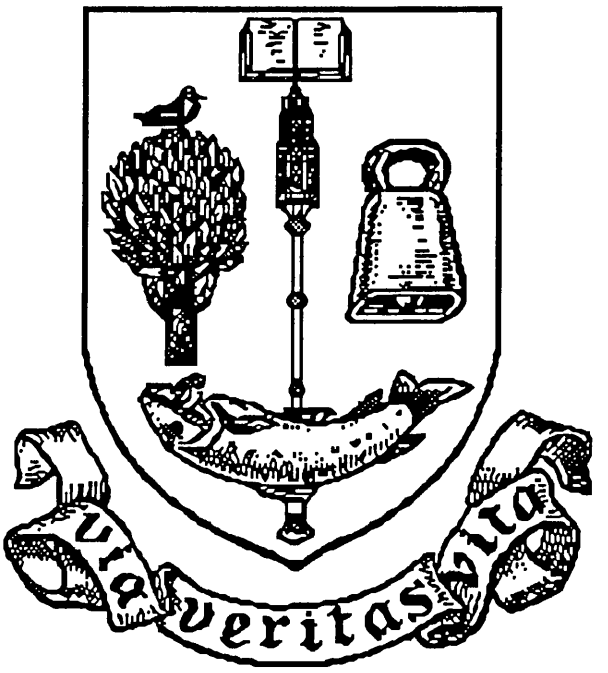
Studies of the Shikimate Pathway Enzyme, 3-Dehydroquinase

A thesis
submitted to the
University of Glasgow

for the degree of
DOCTOR OF PHILOSOPHY

IN
BIOCHEMISTRY



BY
RANJIT KUMAR DEKA

Department of Biochemistry
University of Glasgow
Glasgow, G12 8QQ
SCOTLAND

FEBRUARY, 1993

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

I here by declare that this thesis embodies the results of my own special work, that it has been composed by myself and that it does not include work forming part of thesis presented successfully for a degree in this or another University.

Date: 19th Feb. 1993 Signature:

For my parents

ACKNOWLEDGEMENTS

I would like to thank the following people and organisations who have made the work described in this thesis possible.

Professor J.R. Coggins for his guidance, patience, enthusiasm and instruction during the course of this work. Dr. Colin Kleanthous, Dr. David Mousdale and Dr. Ian Anton for all their help and advice with the project. All the members of the D-Floor and B4a research groups, past and present, without whom the project could not have been completed.

Mr. John Greene for the technical assistance and for the maintenance and repair of apparatus essential for this work.

The Government of Assam, India for financial support.

Dr. Maggi Cusack, Department of Pure and Applied Geology, University of Glasgow, Dr. David Campbell, Department of Biochemistry, University of Dundee and Mr. Brian Dunbar, Department of Molecular and Cell Biology, University of Aberdeen for protein and peptide sequencing.

Dr. Nicholas Price, Department of Biological and Molecular Sciences, University of Stirling, for assistance with measurement of CD spectra.

Dr. Callum Livingstone for proof-reading of this thesis.

Finally, I would like to thank all my colleagues in India for their support and help in getting me admitted to Glasgow University for Ph.D. research and especially my friends and all the members of my family whose tremendous support and encouragement inspired me to come abroad to study for my Ph.D.

Contents **Page No.**

Chapter II

Materials and Methods

Chapter III

The essential histidine residue at the active site of *Escherichia coli* dehydroquinase

Chapter IV

Mechanistic properties of type II dehydroquinase from *Aspergillus nidulans*

Chapter V

Characterisation of *Pisum sativum* bifunctional dehydroquinase-shikimate dehydrogenase

Chapter VI

Cloning of a cDNA for the dehydroquinase region of *Pisum sativum* dehydroquinase-shikimate dehydrogenase

List of figures

List of tables

Abbreviations

The abbreviations used in this thesis are set out in the Biochemical Journal "Instruction to Authors", except the following:

A	absorbance
AMP.	ampicillin
ATP	adenosine triphosphate
AUFS	absorbance units full scale
BSA	bovine serum albumin
CAPS	3-[cyclohexylamino]-1-propanesulfonic acid
CD	circular dichroism
(k)Da	(kilo) daltons
DEPC	diethylpyrocarbonate
DEAE	diethylaminoethyl
DHQase	dehydroquinase
DMDS	dimethyldichlorosilane
DNAse	deoxyribonuclease
dNTP	deoxynucleoside 5'-triphosphate
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetate
EPSP	5-enolpyruvyl-shikimate 3-phosphate
EtBr	ethidium bromide
FPLC	fast protein liquid chromatography
HPLC	high pressure liquid chromatography
IPTG	isopropyl-β-D-thiogalactoside
kb	kilo base pairs
l.m.p.	low melting point

Mr	molecular weight
NAD^+	nicotinamide adenine dinucleotide
$NADP^+$	nicotinamide adenine dinucleotide phospate
NADPH	reduced nicotinamide adenine dinucleotide phospate
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEP	phosphoenol pyruvate
Pipes	1,4-piperazinediethanesulfonic acid
PG	phenylglyoxal
PLP	pyridoxal 5'-phosphate
PMSF	phenylmethanesulphonylfluoride
PVDF	polyvinylidene difluoride
Rf	mobolity of a protein compared to the dye front
RNAse	ribonuclease
rpm	revolution per minute
SDHase	shikimate dehydrogenase
SDS	sodium dodecyl sulphate
TE	Tris-acetate/EDTA buffer
TBE	Tris-borate/EDTA buffer
TEMED	N,N,N'N'-tetramethylethylene diamine
TFA	trifluoroacetic acid
TLCK	*N-p*-toluenesulphonyl-L-lysine chloromethylketone
TPCK	*N-p*-toluenesulphonyl-L-phenylalanine chloromethylketone
Tris	Tris (hydroxymethyl) aminomethane
U	units of enzyme activity
UV	ultra violet
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactoside

SUMMARY

1. Chemical modification with diethylpyrocarbonate (DEPC) and kinetic analyses have established a role for a single active site histidine residue as the general base involved in the mechanism of the type I dehydroquinase of *E. coli*.

2. Differential peptide mapping of DEPC-modified dehydroquinase in the absence and presence of substrate identified a peptide containing the active site histidine residue.

3. Sequence comparison of the active site peptide with other type I dehydroquinases allowed the identification of the active site histidine as His-143 in the *E. coli*.enzyme sequence.

4. The type II dehydroquinase of *A. nidulans* shows very different properties (eg. K_m, V_{max}, V_{max} versus pH plots) from those found for the type I enzyme from *E. coli*.

5. The type II enzyme from *A. nidulans* does not work through a Schiff base adduct since it cannot be inhibited by treatment with substrate plus sodium borohydride.

6. The type II dehydroquinase from *A. nidulans* like the type I enzyme from *E. coli*, has at least one essential active site histidine residue.

7. The type II dehydroquinase is very susceptible to phenylglyoxal inactivation indicating the involvement of an arginine residue(s) in its mode of action.

8. The previously published purification of the *P. sativum* dehydroquinase-shikimate dehydrogenase has been scaled up and the yield improved. The purified *P. sativum* dehydroquinase has a similar K_m and pH optimum to the type I dehydroquinase from *E. coli.*

9. A borohydride trapping experiment established that the dehydroquinase from *P. sativum* is a Schiff base forming type I enzyme. Chemical modification with the histidine specific reagent DEPC suggests that, like the *E. coli* type I enzyme, there is a histidine residue essential for catalytic activity.

10. The shikimate dehydrogenase activity of the *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase was not inhibited by the dehydroquinate/sodium borohydride treatment which completely inactivated the dehydroquinase activity. This establishes that the shikimate dehydrogenase and dehydroquinase active sites are spatially distinct.

11. N-terminal equencing of the intact *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase and of peptides generated by Cleveland mapping provided 157 amino acid residues of sequence. All of the sequences determined showed similarities to regions of either the dehydroquinase or the shikimate dehydrogenase of *E. coli.*

12. The N-terminal sequence of the intact *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase is clearly homologous to the N-terminal region of *E. coli.* dehydroquinase; this clearly establishes that in the bifunctional polypeptide the dehydroquinase domain is N-terminal.

13. Using the peptide sequences to aid primer design a simple PCR strategy was developed to amplify fragments of the cDNA encoding the entire dehydroquinase coding region and part of the shikimate dehydrogenase coding region of the *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase.

14. The complete amino acid sequence of the dehydroquinase domain of the *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase was deduced from the nucleotide sequence of the DNA fragments. At the amino acid level there was from 23 to 28% identity between the *P. sativum* dehydroquinase domain and the other known type I dehydroquinase domains; both the known active sites residues, His-143 and Lys-170 (numbering is according to the *E. coli* enzyme), are conserved in the *P. sativum* enzyme. Another residue, Asp-144, which is conserved in all previously studied type I dehydroquinases, is not conserved in the *P. sativum* enzyme which implies that this residue is not involved in the mechanism of action.

