Review Article

The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects

ALASTAIR R. HAWKINS, HEATHER K. LAMB, JONATHAN D. MOORE, IAN G. CHARLES and CLIVE F. ROBERTS

1Department of Biochemistry and Genetics, Catherine Cookson Building, New Medical School, University of Newcastle upon Tyne, Newcastle NE2 4HH, UK
2Department of Cell Biology, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK
3Department of Genetics, Adrian Building, University of Leicester, Leicester LE1 7RH, UK

Purpose of the review

The detailed chemistry and enzymology of the shikimate pathway have been extensively discussed in recent excellent reviews (Pittard, 1987; Bentley, 1990; Braus, 1991). This short review will therefore focus (a) on the web of interconnecting evolutionary and structural relationships within and between the enzymes and regulatory proteins that constitute the pre-chorismate (shikimate) and the related quinate pathways, and (b) the possibility of pathway engineering to divert major new carbon flux from the quinate pathway into the shikimate pathway.

What is the shikimate pathway?

The shikimate pathway includes the production of chorismate, which is the common precursor for the three aromatic amino acids (Trp, Tyr and Phe) and para-aminobenzoic acid (see Fig. 1). This pathway is present in those prokaryotes, microbial eukaryotes and plants studied; however, it is absent in those animal species investigated, making it a potential target for medical and biotechnological interest and potential targets for pathway engineering. Other branches of the pathway and references to them are contained within the review of Bentley (1990).

The pre-chorismate shikimate pathway has two metabolites (3-dehydroquinate, DHQ; dehydroshikimate, DHS) in common with the quinate pathway that is responsible for the catabolism of quinate to protocatechuate (PCA) in many fungi and some bacteria (Tresguerres et al., 1970a,b; Ingledew et al., 1971; Grund et al., 1990) (see Fig. 1). Quinate is an abundant carbon source for many fungi comprising around 10% by weight of decaying leaf litter. The three enzymes necessary for the catabolism of quinate to PCA are induced by the presence of quinate, but this induction is subject to carbon catabolite repression when quinate is supplied as a mixture with other preferred carbon sources such as glucose. The induction of the production of the quinate pathway enzymes is therefore highly regulated and the molecular mechanisms involved have been intensively studied by genetic and molecular techniques in the filamentous fungi Neurospora crassa and Aspergillus nidulans (Giles et al., 1985; Grant et al., 1988). The metabolites DHQ and DHS, however, are interconverted by two entirely separate 3-dehydroquinase isoenzymes that have different phylogenetic origins and
distinct genetic mechanisms regulating the corresponding gene expression (see below).

Pathway organization

The enzymes catalysing the pre-chorismate section of the shikimate pathway in bacteria are monofunctional with no evidence of in vivo association and are encoded by genes that tend to be dispersed throughout the genome (Pittard, 1987). This trend is not absolute, as at least in Mycobacterium tuberculosis, Amycolatopsis methanolica and Staphylococcus aureus the genes encoding DHQ synthase and 3-dehydroquinase are closely linked (Garbe et al., 1991; Euverink et al., 1992; O'Connell et al., 1993). There is evidence that plants in general have the 3-dehydroquinase and shikimate dehydrogenase enzymic activities fused into a single bifunctional protein (Koshiba, 1978; Mousdale et al., 1987). In many microbial eukaryotes, the enzymes catalysing steps two to six are fused into a single polypeptide (the AROM protein) that is active as a dimer of identical subunits (Smith & Coggins, 1983). The genes encoding these proteins in A. nidulans and Saccharomyces cerevisiae have been isolated by molecular cloning and their nucleotide sequences determined (Hawkins et al., 1985; Charles et al., 1986; Hawkins, 1987; Duncan et al., 1987). Comparison of the predicted amino acid sequences encoded by these genes with those of the corresponding genes from an array of bacterial species strongly suggests that the genes are related to common ancestors and that the genes encoding the pentafunctional proteins evolved by multiple gene fusions (Hawkins, 1987; Duncan et al., 1987). Amino acid sequence comparisons between the 3-dehydroquinase isoenzymes active in the shikimate (biosynthetic) and quinate (catabolic) pathways of filamentous fungi strongly suggested that these isoenzymes are phylogenetically distinct and have arisen by convergent evolution (Da Silva et al., 1986). Most recent biochemical analyses of the shikimate pathway 3-dehydroquinases from Escherichia coli and Salmonella typhi and the quinate pathway 3-dehydroquinase from A. nidulans, have confirmed that they are representatives of two distinct classes of 3-dehydroquinase (Kleanthous et al., 1992; Moore et al., 1993). Type I 3-dehydroquinases are dimeric enzymes with a subunit $M_i$ of around 27000, whereas type II 3-dehydroquinases are dodecamers with a subunit $M_i$ of 16–18000. Type I enzymes involve Schiff base formation as part of the reaction mechanism and abstract water from DHQ in cis, whereas type II enzymes do not form a Schiff base and abstract water from DHQ in trans, the opposite stereocemical outcome from type I enzymes, further strengthening the convergent evol-
Hepatitis low level similarity with a DNA polymerase from known as strains mutant in these genes have not been pathways have been most extensively characterized in A. nidulans. The mutant phenotype of the protein, the product of the gene cluster is positively mediated by the QUTR protein, the product of the qutA gene, and negatively mediated by the QUTR protein, the product of the qutR gene (Grant et al., 1988). In the context of this review, the most important feature of the regulation is that the QUTR protein is proposed to repress transcription by binding to the activator protein stoichiometrically and to negate its repressing function by being able to recognize and bind quinate, DHQ and DHS (Hawkins et al., 1992, 1988a).

