and Haendler, H. (1966) *Biochem. Prep.* **11**, 21–26
- Hawkins, A. R. (1987) *Curr. Genet.* **11**, 491–498
- Hawkins, A. R. and Smith, M. (1991) *Eur. J. Biochem.* **196**, 717–724
- Hawkins, A. R., Reinert, W. R. and Giles, N. H. (1982) *Biochem. J.* **203**, 769–773
- Hawkins, A. R., Da Silva, A. J. S. and Roberts, C. F. (1985) *Curr. Genet.* **9**, 305–311
- Hawkins, A. R., Lamb, H. K., Smith, M., Keyte, J. W. and Roberts, C. F. (1988) *Mol. Gen. Genet.* **214**, 224–231
- Hawkins, A. R., Lamb, H. K. and Roberts, C. F. (1992) *Gene* **110**, 109–114
- Hawkins, A. R., Lamb, H. K., Moore, J. D. and Roberts, C. F. (1993) *Gene*, in the press
- Kinghorn, J. R. and Hawkins, A. R. (1982) *Mol. Gen. Genet.* **186**, 145–152
- Kleanthous, C., Deka, R., Davis, K., Kelly, S. M., Cooper, A., Harding, S. E., Price, C., Hawkins, A. R. and Coggins, J. R. (1992) *Biochem. J.* **282**, 687–695
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Lamb, H. K., Bagshaw, C. R. and 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. and Hawkins, A. R. (1992) *Biochem. J.* **284**, 181–187
- Lambert, J. M., Boocock, M. R. and Coggins, J. R. (1985) *Biochem. J.* **226**, 817–829
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Moore, J. D. and Hawkins, A. R. (1993) *Mol. Gen. Genet.*, in the press
- Moore, J. D., Hawkins, A. R., Charles, I. G., Deka, R., Coggins, R., Cooper, A., Kelly, S. M. and Price, N. C. (1993) *Biochem. J.* **295**, 277–285
- Mousdale, D. M., Campbell, M. S. and Coggins, J. R. (1987) *Phytochemistry* **26**, 2665–2670
- Schneider, A., Harris, J., Kleanthous, C., Coggins, J. R., Hawkins, A. R. and Abell, C. (1993) *Bio-Org. Med. Chem. Lett.* **3**, 1399–1402
- Schweizer, M., Case, M. E., Dykstra, C. C., Giles, N. H. and Kushner, S. R. (1981) *Gene* **14**, 24–32
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
- van den Hombergh, J. P. T. W., Moore, J. D., Charles, I. G. and Hawkins, A. R. (1992) *Biochem. J.* **284**, 861–867
- Welch, G. R. and Gaertner, F. H. (1976) *Arch. Biochem. Biophys.* **172**, 476–489
- White, P. J., Young, I., Hunter, I. S., Nimmo, H. G. and Coggins, J. R. (1990) *Biochem. J.* **165**, 735–738