CHAPTER I

The shikimate pathway: a general introduction

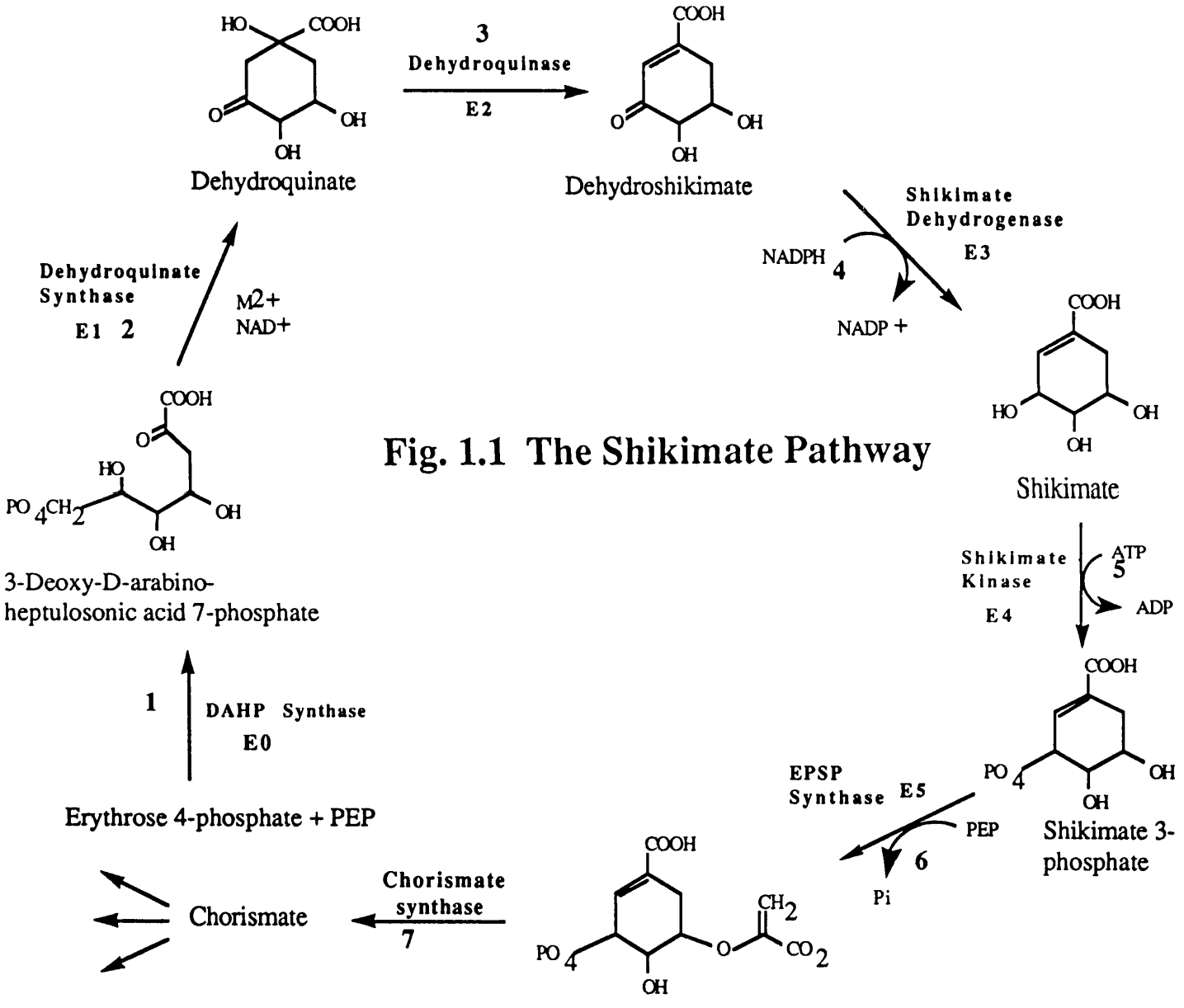


Fig. 1.1 The Shikimate Pathway

1.1 The shikimate pathway: a general introduction

In plants and microorganisms, chorismate is the common metabolic precursor of the aromatic amino acids, phenylalanine, tryptophan and tyrosine, as well as such vital and important metabolites as vitamin E and K, folic acid, ubiquinone, plastiquinone and enterochelin. Many microbial and plant secondary metabolites containing aromatic rings are also derived from the shikimate pathway intermediates. It has been estimated that between 10 and 35% of the dry weight of some higher plants is generated via the shikimate pathway (Boudet *et al.*, 1985; Coggins, 1986).

The pathway derives its name from the first isolated intermediate on the pathway - "shikimic acid" - which was obtained from the plant *Illicium regiosum* (Japanese name *Shikimi-no-ki*, hence the name of the compound) (Eykmann, 1885). However, the complete pathway was not elucidated until the work of Davis (1951, 1955) and Sprinson (1960) who used bacterial auxotrophic mutants to isolate the pathway intermediates (Davis, 1955; Levin and Sprinson, 1964; Gibson and Pittard, 1968).

The shikimate pathway consists of seven discrete enzymatic steps begining with the condensation of two products of carbohydrate metabolism, phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4-P) and finally giving rise to chorismic acid, which is the common precursor of almost all aromatic compounds (Haslam, 1974; Weiss and Edwards, 1980) (Fig. 1.1).

Mammals lack this biosynthetic ability and require a supplementary dietary intake of aromatic amino acids and vitamins. Since the nutritionally essential amino acids tryptophan and phenylalanine are synthesised by the shikimate pathway in plants and microorganisms,

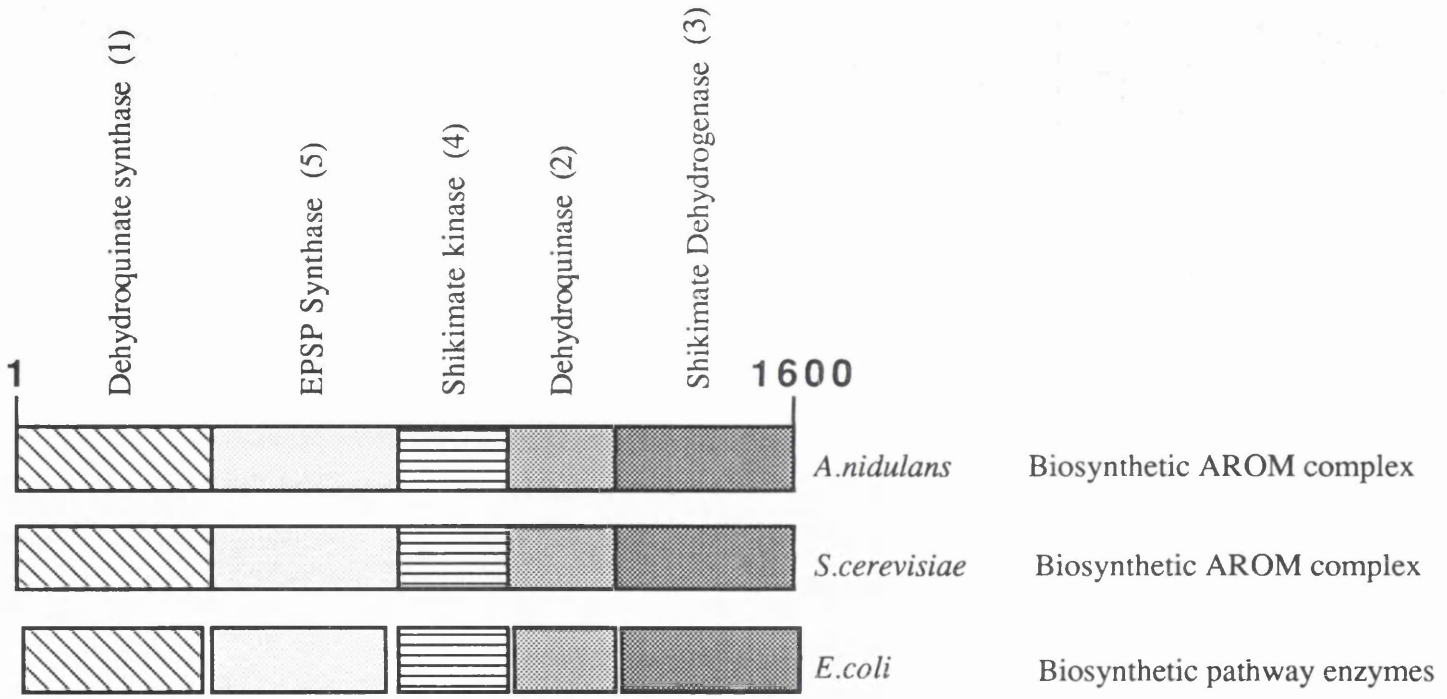
this pathway has prime importance. Attention has focussed on this pathway as a potential target for both herbicides and antimicrobial agents, since inhibitors of the enzymes of this pathway should prove to be harmless to animals (Coggins, 1986). Indeed the very commercially important herbicide glyphosate (Roundup®) is known to act through inhibition of the penultimate enzyme of the prechorismate portion of the shikimate pathway, EPSP synthase (Steinrucken and Amrhein, 1980; Boocock and Coggins, 1983). This finding underlined the importance of detailed mechanistic study of the enzymes of this pathway to aid the rational design of both novel herbicides and antimicrobial agents.

Detailed studies of the shikimate pathway enzymes have been hampered principally because the enzymes were present at relatively low levels, particularly in plants, and they were very difficult to purify (Coggins, 1986; Bentley, 1990; Mousdale and Coggins, 1992). However, the advancement of molecular biological techniques in particular recombinant DNA technology, has enabled a number of the shikimate pathway enzymes to be cloned and overexpressed (Bentley, 1990). As a result, a number of purified shikimate pathway enzymes are available in milligram quantities, or even hundreds of milligram quantities (Bentley, 1990) and at least two shikimate pathway enzymes, EPSP synthase from *Escherichia coli* (Stallings *et al.*, 1991) and dehydroquinase from *Salmonella typhi* (Boys *et al.*,1992) have been crystallised.

1.2 Organisation of the shikimate pathway enzymes

The reaction sequence in the prechorismate pathway is identical in all organisms investigated so far, but there are considerable differences in enzyme organisation (Fig. 1.2) between organisms of different taxonomic groups (Bentley, 1990). For example, in *Escherichia coli*, the seven enzymatic activities of the prechorismate pathway are associated with a monofunctional polypeptides (Berlyn and Giles, 1969) and their genes

Fig. 1.2 Organisation of Enzymes on the Shikimate Pathway



are unlinked in the bacterial genome (Pittard and Wallace, 1966). In contrast, in fungi the five enzyme activities converting DAHP to EPSP reside on a single pentafunctional polypeptide, the *arom* complex, which is the product of a single gene (Lumsden and Coggins, 1977; Gaertner and Cole, 1977). In higher plants (apart from *Euglena gracilis*, the organisation of which resembles that found in fungi, Ahmed and Giles, 1969; Patel and Giles, 1979), two activities (dehydroquinase-shikimate dehydrogenase) are carried on a bifunctional polypeptide (Polley, 1978; Koshiba, 1979; Mousdale *et al.*, 1987), while the other five enzymes are monofunctional (for review see Bentley, 1990; Mousdale and Coggins, 1992).

A schematic representation of the structural organisation of the shikimate pathway in number of organisms is shown in Fig. 1.2.

1.3 The individual enzyme of the shikimate pathway

Since the establishment of the site of action of the herbicide glyphosate, the enzymes of the shikimate pathway have attracted attention as potential sites for the action of both herbicides and antimicrobial agents. This has stimulated much work on the mechanisms of the individual enzymes. Recent work on the seven enzymes is summarised below with a more extensive section on dehydroquinase which is the major topic of this thesis.

1.3.1 DAHP synthase

The oxidation of glucose via the glycolytic and pentose phosphate pathways yields phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4-P), respectively. The first enzyme of the shikimate pathway, DAHP synthase (3-Deoxy-D-arabino-heptulosonate 7-

phosphate synthase, EC 4.1.2.15), catalyses the condensation of these two compounds to give 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) and inorganic phosphate (Srinivasan and Sprinson, 1959a) (Step 1 in Fig. 1.1). The role of this enzyme as an important site of regulation has been clearly demonstrated in bacteria, where modulation of the enzyme activity by feedback inhibition is the major mechanism of controlling carbon flow into the shikimate pathway (Ogino *et al.*, 1982).