The shikimate and quinate pathway enzymes and regulatory proteins are based around a small group of conserved structural modules that have evolved pathway specific functions

Amino acid sequence comparisons have revealed a wealth of information showing that the enzymes and regulatory proteins comprising the shikimate and quinate pathways share a long and complex history (see Fig. 2). The bifunctional qutB encoded quinate pathway quinate/shikimate dehydrogenase is related to a common ancestor of the prokaryotic shikimate pathway shikimate dehydrogenases and to the fungal shikimate dehydrogenases found in the pentafucntional AROM protein (Hawkins et al., 1988). Unexpectedly however, the amino acid sequences of the qa-1S (N. crassa) and qutR (A. nidulans) encoded fungal quinate pathway transcription repressor proteins are similar throughout their entire length to the extreme C-terminus (amino acids 760–863 in A. nidulans) of the EPSP synthase domain, and the three complete C-terminal domains shikimate kinase, 3-dehydroquinase and shikimate dehydrogenase of the pentafunctional AROM proteins (amino acids 864–1063 in A. nidulans) from fungi (Anton et al., 1987; Hawkins et al., 1992). Indeed, one basic structural motif has formed the basis for the shikimate pathway aroM encoded shikimate dehydrogenase, the quinate pathway qutB encoded quinate/shikimate dehydrogenase and approximately one-third of the quinate pathway qutR encoded repressor protein that controls the expression of the qutB gene (see Fig. 2). The domain of the QUTR protein equivalent to the shikimate kinase domain of the AROM protein is of particular interest, as it contains the motif GX₆GKS/T (where X = any amino acid) which is found in purine nucleotide binding proteins (Walker et al., 1982). This observation opens the possibility that regulation of transcription of the qut gene cluster may involve phosphorylation events (see below), a view strengthened by the close similarity of the qutG encoded QUTG protein to bovine myo-inositol monophosphatase (Lamb et al., 1990). In addition, these sequence alignments provide a potential molecular explanation for the genetically based assertion that the quinate pathway qutR encoded transcription repressor protein can recognize and bind quinate, DHQ and DHS, as these metabolites or closely related compounds can be metabolized by the three C-terminal AROM protein domains. The simplest explanation is that the repressor proteins have retained the ability to recognize and bind these
enzymes can recognize and negate the activity of a pathway specific genes (Hawkins et al., 1993). This hypothesis requires that the two N-terminal domains of the AROM polypeptide fold to give a stable protein, and it has recently been demonstrated that the two N-terminal domains (DHQ synthase and EPSP synthase) of the AROM polypeptide are able to fold and function independently of the three C-terminal domains (Hawkins & Smith, 1991; Moore & Hawkins, 1993). This observation therefore gives added credence to the proposal that the activator protein is homologous to this functionally independent part of the AROM polypeptide. The overall sequence similarity of the A. nidulans QUTA protein compared with the AROM protein is 23%, with localized regions as high as 35%; this situation is similar to the sequence similarity of the repressor protein with the three C-terminal domains of the AROM protein.