The mechanism of this condensation reaction is still not understood. Srinivasan and Sprinson (1959b) suggested that nucleophilic attack at the phosphorous atom of PEP resulting in P-O cleavage and generation of a reactive *enol* pyruvate anion which can rapidly add to E4-P. Later it was shown that this condensation proceeded with C-O cleavage rather than P-O cleavage (Floss *et al.*, 1972; DeLeo *et al.*, 1973). Ganem (1978) suggested a reaction mechanism which involving addition of a enzyme-thiol group to PEP as many DAHP synthases have been shown to be susceptible to inhibition by thiol modifying reagents (Staub and Denes, 1969a, b; Gorisch and Lingens, 1971; Huisman and Kosuge, 1974; Sugimoto and Shiio, 1980; Ray and Bauerle, 1991).

Three isozymes of DAHP synthase have been found in *E. coli, Salmonella typhimurium* (Davis, 1951; Pittard, 1987), *Neurospora crassa* (Nimmo and Coggins, 1981) and *Saccharomyces cerevisiae* (Paravicini *et al.*, 1988) which are each inhibited by one of the three aromatic amino acids. Thus, in this way the metabolic flow through the shikimate pathway is regulated by feedback inhibition on DAHP synthase.

Plant DAHP synthases are not inhibited by the aromatic amino acids; in contrast, the enzymes from carrot cells and potato tubers are activated by tryptophan and tyrosine (Suzich *et al.*,1985; Pinto *et al.*,1986). The enzyme activity varies during the growth cycle of carrot (Suzich *et al.*, 1984) and potato (Pinto *et al.*, 1988) cells in suspension culture. DAHP synthase is induced in potato cells that are challenged with herbicide

glyphosate (Pinto *et al.*, 1988), in parsley cells that are challenged with a fungal elicitor (McCue and Conn, 1989) and by wounding potato tubers (Dyer *et al.*, 1989; Muday and Herrmann, 1992) and *Arabidopsis thaliana* (Keith *et al.*, 1991). Thus plant DAHP synthases are developmentally and environmentally regulated.

Two DAHP synthase isozymes have been distinguished in plants through their metal requirements: a Mn^{2+}-activated form, which is *plastidic* and a Co^{2+} requiring form (which also requires another divalent cation, probably Mg^{2+}) which is *cytosolic* (Rubin and Jensen, 1985; Ganson *et al.*, 1986; Morris *et al.*, 1989). Interestingly, the cytosolic form of DAHP synthase from *Nicotiana silvestris* is inhibited by millimolar concentrations of glyphosate (Ganson and Jensen, 1988) like that of mung bean DAHP-synthase-Co^{2+} form (Rubin *et al.*, 1982); while the Mn^{2+}-activated form (plastidic form) is resistant to glyphosate inhibition (Ganson and Jensen, 1988).

1.3.2 3-Dehydroquinate synthase

3-Dehydroquinate synthase (EC 4.6.1.3) catalyses the second reaction of the shikimate pathway, namely the enzymic self-condensation of DAHP to give 3-dehydroquinate, the first cyclic intermediate on the pathway (Step 2 in Fig. 1.1). This reaction was first demonstrated in 1963 (Srinivasan *et al.*, 1963), following the earlier isolation and identification of dehydroquinate as a pathway intermediate (Weiss *et al.*, 1953).

By using analogues and radiolabelled substrates, the reaction was shown to involve an intramolecular exchange of the DAHP ring oxygen with C-7, accompained by an oxidation at C-6 and a reduction at C-2 (LeMarechal and Azerad, 1976; Rotenberg and Sprinson, 1970, 1978). Later it was suggested that an enzyme group is not required for

deprotonation of C-6 and instead, the phosphate group of DAHP itself functions as the base with the subsequent elimination of phosphate (Bartlett and Satake, 1988; Widlanski *et al.*, 1989). It was shown that both NAD^+ and divalent metal ions were essential for 3-dehydroquinate synthase activity, especially for NAD^+-linked oxidation and reduction of the C-5 hydroxyl group of DAHP (Srinivasan *et al.*, 1963). Lambert *et al.* (1985) noted that the enzyme recognised a pyranose form of DAHP rather than a straight chain conformation.

3-Dehydroquinate synthase has been purified to homogeneity from *E. coli* (Frost *et al.*, 1984; Millar and Coggins, 1986), *Bacillus subtilis* (Frost *et al.*, 1984) and *Phaseolus mungo* (Pompliano *et al.*, 1989). The *E. coli* enzyme is monomeric, whilst the plant enzyme is dimeric. In *B. subtilis* the enzyme is associated with chorismate synthase and a flavin reductase in a multifunctional complex. The 3-dehydroquinate synthase activity of the *arom* complex from *N. crassa* has been co-purified in constant ratio to the other activities and shown to be Zn^{2+} dependent (Lambert *et al.*, 1985).

1.3.3 3-Dehydroquinase

This is the major topic of this thesis and is discussed in Section 1.5.

1.3.4 Shikimate dehydrogenase

Shikimate dehydrogenase (Shikimate: $NADP^+$ oxido-reductase, EC 1.1.1.25) catalyses the stereospecific reduction of 3-dehydroquinic acid to shikimic acid utilizing hydride transfer from NADPH (Yaniv and Gilvarg, 1955; Dansette and Azerad, 1974) (Step 4 in Fig. 1.1).

Shikimate dehydrogenase, which has been purified to homogeneity from *E. coli* is a monomer of Mr 30,000 (Chaudhuri and Coggins, 1985) and the *aro*E gene encoding this protein has been cloned, sequenced and overexpressed (Anton and Coggins, 1988). The predicated polypeptide has 272 amino acid residues and is homologous to the shikimate dehydrogenase domains of the *arom* complexes of *Aspergillus nidulans* (Charles *et al.*, 1986) and *S. cerevisiae* (Duncan *et al.*, 1987). Interestingly, the sequence contains a consensus nucleotide-binding region (probably the $NADP^+$ binding site) (Anton and Coggins, 1988).

As mentioned in Section 1.3.3, the shikimate dehydrogenase activity resides with 3-dehydroquinase on a bifunctional polypeptide in a variety of plants (see Mousdale and Coggins, 1992 for recent review).

Baillie *et al.* (1972) have synthesised a number of analogues of 3-dehydroshikimic acid and found them to be fairly potent inhibitors (K_i's in the range 70 to 800μM), but they failed as herbicides *in vivo*.

1.3.5 Shikimate kinase

Shikimate kinase (EC 2.7.1.71) catalyses the fifth step (Step 5 in Fig. 1.1) in the shikimate pathway, namely the energy requiring irreversible phosphorylation of shikimate to shikimate 3-phosphate with ATP as co-substrate (Weiss and Mingioli, 1956).

E. coli and *S. typhimurium* possess two shikimate kinases, shikimate kinase I and shikimate kinase II (Berlyn and Giles, 1969; Ely and Pittard, 1979). Kinetic analysis of the *E. coli* isozymes showed that the K_m of shikimate kinase I for shikimate is about 100-fold higher than that for shikimate kinase II, suggesting that shikimate kinase II is the

principal enzyme which phosphorylates shikimate in the cell (DeFeyter and Pittard, 1986). The shikimate kinase II from *E. coli* K12 has been purified and overexpressed by cloning *aro*L gene and found to be a monomeric enzyme containing 173 amino acid residues (Millar *et al.*, 1986; DeFeyter and Pittard, 1986). More recently the gene (*aro*K) encoding shikimate kinase I has been identified in *E. coli* and and sequence studies show 34% homology to shikimate kinase II (Løbner-Olesen and Marinus, 1992). Both these sequences contain a region (GXXXXGKT/S) homologous with other kinases and ATP-requiring enzymes (Walker *et al.*, 1982) that is likely to be part of the ATP-binding site (Millar *et al.*, 1986; Løbner-Olesen and Marinus, 1992). Unlike shikimate kinase II, the shikimate kinase I appears to be independent of *tyr*R regulation (Løbner-Olesen and Marinus, 1992).

The *E.coli* shikimate kinases are also homologous to the sequences of the shikimate kinase domains of the *A. nidulans* (Charles *et al.*, 1986) and *S. cerevisiae* (Duncan *et al.*, 1987) *arom* complexes, particularly in the region of the ATP-binding motif (Millar *et al.*, 1986).

Shikimate kinase activity has been detected in several plant species (see Mousdale and Coggins, 1992). The enzyme has been purified to near homogeneity from spinach chloroplasts and is a monomer of Mr 31,000 (Schmidt *et al.*, 1990). The first plant shikimate kinase cDNA encoding 300 amino acid residues has been cloned and sequenced from tomato; the deduced amino acid sequence shows homology to bacterial and fungal shikimate kinases and contains an ATP-binding region (Schmid *et al.*, 1992).

1.3.6 EPSP synthase

EPSP synthase (5-enolpyruvyl-shikimate 3-phosphate synthase, EC 2.5.1.19) catalyses the reversible transfer of phosphoenolpyruvate to shikimate 3-phosphate to give 5-enolpyruvyl-shikimate 3-phosphate (EPSP) and inorganic phosphate (Levin and Sprinson, 1964) (Step 6 in Fig. 1.1). This reaction has highlighted the role of the shikimate pathway in plant metabolism, since EPSP synthase is the major target for inhibition by the commercially important, broad spectrum, nonselective, postemergence herbicide glyphosate (Roundup®) (Steinrucken and Amrhein, 1980; Boocock and Coggins, 1983).

Two decades ago, Sprinson and his colleagues suggested a mechanism which involves the protonation of C-3 of PEP and the transfer of the enolpyruvate moiety of PEP onto the hydroxyl group at C-5 on the acceptor molecule, shikimate 3-phosphate, to form a tetrahedral intermediate. The cleavage of the C-O bond (rather than of the O-P bond) at C-2 of PEP facilitates the elimination of the phosphate (Levin and Sprinson, 1964; Bondinell *et al.*, 1971). The "Sprinson mechanism" has recently been conclusively validated by the isolation of the predicted tetrahedral intermediate (Anderson *et al.*, 1988a, b).

Extensive chemical modification studies of EPSP synthase have implicated a large number of amino acid residues in the active sites. These include lysine (Lys-22 for *E. coli*, and Lys-23 for *Petunia hybrida*) (Huynh *et al.*, 1988), arginine (Arg-28 and Arg-131 for *P. hybrida*) (Padgette *et al.*,1988b), glutamate (Glu-418 for *E. coli*) (Huynh, 1988), two cysteines (Cys-288 and Cys-408 for *E. coli*) (Padgette *et al.*, 1988a) and a histidine residue with a pK_a of 6.8 in *E. coli* (Huynh, 1987). Assignment of roles for these residues must await determination of the high resolution three-dimensional structure of the enzyme. EPSP synthase has been crystallized by the Monsanto group and the backbone structure determined at 3-Å resolution (Stallings *et al.*, 1991).

EPSP synthase from *E. coli* has been purified, sequenced and overexpressed (Lewendon and Coggins, 1983; Duncan *et al.*, 1984; Duncan *et al.*, 1985) and is a monomer of Mr 46,112. The crucial step in the purification from wild-type *E. coli* was substrate elution from phosphocellulose. Shikimate 3-phosphate and PEP together were most effective in eluting the activity and this indicates an ordered reaction in which shikimate 3-phosphate binds to the enzyme first (Boocock and Coggins, 1983).

This purification was followed by others from a range of bacteria and plants, including *S. typhimurium* (Stalker *et al.*, 1985), *Klebsiella pneumoniae* (Steinrucken and Amrhein, 1984), pea seedling (Mousedale and Coggins, 1984) and *Anabaena variabilis* (Powell *et al.*, 1992), with sequence data also being obtained for the first two. Subsequently, other sequences were obtained from *A. thaliana* (Klee *et al.*, 1987), *Bordetella pertussis* (Maskell *et al.*, 1988), petunia and tomato (Gasser *et al.*, 1988), pea (Granger, 1989), tobacco (Wang *et al.*, 1991), *Aspergillus nidulans* (Charles *et al.*, 1986) and yeast (Duncan *et al.*, 1987). There is considerable homology between the plant, bacterial and fungal sequences.

As mentioned above, EPSP synthase is the site of action of glyphosate but despite a large amount of research the precise nature of this effect is not clear. It was shown for the EPSP synthase component of the *N. crassa arom* complex that the inhibition of EPSP synthase by glyphosate is competitive with PEP and uncompetitive with S3-P (Boocock and Coggins, 1983). Substrate binding is ordered, with S3-P binding first, followed by PEP (Amrhein *et al.*, 1981; Boocock and Coggins, 1983). As glyphosate is not a close analogue of PEP and does not inhibit other PEP-utilizing enzymes, it is thought that glyphosate acts as a transition-state analogue of PEP (Anton *et al.*, 1983; Amrhein, 1986). Recent studies on glyphosate and its complex with EPSP synthase using ^{13}C, ^{15}N and ^{31}P NMR chemical shift perturbations have revealed a generally linear conformation of glyphosate bound within the enzyme 'dead-end' complex (Enzyme-S3-P-glyphosate

complex) (Castellino *et al.*, 1989). Indeed, two phosphate analogues of the EPSP synthase tetrahedral intermediate have been synthesised which showed a K_i of 15nM for the *P. hybrida* EPSP synthase. These are the most potent inhibitors so far reported (Alberg and Bartlett, 1989).

Recently, cell cultures have been employed in a number of plant species to select for glyphosate tolerance and in each case examined, the herbicide tolerance was accompanied by an increased level of EPSP synthase activity (Nazfiger *et al.*, 1984; Smart *et al.*, 1985; Smith *et al.*, 1986; Steinrucken *et al.*, 1986; Dyer *et al.*, 1988). In glyphosate-tolerant cell lines of petunia, carrot and tobacco, overproduction of EPSP synthase has been shown to be the result of amplification of genes encoding this enzyme (Shah *et al.*, 1986; Hauptmann *et al.*, 1988; Goldsbrough *et al.*, 1990; Wang *et al.*, 1991).

1.3.7 Chorismate synthase

Chorismate synthase (EC 4.6.1.4) catalyses the formation of chorismate - the last precursor for all three aromatic amino acids - by the *trans* 1,4-elimination of *ortho*phosphate from EPSP (Gibson and Jackman, 1963; Gibson and Gibson, 1964; Hill and Newkome, 1969; Onderka and Floss, 1969; Floss *et al.*, 1972) (Step 7 in Fig. 1.1). The enzyme strictly requires a reduced flavin as cofactor, although no net change in redox state of the substrate is observed.

Chorismate synthase can either be monofunctional, as in plants (Mousdale and Coggins, 1986; Schaller *et al.*, 1990) and in bacteria (White *et al.*, 1988), or it can be bifunctional , as in *N. crassa* (Gaertner, 1987; White *et al.*, 1988) and *B. subtilis* (Hasan and Nester, 1978).

The *N. crassa* (Welch *et al.*, 1974), and *B. subtilis* (Hasan and Nester, 1978) chorismate synthase required reduced flavin and NADPH for activity, and possess an associated "flavin reductase" or "diaphorase" activity, which generates the required reduced flavin *in situ.* The *E. coli* enzyme does not possess this intrinsic diaphorase activity and can only catalyse the formation of chorismate when supplied with exogenous reduced flavin (Morrel *et al.*, 1967; White *et al.*, 1987).

The *aro*C gene encoding chorismate synthase in *E. coli* has been cloned, sequenced, and overexpressed and the enzyme was found to be a tetramer of subunit Mr 38,000 (White *et al.*, 1988). Comparison with the *N. crassa* enzyme, which is a tetramer of subunit Mr 50,000, suggest that the difference in size may be due to the "missing" flavin reductase activity (White *et al.*, 1988).

Monofunctional chorismate synthases have been purified and characterised from *Pisum sativum* (Mousdale and Coggins, 1986) and *Corydalis sempervirens* (Schaller *et al.*, 1990) and appear to be analogous to the *E. coli* type.

Recently *Euglena gracilis* chorismate synthase has been purified and characterised and found to be very similar to other chorismate synthases in terms of cofactor specificity, kinetic properties, isoelectric points and pH optima. Like that of bifunctional *N. crassa* enzyme (Welch *et al.*, 1974; White *et al.*, 1988) it also associated with flavin reductase activity (Schaller *et al.*, 1991).

The substrate analogue *(6S)* and *(6R)*-6-fluoro-5-enolpyruvylshikimate 3-phosphate have been shown to be potent inhibitors of *N. crassa* chorismate synthase (Balasubramanian *et al.*, 1991). The *(6R)* compound has also been shown to form a stable flavin mononucleotide free radical with the *E. coli* enzyme (Ramjee *et al.*, 1992).

1.4 The shikimate pathway in higher plants

The shikimate pathway is identical in plants and microorganisms, as evidenced by the demonstration of the individual reactions, the incorporation of intermediates and the purification of the relevant enzymes (Mousdale and Coggins, 1985; Kishore and Shah, 1988; Coggins, 1989; for recent reviews see Bentley. 1990; Mousdale and Coggins, 1992). It is accepted that this pathway is the primary route leading to aromatic amino acids, phenylquinones, lignins, and flavonoids in plants. Up to 60% of the ultimate plant mass (dry weight) may be comprised of molecules that once traversed this pathway (Jensen, 1986a).

In higher plants the subcelluar location of the pathway enzymes has not been fully resolved. Most of the earlier research was focused on the pathway itself and not its localisation (Koshiba, 1979; Polley, 1978).

The synthesis of aromatic amino acids by the plastidic shikimate pathway is now well established in photosynthetic tissues from the incorporation of $^{14}CO_2$ and [^{14}C] shikimate into tyrosine, phenylalanine, and tryptophan (Bickel *et al.*, 1978; Buchholz *et al.*, 1979; Buchholz and Schultz, 1980) by spinach chloroplast fractions. Later all the seven shikimate pathway enzymes, previously characterised in bacteria and fungi, were detected in the stroma of density gradient purified, washed chloroplasts from pea seedlings (Mousdale and Coggins, 1985), confirming that chloroplasts have all the enzymes for chorismate synthesis and are a major site for the biosynthesis of aromatic amino acids.

An additional, independently regulated shikimate pathway localised in the cytoplasm of higher plants has been proposed (Jensen, 1986b) which may provide precursor amino acids for the biosynthesis of protein in the cytoplasm. To date shikimate kinase (Schmid *et al.*, 1992), DAHP synthase (Dyer *et al.*, 1990), EPSP synthase (Gasser *et al.*, 1988;

Shah *et al.*, 1986), chorismate synthase (Schaller *et al.*, 1991), and the β subunit of the tryptophan synthase (Berlyn *et al.*, 1989; Last *et al.*, 1991) have been cloned from higher plant species. All the cDNAs and genes of these enzymes encode plastidic enzymes with targetting sequences, which lends little support to the Jensen hypothesis (Jensen, 1986b) that a complete cytoplasmic shikimate pathway exists for the biosynthesis of secondary metabolites in addition to the plastidic pathway. On the other hand, there is evidence for the occurrence of isozymes of the shikimate pathway enzymes (see Mousdale and Coggins, 1992 for a recent review) and in particular for the existence of cytosolic isozymes, which differ from the chloroplastic enzymes. The occurrence of these isozymes provides some support for a second, either complete or truncated, shikimate pathway as reported by Jensen (1986b). More knowledge is required to clarify the situation.

1.5 3-Dehydroquinase

3-Dehydroquinase catalyses the dehydration of 3-dehydroquinic acid to 3-dehydroshikimic acid in two separate metabolic pathways: one involved in biosynthesis and other involved in catabolism (Fig. 1.1 and 1.4) (Haslam, 1974; Giles *et al.*, 1985). The first dehydroquinase to be characterised was the biosynthetic enzyme which forms part of the shikimate pathway in *E. coli*; this enzyme catalyses the third step which involves the elimination of a water molecule with the introduction of first double bond into the aromatic ring system (Salamon and Davis, 1953; Srinivasan *et al.*, 1956; Hanson and Rose, 1963; Turner *et al.*, 1975). Later catabolic dehydroquinases, which form a part of the quinate pathway (see Section 1.5) of fungi, were characterised (Giles *et al.*, 1985).

1.5.1 Classification of dehydroquinases

The first distinction between two classes of dehydroquinases was made on the basis of their pathway origin. The biosynthetic enzyme from *E. coli*, was found to be heat labile and to consist of two identical subunits of Mr 27,500 (Chaudhuri *et al.*, 1986). In contrast the catabolic dehydroquinase of *N. crassa* was found to be heat stable and to be highly oligomeric consisting of twelve identical subunits of Mr 16,000 (Giles *et al.*, 1985). The *E. coli* - like biosynthetic dehydroquinases were classified as type I enzymes and the catabolic (*N. crassa* and *A. nidulans*) - type enzymes were classified as type II enzymes (White *et al.*, 1990). The two classes of enzyme were found to be unrelated at the level of primary structure (Garbe *et al.*, 1991). It was shown that the biosynthetic dehydroquinases of the *arom* multifunctional enzymes were closely related at the sequence level to the type I enzyme (Charles *et al.*, 1985; Duncan *et al.*, 1987; Chaudhuri *et al.*, 1991). Later White *et al.* found that there is another type of dehydroquinase in bacteria, which is also heat stable and very similar to the fungal catabolic dehydroquinase, both at level of quaternary structure and sequence (White *et al.*, 1990). Thus it seems that the type II dehydroquinases, previously thought to be exclusively catabolic, can also function in a biosynthetic context.