A speculative model for the molecular control of transcription of the qut gene cluster

This sequence similarity suggests a possible mechanism for how the repressor protein is able to recognize and bind to the activator. Early pioneering genetic studies demonstrated that the native AROM protein in N. crassa is a dimer of identical subunits, and that the dimerization metabolites, but have lost enzymic activity with respect to them. These observations about the QUTR and equivalent N. crassa QA-1S proteins, whilst providing a potential basis to explain the phenotypes of qa-1 mutants, also raise the fundamental question of how a protein that is related to three shikimate pathway enzymes can recognize and negate the activity of a eukaryotic transcriptional activator protein by binding to it. Most remarkably, it has recently been reported that the qutA encoded quinate pathway activator protein is homologous with the two amino terminal domains of the pentafunctional AROM protein, DHQ synthase and EPSP synthase (Hawkins et al., 1993a). The fundamental change is the insertion into the DHQ synthase like domain of an amino acid sequence motif that is able to co-ordinate zinc atoms and form a putative zinc binuclear cluster domain, thereby facilitating DNA binding (see Fig. 2). Immediately adjacent to the C-terminus of the zinc binuclear cluster domain is a specificity region which acts to direct binding to defined upstream activator sequences in the promoters of quinate pathway specific genes (Hawkins et al., 1988). This observation therefore gives added credence to the proposal that the activator protein is homologous to this functionally independent part of the AROM polypeptide. The overall sequence similarity of the A. nidulans QUTA protein compared with the AROM protein is 23%, with localized regions as high as 35%; this situation is similar to the sequence similarity of the repressor protein with the three C-terminal domains of the AROM protein.

---

Fig. 2. Modular structure of the enzymes and regulatory proteins comprising the quinate and pre-chorismate shikimate pathways in A. nidulans. The boxes denoted are designate genes from bacteria encoding monofunctional shikimate pathway enzymes: ar0A, E. coli EPSP synthase (EPSPS); ar0B, E. coli DHQ synthase (DHS); ar0D, E. coli type I 3-dehydroquinase (DQ); ar0E, E. coli shikimate dehydrogenase (SDH); ar0L, E. coli shikimate kinase (SK); and ar0Q, M. tuberculosis type II 3-dehydroquinase (DQII). AROM designates the modular structure of the AROM protein of A. nidulans which is specified by the ar0A gene that arose by the fusion of the bacterial ar0A, B, D, E and L genes. The boxes denoting the genes of the qut pathway are: qutA, activator; qutB, quinate/shikimate dehydrogenase (QDH); qutD, permease; qutE, type II 3-dehydroquinase (DQII); qutG, possibly a phosphatase; and qutR, repressor. The genes encoding the quinate pathway activator and repressor proteins are proposed to have arisen by duplication of the ar0A gene followed by cleavage in the DNA sequence specifying the C-terminus of the EPSP synthase domain (ar0A equivalent). The ar0A sequence encoding the DHQ synthase domain in the duplicated copy is proposed to have undergone an event in which the DNA sequence encoding a pre-formed zinc binuclear cluster motif (designated Zn) has become inserted into its N-terminus. The box denoted ar0Q designates the type II 3-dehydroquinase from M. tuberculosis. The arrows between boxes are there to indicate that genes denoted by the boxes are related to one another. The use of A. nidulans, E. coli and M. tuberculosis is to give specific examples as the homologies between the various genes are evident across a wide range of species. Refer to the text for references to these and other examples in the literature.
interface involved contacts between the two DHQ synthase domains and between the two shikimate dehydrogenase domains (Case et al., 1969; Case & Giles, 1971). Later work confirmed that the native AROM protein in N. crassa is a dimer and demonstrated through the use of limited proteolysis peptide mapping and suberimidate cross-linking studies that proteolytically cleaved AROM protein retains its native quaternary structure (Lumsden & Coggins, 1978; Smith & Coggins, 1983). The cross-linking studies further strongly suggest that the two individual polypeptides comprising the native AROM protein adopt a compact globular shape with the N- and C-terminal domains (DHQ synthase and shikimate dehydrogenase respectively) closely juxtaposed (Smith & Coggins, 1983). These observations suggest a mechanism for how the repressor protein is able to recognize and bind to the activator. As the activator and repressor appear to have evolved by splitting a duplicated copy of the gene encoding the pentafunctional AROM polypeptide in half, the two proteins may come into close juxtaposition with the shikimate dehydrogenase domain of the repressor in close proximity with the N-terminus (the DHQ synthase like domain) of the activator in a manner similar to the AROM protein. This arrangement would give the potential for the repressor 3-dehydroquinase like and/or shikimate kinase like domains to occlude the surface acidic residues located near the C-terminus of the activator that are proposed to be important for its interaction with the accessory transcription complex (Ptashne, 1988). These observations also reinforce the link between the fact that the N-terminal domain of the repressor protein is homologous to a metabolic kinase (shikimate kinase) and that the quinate inducible qutG gene encodes a protein that is homologous with bovine myo-inositol monophosphatase. It is hard to believe (though not impossible) that the metabolic kinase like domain of the QUTR repressor protein could have evolved into a protein kinase; similarly, it is unlikely that the qutG encoded myo-inositol phosphatase like protein is a protein phosphatase. The discovery that the QUTA activator protein is homologous with the two N-terminal domains of the AROM protein (DHQ synthase and EPSP synthase), however, suggests a speculative model for how transcriptional regulation may be influenced by a reversible phosphorylation event. The C-terminal half of the QUTA activator is homologous with the EPSP synthase domain of the AROM protein, and this domain is able to recognize shikimate phosphate as a substrate. It is therefore possible that the C-terminal half of the QUTA protein could bind a similar metabolite that is subject to reversible phosphorylation by the products of qutG encoded protein and the shikimate kinase like domain of the QUTR repressor protein. These events may generate a phosphorylated isoform of the activator that is transcriptionally hyperactive in a manner analogous to the GAL4 activator in S. cerevisiae (Mylin et al., 1990).

The quinate pathway may be the first unequivocally documented case where it is shown that virtually an entire catabolic pathway (the only exception is the qutC gene which shows no extensive homology to any protein in known databases) including regulatory proteins has been constructed by co-opting preformed domains from related pathways and evolving pathway specific functions (see Fig. 2). Four genes, qutA, R, E and B have been co-opted from the shikimate pathway, the qutD encoded permease is almost certainly a modified sugar transporter (Whittington et al., 1987; Geever et al., 1989), qutG is a member of a large group of phosphatases involved in dephosphorylating sugars (Neuwald et al., 1991) and qutH shows a low level homology with the Hepatitis B DNA polymerase and has a conserved zinc binuclear cluster motif (Geever et al., 1989; Lamb et al., 1990, 1992).

**Domain structure within the AROM protein**

Domain structure and interaction within the A. nidulans pentafucntional protein have been investigated using *in vivo* and *in vitro* techniques. Sub-fragments of the aromA gene (encoding the AROM protein), have been expressed in appropriate *E. coli* mutant strains lacking particular shikimate pathway enzymes using prokaryotic expression vectors to complement the *aro* mutations and allow growth in minimal medium (Hawkins & Smith, 1991). These initial studies defined the approximate limits of the functional domains within the AROM protein, and subsequent work using polymerase chain reaction (PCR) amplified DNA was able to more precisely delineate the limits of these domains (van den Hombergh et al., 1992; Moore & Hawkins, 1993). The complete AROM protein has been overproduced in A. nidulans by around 150-fold, constituting around 6% of the soluble cell protein (Moore et al., 1992). The two N-terminal domains DHQ synthase and EPSP synthase have been substantially overproduced in *E. coli* as monofunctional domains (DHQ synthase), bifunctional domains (DHQ synthase and EPSP synthase) or as fusion proteins with glutathione S-transferase (EPSP synthase) (van den Hombergh et al., 1992; Moore & Hawkins, 1993). An *in vivo* protein based complementation test in *E. coli* has also identified an interaction between the N-terminal domains of DHQ synthase and EPSP synthase that stabilizes the activity of the latter enzyme (Moore & Hawkins, 1993). The type I 3-dehydroquinase domain of the *A. nidulans* AROM protein has also been overproduced in and purified from *E. coli*, and has substantially different kinetic parameters compared to the same domain of the
The physiology of the AROM protein