1.5.2 Reaction mechanism of the type I dehydroquinase

The type I enzyme from *E. coli* is the best studied enzyme of this class. Early experiments had shown that the elimination was *cis* (Hanson and Rose, 1963). Walsh (1979) proposed that the reaction proceeds through the formation of the Schiff base with the carbonyl group of substrate, 3-dehydroquinate, and the ε-amino group of an active site lysine residue followed by abstraction of the *pro*-R hydrogen at C-2 position by a general base (Fig. 1.3). The first evidence for the Schiff base involvement came in 1974, when

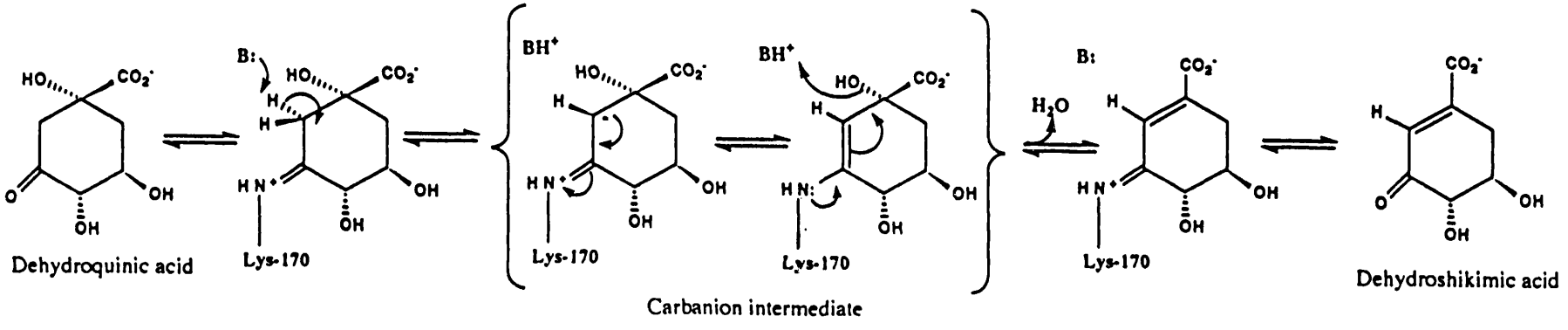


Fig. 1.3 Mechanism of the type I 3-dehydroquinase from *E. coli*

Butler and co-workers demonstrated that borohydride treatment of a crude preparation of dehydroquinase in the presence of substrate lead to inactivation of the enzyme (Butler *et al.*, 1974). The Walsh mechanism required the stereospecific deprotonation of the imine intermediate by a basic group on the enzyme. pH activity studies by Chaudhuri *et al.* (1986) suggested that a group with a pK_a of 6.1 was required in a deprotonated form for maximal activity. Inactivation of the enzyme by the histidine - directed reagent DEPC suggested that an imidazole side-chain carried out this deprotonation (Chaudhuri *et al.*, 1986). This residue is shown as B-H^+ in the proposed mechanism of Walsh (1979) (Fig. 1.3).

1.5.3 Structure-function studies of the type I dehydroquinases

In *E. coli*, dehydroquinase is encoded by the *aro*D gene. This gene has been precisely located, cloned and sequenced (Duncan *et al.*, 1986). It has also been overexpressed by cloning it into the high copy number plasmid pKD201 and as a result the enzyme is now available in hundreds of milligram quantities. The enzyme is a dimer of identical subunits; each chain consists of 252 amino acid residues (Chaudhuri *et al.*, 1991).

The lysine residue implicated by Butler *et al.* (1974) has been identified as Lys-170 by inactivation of the enzyme using sodium borotritide, followed by tryptic digestion and sequencing of the labelled peptide (Chaudhuri *et al.*, 1991). Recent electrospray mass spectrometry experiments have shown that only the product Schiff base intermediate (and not the substrate Schiff base) is trapped by borohydride (Shneier *et al.*, 1991), which may be a consequence of the equilibrium constant for the catalysed reaction (K_{eq}=15) (Mitsuhashi and Davis, 1954).

The histidine residue implicated by the pH/activity profile and DEPC-mediated inactivation of dehydroquinase (Chaudhuri *et al.*, 1986) has not been identified. An objective of this work was to identify the histidine residue involved in proton abstraction.

Two other residues which must be in or near the active sites have been identified by group-specific chemical modification of *E. coli* dehydroquinase with iodoacetate. These are Met-23 and Met-205 (Kleanthous *et al.*, 1990; Kleanthous and Coggins, 1990). It is not known whether these residues participate in the mechanism directly or are simply in the vicinity of the enzyme active site.

Bugg *et al.* (1988) examined the mechanism and active site structure of *E. coli* dehydroquinase with a series of substrate analogues incorporating reactive functional groups. Some of these analogues irreversibly inhibited the enzyme but it was not possible to identify the sites of modification.

In *N. crassa* the type I enzyme is part of the pentafunctional *arom* complex (Lumsden and Coggins, 1977; Smith and Coggins, 1983; Lambert *et al.*, 1985; Chaudhuri *et al.*, 1991). Like the *E. coli* enzyme, the dehydroquinase domain of the *N. crassa arom* complex works through an imine intermediate and the active site lysine residue has been identified (Chaudhuri *et al.*, 1991).

Other type I enzymes studied from *A. nidulans* (Charles *et al.*, 1985), *S. cerevisiae* (Duncan *et al.*, 1987) and *S. typhi* (Servos *et al.*, 1991) have similar characteristics to the *E. coli* and *N. crassa* enzymes. They are heat labile and they all contain a conserved active site lysine residue which is involved in Schiff base formation (Chaudhuri *et al.*, 1991).

The type I dehydroquinase from *S. typhi* has been crystallised. The crystals diffract to a resolution of 2.3 Å and suitable for high resolution structure determination (Boys *et al.*, 1992). This work is presently in progress.

1.5.4 Type II dehydroquinases

The heat stable, multimeric type II dehydroquinases have been studied both from the quinate pathway of fungi (Giles *et al.*, 1985) and the shikimate pathway of bacteria (White *et al.*, 1990; Garbe *et al.*, 1991). More recently, a dual function, quinate inducible type II dehydroquinase has been identified from *Amycolatropsis methanolica*, which is involved in both the catabolism of quinate and the biosynthesis of aromatic amino acids (Euverink *et al.*, 1992).

Unlike the type I enzymes, the type II enzymes are not inhibited by borohydride plus substrate and hence cannot work through the Schiff base mechanism. To learn about the mechanistic properties of this class of enzyme, some work was done on the type II dehydroquinase from *A. nidulans*.

1.5.5 Plant dehydroquinases

Dehydroquinase has been purified from a number of plants, and in all cases it co-purifies with the shikimate dehydrogenase activity (Polley, 1978; Fiedler and Schultz, 1985; Mousdale *et al.*, 1987). Purified dehydroquinase-shikimate dehydrogenases are monomeric with molecular weights varies from 48,000 to 60,000 according to plant species (Polley, 1978; Fiedler and Schultz, 1985; Mousdale *et al.*, 1987). No plant

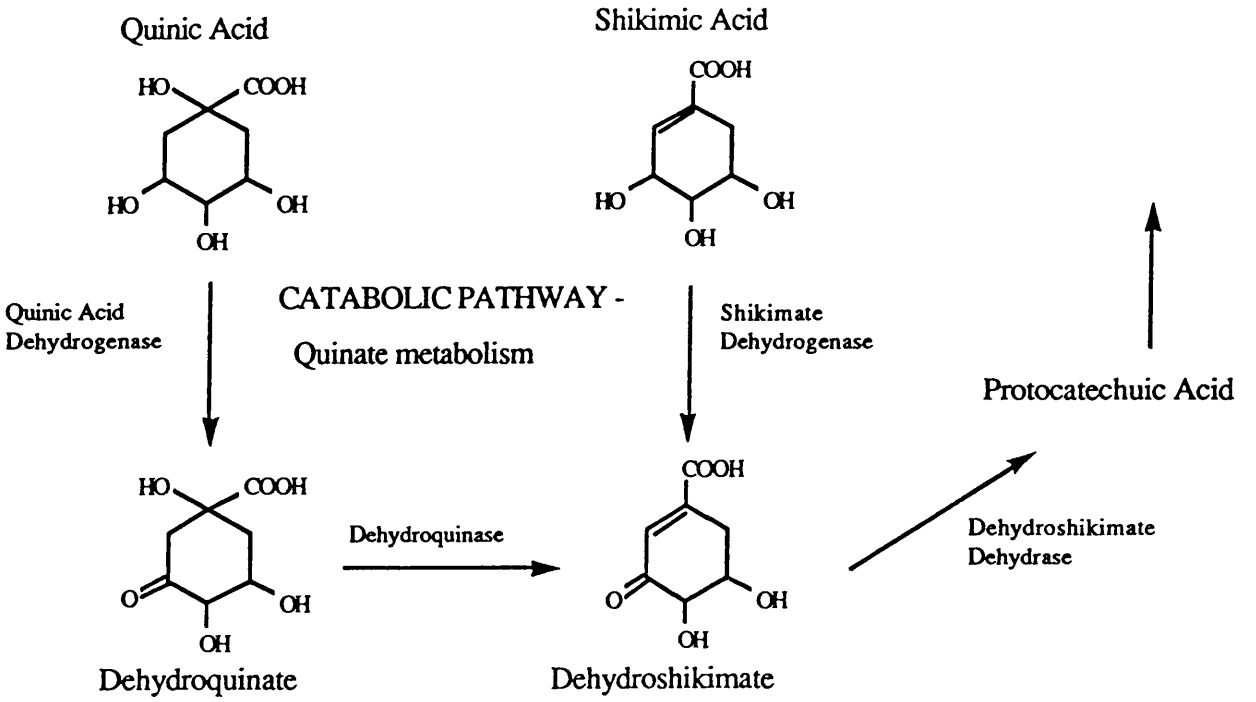


Fig. 1.4 The Quinate Pathway

dehydroquinase had been characterised at mechanistic or sequence level when this work began and therefore, one of the major objectives of this thesis was to characterise the *Pisum sativum* dehydroquinase activity.

1.6 The quinate pathway

Quinic acid is a very abundant plant metabolite and it comprises 2-10% of the dry weight of the leaves of higher plants (Bentley, 1990). Soil microganisms utilize this compound as a growth substrate via the quinate pathway (Giles *et al.*, 1985). This pathway has received considerable attention and is only discussed briefly here in the context of the type II dehydroquinase.

Three enzymes are involved in the conversion of quinic acid to protocatechuate in the following order: quinate dehydrogenase, 3-dehydroquinase, and dehydroshikimate dehydratase (Fig. 1.4). Organisms able to grow on both quinate and shikimate use an NADH-linked dehydrogenase to form 3-dehydroquinic acid and 3-dehydroshikimic acid respectively (Giles *et al.*, 1985). Dehydroquinase converts the dehydroquinate to dehydroshikimate; this reaction also occurs, of course, as part of the shikimate pathway for the biosynthesis of aromatic compounds. Part of the dehydroshikimate is used for the aromatic biosynthesis but part is converted to protocatechuate and and catabolised via the β-ketoadipate pathway (Giles *et al.*, 1985). Both *N. crassa* and *A. nidulans* contain a quinate inducible dehydroquinase (type II enzyme), in addition, to the biosynthetic dehydroquinase (type I enzyme). These are encoded by distinct genetic loci that are subject to separate physiological and genetical controls. The type I enzymes are encoded by the complex of *arom* locus which encodes the pentafunctional *arom* polypeptide chain (Lumsden and Coggins, 1977; Charles *et al.*, 1985). On the other hand, the type II quinate inducible dehydroquinases are encoded by the *qa*-2 in *N. crassa* and by the

QUTE gene in *A. nidulans*. These genes are part of a tightly linked cluster of seven genes which comprise the quinic acid utilization gene cluster in these species (Giles *et al.*, 1985; Grant *et al.*, 1988). Lamb *et al.*, (1991) have shown that the common metabolites 3-dehydroquinate and dehydroshikimate leak from the *arom* protein at a rate comparable with the flux of metabolites through the shikimate pathway and can be utilized by the enzymes of both pathways.

Recently it has shown that many nocardiform actinomycetes employ a single type II dehydroquinase (dual function) involved in both the catabolism of quinate and the biosynthesis of aromatic amino acids (Cain, 1981; Euverink *et al.*, 1992).

1.7 Aims of this project

The main objective of my Ph.D. project was to study 3-dehydroquinase both in plants and microorganisms. The specific aims were:

(1) ***To identify the general base histidine residue in the proposed mechanism of the type I dehydroquinase from Escherichia coli.***

(2) ***To study the mechanistic properties of the type II dehydroquinase from Aspergillus nidulans.***

(3) ***To scale-up the purification of the bifunctional dehydroquinase-shikimate dehydrogenase from Pisum sativum and to characterise the dehydroquinase activity.***

(4) ***To isolate a cDNA clone for at least the dehydroquinase domain of the bifunctional dehydroquinase-shikimate dehydrogenase.***

CHAPTER II

Materials and Methods

2.1 Materials and Reagents

2.1.1 Chemicals and biochemicals

Polyvinylidene difluoride (PVDF) membranes (Immobilon-P Transfer Membranes, as first described by Matsudaira, 1987) were obtained from Millipore, Bedford, Massachusetts, USA.

Ampicillin, benzamidine, CAPS (3-[cyclohexylamino]-1-propanesulfonic acid), Bis-tris, Bis-trispropane buffer, Pipes (1,4-piperazinediethanesulfonic acid) buffer, bromophenol blue, Coomassie brilliant blue R-250, ethidium bromide, shikimic acid, quinic acid, DEPC (diethylpyrocarbonate), hydroxylamine-hydrochloride (NH_2OH-HCl), sodium borohydride ($NaBH_4$), Na-salt of thioglycolic acid (mercaptoacetic acid), PLP (pyridoxal 5'-phosphate), phenylglyoxal (PG) and Amberlite MB-1 were obtained from Sigma Chemical Co., Poole, Dorset, UK.

Bactotryptone, yeast extract and Bactoagar (agar) were obtained from Difco, Detroit, USA.

DTT (dithiothreitol), $NADP^+$, NADPH, phenylmethanesulphonyl fluoride (PMSF), N,N,N',N'- tetramethylethylene diamine (TEMED), and Tris buffer were obtained from Boehringer Mannheim, Lewes, Sussex, UK.

Polyethylene glycol 6000 (Biochemical grade), were obtained from BDH Chemicals, Poole, Dorset, UK.

Agarose, low melting point (l.m.p.) agarose, isopropyl-β-D-thiogalactoside (IPTG), phenol (ultrapure), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) were obtained

from BRL, Gibco Ltd., Paisley, Scotland, UK.

N,N-Dimethylformamide was obtained from Aldrich Chemical Co. Ltd., Poole, Dorset, UK.

Acrylamide, bisacrylamide and SDS were obtained from FSA Laboratory Supplies, Loughborough, Leicestershire, UK.

Dehydroquinic acid was prepared by the oxidation of quinic acid (Sigma) following the procedure of Grewe and Haendler (1966) and the product isolated as the ammonium salt. The purity of the product was judged enzymatically or by h.p.l.c. (Mousdale and Coggins, 1985; Coggins *et al.*, 1987). The product was $\geq$ 95% pure and stored as a dry powder at -20°C.

HPLC grade acetonitrile was from Rathburn, Walkerburn, Scotland, UK and HPLC grade trifluoroacetic acid (TFA) and water were from Pierce and Warriner, Chester, Cheshire, UK. HPLC grade sodium phosphate was from BDH Chemicals.

[α-^{35}S]-dATP was obtained from Amersham International Plc., Amersham, Buckinghamshire, UK.

All other chemicals were of analytical reagent grade and were obtained from one of the following suppliers: Aldrich Chemical Co. Ltd.; BDH Chemicals; Formachem Ltd., Strathaven, Scotland, UK; FSA Laboratory Supplies; Koch-Light Ltd., Colnbrook, Buckinghamshire, UK.

2.1.2 Enzymes and proteins

Bovine serum albumin (BSA), DNase I, ribonuclease A (RNase) and lysozyme were obtained from Sigma Chemical Co.

A kit for molecular weight (standard markers) determination was obtained from Pharmacia, Milton Keynes, Buckinghamshire, UK.

Endoproteinase Glu-C (also called protease V8 from *Staphylococcus aureus*, Strain V8) sequencing grade was supplied by Boehringer Mannheim. TPCK treated bovine trypsin, TLCK treated bovine α-chymotrypsin and subtilisin were supplied by Sigma Chemical Co.

All restriction enzymes and T4 DNA ligase were obtained from BRL, Gibco Ltd. and used according to their instructions, unless otherwise indicated.

Taq DNA polymerase (with a Mg^{++} free buffer) was obtained from Promega, Chilworth Research Centre, Southampton, UK.

Deoxynucleotide triphosphate (dNTP) were obtained from Pharmacia.

2.1.3 Chromatography media

DEAE-Sephacel, 2'5'ADP-Sepharose® 4B and Sephadex® G-50 were supplied by Pharmacia. DEAE-cellulose (DE-52) was obtained from Whatman Biochemicals, Maidstone, Kent, UK. Bio-Gel Hydroxylapatite was purchased from Bio-Rad Laboratories Ltd., Watford, Hertfordshire, UK.

2.1.4 Prepacked media

Pharmacia columns PD-10 are prepacked, disposable columns containing Sephadex® G-25 M; these were used for rapid desalting and buffer exchange.

Prepacked Mono Q column was obtained from Pharmacia and utilized on a Pharmacia FPLC® System. μ Bondapak C18 reverse phase columns were purchased from Waters Chromatography, Watford, Hertfordshire, UK and attached to a Beckman System Gold HPLC apparatus (Beckman Instruments Inc., High Wycombe, Buckinghamshire, UK).

2.1.5 Bacterial strain, plasmid and bacteriophage

The bacterial strains are derivatives of *E.coli* K12. The bacterial strain used (for M13 propagation) during this study is shown below:

E. coli TG1	*sup*E *hsd*Δ5 *thi* Δ(*lac-proAB*)	
	F' [*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]	Gibson, 1984

Bacterial stocks were stored at -70°C as a frozen stock, prepared by mixing 1ml of over night L-Broth (LB) culture with an equal amount of sterile glycerol. Short term storage was on a suitable sealed plate at 4°C or as a 10ml LB culture at 4°C.

The plasmid pAT153 and λ-DNA were cleaved with appropriate restriction enzymes and the mixture of fragments generated used as molecular weight size markers during agarose gel electrophoresis of DNA. λ/*Bst* EII obtained from Appligene, Chester-Le-Street, Durham, UK, was also used as a DNA molecular weight marker.

M13 mp18 and mp19 vectors (bacteriophage) (Yanisch-Perron *et al.*, 1985) were used for DNA sequencing.

2.2 General Laboratory methods

General molecular biological techniques were carried out essentially as described in Sambrook *et al.* (1989) and the general methods for handling proteins and enzymes were as described in *Methods in Enzymology*: *Guide to protein purification* , Vol.182 pp.19-83. (Murray P. Deutscher, ed., 1990).

2.2.1 pH measurement

pH measurements were made with a Radiometer Model 26 pH meter (Copenhagen, Denmark), using a Russell (Auchtermuchty, Scotland, UK) combination electrode calibrated at room temperature.

2.2.2 Conductivity measurement

Conductivity measurements were made at 4°C with a Radiometer Model CDM2e conductivity meter.

2.2.3 Protein estimation

Protein was determined by the method of Bradford (1976), with BSA as standard.

2.2.4 Lyophilization

Solutions of proteins, peptides, etc. were collected either into Reacti Vials (Pierce), polypropylene tubes (Eppendorf®) or 25-1000ml acid washed round-bottom flasks and the contents frozen by immersing the vessels in a dry-ice/ethanol mixture before lyophilisation on a FTS® Systems (Stone Ridge, New York, USA) Flexi-Dry freeze dryer.

2.2.5 Desalting of proteins

Protein samples were rapidly desalted or buffer exchanged using Pharmacia prepacked PD-10 columns (Sephadex® G-25 M) according to suppliers instructions.

2.2.6 Spectrophotometric measurement of nucleic acids

Nucleic acid concentrations were determined spectrophotometrically at 260nm (Sambrook *et al.*, 1989) in a Phillips Model PU8700 spectrophotometer using quartz cuvettes. In a 1 cm path length cuvette an absorbance value of 1.0 corresponds to 50μg/ml for double-stranded DNA and 20 μg/ml for single-stranded oligonucleotides.

An RNA solution with an A_{260} of 1.0 was taken to have a concentration of 40μg/ml.

2.2.7 Siliconizing glassware

Glassware was siliconized (in a fume hood) with 2% (v/v) DMDS in chloroform (for 30-60min), and thoroughly rinsed with distilled water before baking at 100°C for 2-3h.

2.3 Polyacrylamide gel electrophoresis

Electrophoresis in the presence of SDS was performed by the method of Laemmli (1970), with a 3% stacking gel and a 10% running gel. The ratio of acrylamide : bis-acrylamide in all PAGE experiments was 30 : 0.8 and polymerization was induced by 0.03% (v/v) TEMED and 0.05% (w/v) ammonium persulphate. After electrophoresis gels were stained for protein by the Coomassie method (Section 2.3.1).