The contemporary functional significance of why the enzymes catalysing steps two to six in the shikimate pathway are fused into a single multifunctional protein in fungi is the subject of a long running controversy. Early work based on genetic studies and \textit{in vitro} analysis of purified AROM protein proposed that the fused protein acted as a channelling mechanism to keep the quinate and shikimate pathway derived metabolites DHQ and DHS separate (Giles, 1978). Further work additionally ascribed the function of catalytic facilitation to the AROM protein, but these \textit{in vitro} studies were subsequently reported to have been carried out with protein that was deficient in zinc (DHQ synthase requires zinc) and had suffered significant proteolysis (Gaertnert et al., 1970; Welch & Gaertnert, 1975, 1976; Lambert et al., 1985). The channelling hypothesis was challenged and an alternate explanation of ensuring the co-ordinate production of all five enzymes was proposed to explain the functional significance of having a pentafunctional protein (Duncan et al., 1988).

Most recently, this controversy has been addressed in \textit{A. nidulans} by taking an \textit{in vivo} approach in contrast to earlier \textit{in vitro} experiments. Recombinant DNA technology has been used to produce strains of \textit{A. nidulans} that have particular quinate or shikimate pathway enzyme concentrations up-modulated. Strains of \textit{A. nidulans} lacking the \textit{qutE} encoded catabolic type II 3-dehydroquinase are unable to utilize quinate as carbon source (see Fig. 1). If the AROM protein concentration in these strains is increased 5-fold, they are able to utilize quinate as carbon source demonstrating that the AROM protein type I 3-dehydroquinase is able to synthesize DHS from quinate pathway DHQ and subsequently release it, thereby showing that the AROM protein is leaky \textit{in vivo} (Lamb et al., 1991).

If the AROM protein concentration is raised over 5-fold in \textit{A. nidulans} strains with wild-type \textit{qut} gene cluster genes, then these strains are unable to grow in the presence of quinate (even on mixed carbon sources with glycerol) and have the \textit{qut} pathway enzymes induced to a much lower level than in wild-type strains. The proposed cause of this effect is that the overproduced AROM protein is diverting flux of quinate pathway derived DHQ and DHS into the shikimate pathway at an increased rate, leading to one of the shikimate pathway metabolite concentrations reaching toxic levels (Lamb et al., 1991). Reciprocal experiments in which the \textit{qutC} encoded dehydroshikimate dehydratase concentration is manipulated demonstrate that substantial flux of DHS from the shikimate pathway can be diverted into the quinate pathway. Dehydroshikimate dehydratase catalyses the conversion of DHS to protocatechuate and is therefore active at the branch point between the quinate and shikimate pathways (see Fig. 1). With increases of dehydroshikimate dehydratase up to 10-fold over the normal wild-type level, an \textit{N}-fold increase in enzyme concentration gives an \textit{N}-fold increase in pathway flux (Lamb et al., 1992b). This is equivalent to saying that the dehydroshikimate dehydratase enzyme has a shikimate pathway deviation control index \((D)\) of \(-1\), where \(D^p\) is defined as \((\Delta J/J)/(\Delta E/E)\), where \(\Delta\) stands for a large change, \(J\) stands for a pathway flux and \(E\) stands for the enzyme concentration under consideration (Niederberger et al., 1992).

The \textit{A. nidulans} strains overproducing the dehydroshikimate dehydratase display a varying auxotrophic requirement for aromatic amino acids which becomes more pronounced as the enzyme concentration increases (Lamb et al., 1992b). This increasing auxotrophic requirement for aromatic amino acids is most simply interpreted as a consequence of diverting substantial flux from the shikimate pathway into the quinate pathway with an \textit{N}-fold increase in enzyme (dehydroshikimate dehydratase) concentration giving an approximate \textit{N}-fold change in the flux ratio of the two branches of the pathway (Lamb et al., 1991; Hawkins et al., 1993c). Additionally, strains overproducing the dehydroshikimate dehydratase enzyme show levels of the AROM protein elevated up to 5-fold over the wild-type level.
(Lamb et al., 1992b). These observations are consistent with the proposal that the AROM protein and dehydroshikimate dehydratase enzyme are displaying a mass action effect on a substrate (DHS) that approximates pool behaviour, without any need to impute a channelling function for the AROM protein (Lamb et al., 1991; Hawkins et al., 1993c).