Sometimes Phast Gels (Pharmacia) were used for denaturing electrophoresis, utilizing the Pharmacia Phast-System apparatus with "SDS" buffer strips obtained from the supplier.

2.3.1 Protein staining

Protein was visualised on gels by staining with Coomassie blue. The Coomassie reagent was 0.1% (w/v) Coomassie brilliant blue R-250 in 10% (v/v) glacial acetic acid, 50% (v/v) methanol; the destaining reagent was 10% (v/v) acetic acid, 10% (v/v) methanol. These procedures were carried out at 40°C.

2.4 Enzyme assays

2.4.1 3-Dehydroquinase

3-Dehydroquinase (DHQase) activity was determined by monitoring the formation of dehydroshikimate at 234nm ($\varepsilon = 12X10^3M^{-1}\ cm^{-1}$) (Salamon and Davis, 1953; Mitsuhashi and Davis, 1954) at 25°C.

The type I DHQase (*E. coli*) assay mixture contained 100mM-potassium phosphate, pH 7.0 and 100μM ammonium dehydroquinate as substrate.

P. sativum DHQase was assayed in a similar way except 200μM substrate was used.

For the type II DHQase (*A. nidulans*), the assay mixture contained 50mM Tris/acetate, pH 7.0 and 1mM substrate.

2.4.2 Shikimate dehydrogenase

Shikimate dehydrogenase (SDHase) activity was measured in the reverse direction by following the increase in absorbance (due to NADPH formation) at 340nm ($\varepsilon=6.2X10^3M^{-1}cm^{-1}$) during the oxidation of shikimate to dehydroshikimate at 25°C. The assay mixture contained 100mM-sodium carbonate/bicarbonate buffer, pH 10.6, 4mM-shikimic acid, 2mM-$NADP^+$.

2.5 Growth media

Unless otherwise stated all media were sterilised by autoclaving at 15 psi for 30min.

2.5.1 Rich media

The 'rich' media used for growth of microorganisms during the course of this project are listed below (the abbreviations LB and LA are used extensively in the text for L-broth and L-agar):

Where appropriate, glucose (20% v/v) was autoclaved separately (5psi for 50min) and added to the correct final concentration.

Medium (liquid)	Composition (per litre)
L-broth (LB)	10 g bactotryptone 5 g yeast extract 10 g NaCl (+ 5 ml 20% (w/v) glucose)
2 x YT	16 g bactotryptone 10 g yeast extract 5 g NaCl

Medium (solid)	Composition (per litre)
L-agar	as LB, + 15 g agar
H-agar	10g bactotryptone 8g NaCl 12g agar
H-top agar	10g bactotryptone 8g NaCl 8g agar

2.5.2 Minimal media

Minimal medium (MM or M-9 salts) was prepared as follows:

M-9 salts (minimal medium)	6g Na_2HPO_4 (MW 131.96), 3g KH_2PO_4 (MW 136.09), 0.13g $MgSO_4 \cdot 7H_2O$, 1g NH_4Cl, H_2O to 1 litre
M-9 agar) (minimal agar plates)	As M-9 salts + 15g agar (Oxoid)

To prepare plates the M-9 agar medium was autoclaved to sterilize, cooled to approximately 55°-60°C and then 1ml 100mM $CaCl_2$ (previously autoclaved), 10ml 20% (w/v) glucose (previously autoclaved), and 2ml 10mg/ml thiamine (filter sterilized) were added and the volume adjusted to 1000ml with sterile water before pouring the plates. The plates can be stored for three months at 4°C.

2.5.3 Selection supplements

The following supplements were added to rich media to select and identify recombinant organisms.

(a) antibiotics: Ampicillin (Amp) was used at a final concentration of 50 μg/ml. A stock solution of 25 mg/ml was filter sterilised and stored at -20°C. Hot L-agar was cooled to 55°C before ampicillin was added. L-Amp plates were stable for at least 4 weeks if stored at 4°C.

(b) chromogenic substrates: IPTG was used at a final concentration of 20 μg/ml. A stock solution of IPTG of 20 mg/ml was filter sterilised and stored at -20°C. A stock solution of X-gal (20%) was made up in *N,N*-dimethylformamide and stored at -20°C.

2.6 Enzyme preparation

2.6.1 Type I DHQase from *E. coli*

DHQase was purified from an overproducing strain of *E. coli* AB2848/pKD203, according to the procedure of Chaudhuri *et al.* (1986).

The protein concentration of a sample of the purified DHQase was determined by amino acid analysis, and the extinction coefficient at 280nm was found to be 0.69mg/ml.

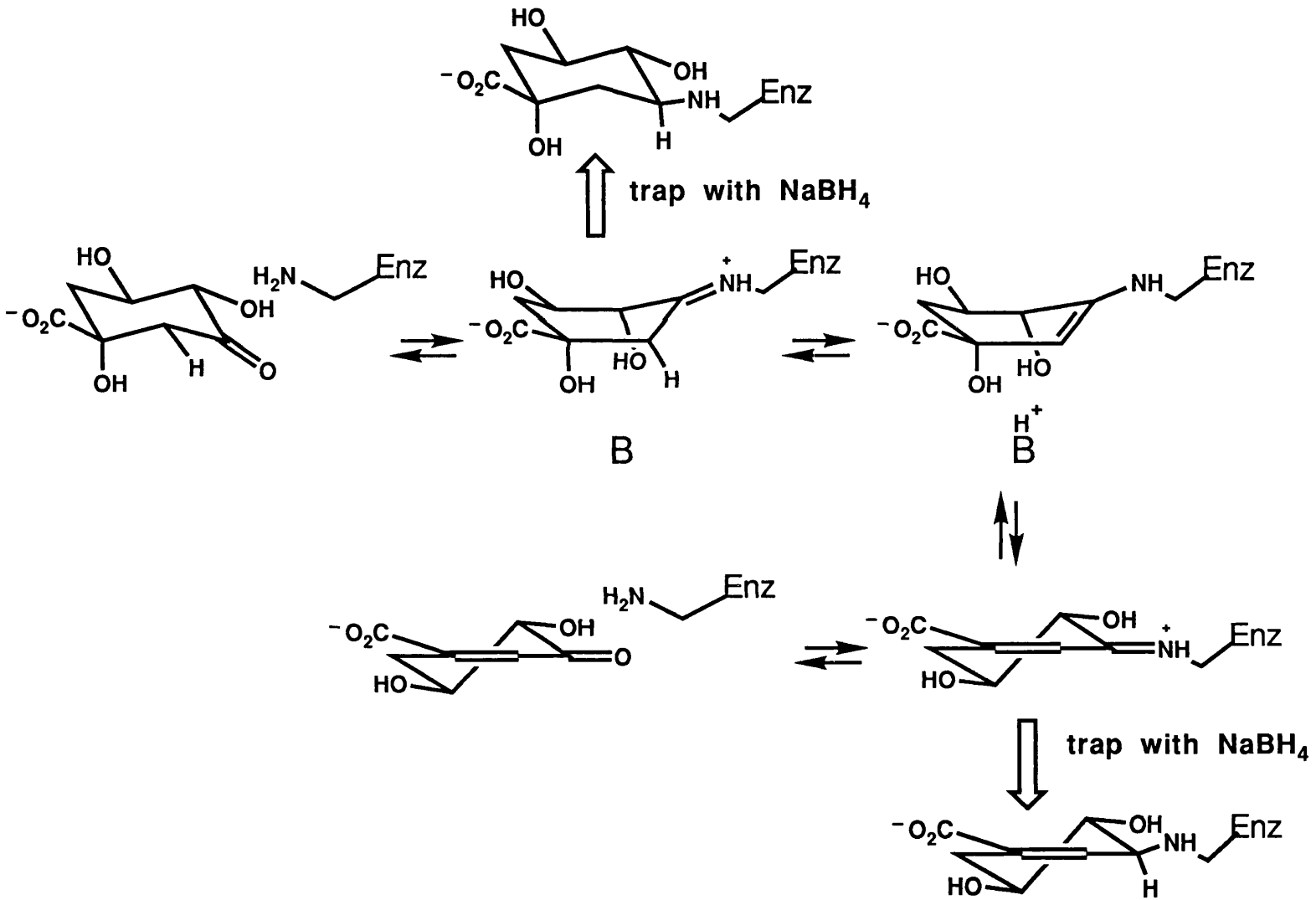
2.6.2 Type II DHQase from *A. nidulans*

The type II DHQase from *A. nidulans* was overexpressed in *E. coli* and purified as described by Beri *et al.* (1990). There was no contamination from host type I DHQase as the purification scheme used a heat treatment step (70°C for 15 min), which denatures all the type I enzyme but does not affect the type II enzyme.

2.6.3 Preparation of substrate-trapped DHQase

Reduction of the enzyme-bound Schiff base intermediate was carried out at room temperature as described below using sodium borohydride. Several aliquots (50μl) of

Fig. 2.1 Trapping of Schiff base intermediates with sodium borohydride
Large arrows signify the trapping of imine intermediates with sodium borohydride



sodium borohydride (20mg/ml dissolved in 40mM NaOH) were added to two separate samples of DHQase (1.7mg/ml) which had been dialyzed into potassium phosphate buffer (50mM pH 7.0). To one sample was added 1mM dehydroquinate (in water) 5min prior to the first addition of borohydride and to the other was added water, to serve as control for the addition of the reducing agent. Three aliquots of borohydride were required to inactivate the enzyme to less than 1% of the starting activity. Activity loss was due to the reduction of the Schiff base formed at lysine-170 of DHQase (Chaudhuri *et al.*, 1991). No activity was lost in the water control. Excess ligand was removed by dialysis. DHQase is a single substrate enzyme, and the equilibrium constant has been estimated as 15 (Mitsuhashi and Davis, 1954). Addition of dehydroquinate prior to the addition of borohydride would be expected to yield a mixed population of inactivated DHQase molecules, some bound with substrate and some with product (Fig. 2.1). Through out this thesis these species will be referred as either ligand-linked or substrate-trapped DHQase.

2.7 Characterization of DEPC-modified DHQase from *E. coli*

2.7.1 Determination of DEPC concentration

The concentration of stock DEPC was determined by reaction with 10mM imidazole in 50mM potassium phosphate buffer (pH 7.5) and the resulting increase in absorbance at 240nm (ε = 3200 M^{-1} cm^{-1}) (Ovadi *et al.*, 1967). The reagent was freshly diluted with ice-cold absolute ethanol before each experiment.