These experiments whilst unequivocally demonstrating that the AROM protein is extremely leaky in vivo are dealing with the fungus growing under ideal laboratory conditions and do not rule out the possibility that there is a low level channelling property associated with the AROM protein. In this regard, analysis of strains overproducing the dehydroshikimate dehydratase up to 30-fold over the wild-type level have given evidence that the AROM protein does have a very low selective flux (channelling) effect due to the close juxtaposition of five active sites, but that this effect is only of physiological significance when the organism is growing in its natural environment where nutrient supply, aeration and water supply will not be optimal (Lamb et al., 1992b).

Pathway manipulation

In A. nidulans, when quinate is supplied as the sole carbon source, there is an approximately 30-fold difference in the level of flux through the quinate pathway compared to the shikimate pathway as judged by the experiments up-modulating the dehydroshikimate dehydratase enzyme (Lamb et al., 1992b). The quinate pathway is an entirely dispensable high flux pathway as A. nidulans and related fungi have a wide array of alternate carbon sources that they can utilize for energy and growth. It should therefore be possible to divert de novo carbon flux in the form of the metabolites DHQ and DHS from the quinate pathway into the lower flux shikimate pathway thereby circumventing the end-product inhibition on the DAHP synthases (Doy, 1968; Nimmo & Coggins, 1981; Braus, 1991), converting the latter into a high flux pathway possibly allowing the increased production of metabolites that are of commercial interest. Experiments with A. nidulans over-producing a mutant form of the AROM protein lacking the 3-dehydroquinate synthase enzyme activity have shown that substantial flux of DHQ and DHS can be diverted from the quinate pathway into the shikimate pathway (Lamb et al., 1991; Moore & Hawkins, 1993). These experiments showed that (a) it is possible to circumvent the end-product inhibition on the DAHP synthases, the first step in the pathway and so increase pathway flux, and (b) that this increase in flux is limited to the pre-chorismate section of the shikimate pathway leading to one or more pathway intermediates reaching toxic levels. This inability to increase overall pathway flux by up-modulating a single block of pathway enzymes is to be expected from a consideration of the principles of metabolic control analysis (MCA). MCA theoretically deals with infinitesimal changes in enzyme and metabolite concentrations, and in fluxes, but in practice small but finite changes are introduced into experimental systems. The changes in enzyme concentration and fluxes reported for in vivo quinate and shikimate pathway manipulations in A. nidulans clearly cannot be considered as small (Lamb et al., 1991, 1992b); however, recent theoretical developments in MCA based on experiments with the shikimate pathway in S. cerevisiae are pointing the way forward (Niederberger et al., 1992). Essentially these new observations show that in order to increase substantial overall pathway flux then ideally all of the enzymes in the pathway should be up-modulated simultaneously (Niederberger et al., 1992). This requirement presents technically difficult obstacles unless the genes encoding all the enzymes in a chosen pathway have been isolated by molecular cloning. This is a greater or lesser problem depending on the tractability of the organism under study, but a second way of circumventing this problem is to express genes that have previously been isolated from other species encoding the required enzyme activities in the organism under study. This approach has been successfully employed in A. nidulans where a 158 bp truncated promoter from the quitE gene containing a putative upstream activator sequence recognized by the QUTA protein has been engineered into an A. nidulans expression vector containing the quitE transcription termination signals (Lamb et al., 1991; Moore et al., 1992; Moore & Hawkins, 1993). This plasmid successfully places any gene appropriately fused to the promoter under the control of the positively acting QUTA transcription regulator protein allowing quinate inducible production of heterologous proteins in A. nidulans (Lamb et al., 1991; Moore et al., 1992; Moore & Hawkins, 1993). Later versions of this expression plasmid containing the constitutive high level A. nidulans phosphoglycerate kinase promoter (Clements & Roberts, 1986; H. K. Lamb & A. R. Hawkins, unpublished data), should allow up-modulation of all the steps in the shikimate pathway by expression of appropriate heterologous genes in A. nidulans. This approach may allow the diversion of major new carbon flux from the dispensable quinate pathway into the shikimate pathway turning the latter into a high flux pathway.

References