2.7.2 DEPC inactivation reaction

Inactivation was carried out by incubating DHQase (2-3μM monomer concentration) with DEPC in 50mM potassium phosphate buffer, pH6.0 at 25°C. The final concentration of ethanol in the reaction mixture (1ml) was kept below 5% (v/v), concentrations of which do not affect enzyme activity. Aliquots (10μl) were withdrawn at various time intervals and quenched with 40μl, 100 mM ice-cold imidazole buffer (pH7.0) followed by dilution with 100mM ice-cold potassium phosphate (pH 7.0) and residual enzyme activity determined as in Section 2.4.

2.7.3 Substrate protection against DEPC inactivation

Protection experiments simply included increasing concentrations of ammonium dehydroquinate dissolved in water. Since DHQase is a single substrate enzyme an equilibrium mixture (1:15; Mitsuhashi and Davis, 1954) of substrate and product is generated which is referred to as protecting ligand.

2.7.4 Reversal of DEPC inactivation

For reversal experiments (decarbethoxylation), the protein samples previously treated with DEPC (in phosphate buffer, pH6.0) were incubated with various concentrations of hydroxylamine (prepared as an aqueous stock solution of the hydrochloride adjusted to pH 7.0 with KOH) for different time intervals at room temperature or at 4°C. Excess hydroxylamine, residual DEPC, and other low molecular weight reactants were then removed either by dialysis or passage through the PD10 column. Control samples were similarly treated with hydroxylamine but contained no DEPC (only ethanol).

2.7.5 pH dependence of DEPC inactivation

The pH dependence of DHQase inactivation by DEPC was determined in 50mM potassium phosphate buffer over the pH range 6.0-7.1 at 25°C.

2.7.6 CD analysis

CD spectra were obtained with a Jasco J-600 automatic recording dichrograph at room temparature. A cell of 1cm path length was used for the measurements. The mean residue weight was calculated to be 108.99 (27466/252 = 108.99) from the amino acid sequence (Chaudhuri *et al.*, 1991) and the mass determined for purified DHQase by electrospray mass spectrometry (Shneier, *et al.*,1991). The specta were recorded by Ms Sharon Kelly in the laboratory of Dr. Nicholas Price at the University of Stirling.

2.7.7 Difference spectroscopy and stoichiometry of carbethoxylation

The extent of N-carbethoxylation of histidine residues was determined by the time dependent increase in the absorbance at 240nm in a 1ml reaction volume in 50mM potassium phosphate buffer (pH6.0, 25°C). A single beam spectrophotometer was used (Phillips/PU8700) equipped with a data storage facility and a thermostatted cuvette holder. Reference spectra (of buffer, enzyme, and ethanol) were collected, over the wavelength range 220nm-350nm, and then subtracted from the sample spectra (containing the same components plus DEPC) collected over the same wavelength range. The number of the modified histidine residues, as a function of time, was calculated from the molar extinction coefficient for N-carbethoxyhistidine (ε_{240} = 3200 $M^{-1}cm^{-1}$; Ovadi *et al.*, 1967).

2.7.8 Borohydride reduction of DHQase followed by DEPC modification

Reduction of Schiff base bound ligand to the enzyme was as described in Section 2.7.3 using sodium borohydride. Excess ligand was removed by dialysis against 50mM potassium phosphate buffer (pH6.0) and the extent of DEPC histidine modification monitored by difference spectroscopy, as described above (Section 2.7.7).

2.8 Peptide isolation and sequencing

2.8.1 Acid-washed glassware

Glassware for protein chemistry was boiled for 1h in 6N HCl or 6N HNO_3 then left steeping overnight. It was then rinsed exhaustively with distilled water and dried in a hot oven. Glass Reacti Vials (Pierce) were used for purified peptide collection and freeze drying.

2.8.2 Preparation of DEPC-modified DHQase

DEPC-modified proteins (2.5mg) both in the presence and absence of added substrate (ammonium dehydroquinate, 2mM final concentration) were prepared by adding a 30-fold molar excess of DEPC over protein monomer in 50mM potassium phosphate buffer, pH6.0 at 25°C. After 30min incubation, >96% activity was lost from the sample incubated in the absence of substrate, whereas 85% of the protected sample remained active.

The DEPC-modified proteins were dialysed against 50mM sodium phosphate buffer, pH 7.0, for 4-6h with several changes of buffer. The residual activity of both samples

remained unaltered by dialysis. Samples were then digested with proteinases as described in Section 2.8.3.

2.8.3 Digestion of DEPC-modified DHQase

Aliquots of DEPC-modified DHQase (both in the presence and absence of substrate) in 50mM sodium phosphate buffer, pH 7.0 were digested with trypsin, chymotrypsin and substilisin (protease : DHQase, 2% w/w) at 30°C. Digestion was carried out in a round bottom acid-washed flask containing a small magnet (acid-washed) for continuous stiring during digestion. At intervals, samples were withdrawn and immediately frozen (dry-ice/ethanol) and stored at -20°C or directly injected onto a reverse phase HPLC column (μ Bondapack C18 column) for separation.

Double digestion: Protein digested with chymotrypsin for $2^1/2$h were treated with TPCK (1.1 fold excess over chymotrypsin), then trypsin was added to the sample and digestion was carried out for another 1h as described above. Trypsin inhibitor was then added to the reaction mixture and the mixture was frozen and stored at -20°C.

2.8.3 Reverse phase HPLC of peptides

Two HPLC separation conditions were used for the purification of the active site peptide. Both utilized a C18 μBondapak reverse phase column (Waters) but differed in the solvent conditions. Peptides were first fractionated using **solvent system I**, where the column was equilibrated in 1mM sodium phosphate in water (pH6.3) and developed with a mixture (70:30, v/v) of 1mM sodium phosphate, pH6.3 and acetonitrile at a flow rate of 1ml/min. Eluting peptides were monitored spectrophotometrically at 220nm and 240nm to

detect peptide bonds and DEPC modified peptides, respectively. Isolated peptide fractions were collected manually, freeze-dried and further purified by **solvent system II**, where the column was equilibrated in a mixture 0.15% trifluoroacetic acid in water (pH 2) and acetonitrile (90:10, v/v) and developed with a linear gradient of this solvent and acetonitrile containing 0.15% trifluoroacetic acid at a flow rate of 1ml/min. Detection of peptides was at 220nm only. Pure peptide fractions were collected manually, freeze dried and stored at -20°C.

2.8.4 Sequence analysis of the active site peptide

The purified active site peptide was sequenced on an Applied Biosystems Model 470A Gas-phase sequencer with on-line PTH analysis (model 120A analyser) by Dr. David Campbell at the University of Dundee.

2.9 Chemical modification of type II DHQase from *A. nidulans*

2.9.1 DEPC inactivation

Purified enzyme solution (generally 1ml) in 50mM sodium bicarbonate buffer, pH 8.5 were treated at 25°C with various concentrations of DEPC. Aliquots were withdrawn and quenched with ice-cold imidazole buffer as described in Section 2.7.2.

The pH dependence of inactivation was investigated in 50mM Pipes buffer in the pH range 6.5 - 7.5.

2.9.2 Phenylglyoxal inactivation

An aliquot of the enzyme was incubated at 25°C alone or in the presence of DHQ, in 100mM sodium bicarbonate buffer, pH 9.4 for ~5min. Phenylglyoxal (PG) was then added to the incubation mixtures and aliquots were subsequently removed at time intervals for assay of DHQase activity.

2.9.3 Pyridoxal 5'-phosphate inactivation

Pyridoxal 5'-phosphate (PLP) inactivation of type II DHQase was carried out in 50mM Tris/acetate, pH 7.0 and at 25°C. PLP was freshly prepared as a 50mM stock in 50mM Tris/acetate, pH 7.0 and protected from light. 1ml reaction mixtures contained 940µl 50mM Tris/acetate pH 7.0, 20µl enzyme (~5mg/ml) and 20µl 50mM PLP. After 30min of incubation another 20µl PLP solution was added to the reaction mixture. Controls were carried out by adding Tris buffer instead of PLP. At time intervals aliquots (10µl) were withdrawn and diluted to 1ml with ice-cold 50mM-Tris/acetate, pH 7.0. The remaining activity was determined as described in Section 2.4.

2.9.4 Attempted borohydride inactivation

Attempted borohydride inactivation of type II DHQase was carried out as described in Section 2.6.3 and in 50mM sodium bicarbonate buffer, pH 8.5.

2.10 Growth of *P. sativum* for DHQase-SDHase purification

As described in Chapter V (Section 5.2)

2.10.1 DHQase-SDHase preparation

Enzyme was purified by modifying the previous protocol described by Mousdale *et al.*, 1987 as described in Chapter V (Section 5.3).

2.11 Characterisation of DHQase-SDHase

2.11.1 Molecular weight determination

SDS PAGE was used to estimate the subunit Mr of the purified protein. A low molecular weight (LMW) calibration kit (Pharmacia) was used to produce standard curves of Rf against log Mr. The standard molecular weight proteins were chosen from the kit given below:

Protein	**Subunit Mr**
Rabbit muscle hosphorylase b	94,000
Bovine serum albumin	67,000
Egg white ovalbumin	43,000
Bovine erythrocyte carbonic anhydrase	30,000
Soybean trypsin inhibitor	20,100
Bovine milk α-lactalbumin	14,400

2.11.2 Inactivation of DHQase-SDHase with DEPC

500µl reaction mixtures containing DHQase-SDHase in 100mM-potassium phosphate, pH 6.0 at 20°C were incubated with DEPC freshly diluted in ice-cold ethanol. The final ethanol concentration was below 5%. Aliquots (50µl) were taken at intervals and quenched with an equal volume of 100mM ice-cold imidazole buffer, pH 7.0 and assayed for both DHQase and SDHase activities. Control experiments were performed with ethanol alone. Substrate protection experiments were carried out in the same manner by adding the substrates dehydroquinate and shikimic acid/$NADP^+$ before the addition of DEPC.

2.11.3 Inactivation of DHQase-SDHase with ammonium dehydroquinate and $NaBH_4$

Samples (100µl) containing DHQase-SDHase in 20mM-Tris/HCl, pH 7.5, 1mM-BE, 0.5mM-DTT and 1mM-ammonium dehydroquinate were incubated at 20°C for 5min before the addition of 10ml freshly prepared $NaBH_4$ (20mg/ml) in 40mM NaOH. Controls were carried out in a similar way with the exception that control I received 10 ml 40mM NaOH instead of $NaBH_4$; control II: protein only; control III did not contain ammonium dehydroquinate.

2.12 Protein and peptide microsequencing

Peptide fragments were electroblotted onto polyvinylidene difluoride (PVDF) membrane and subjected to microsequencing (Matsudaira, 1987) using an Applied Biosystem 477A pulsed-liquid protein sequencer by Dr. Brian Dunbar at the University of Aberdeen. Electroblotted intact protein was sequenced by the same method.

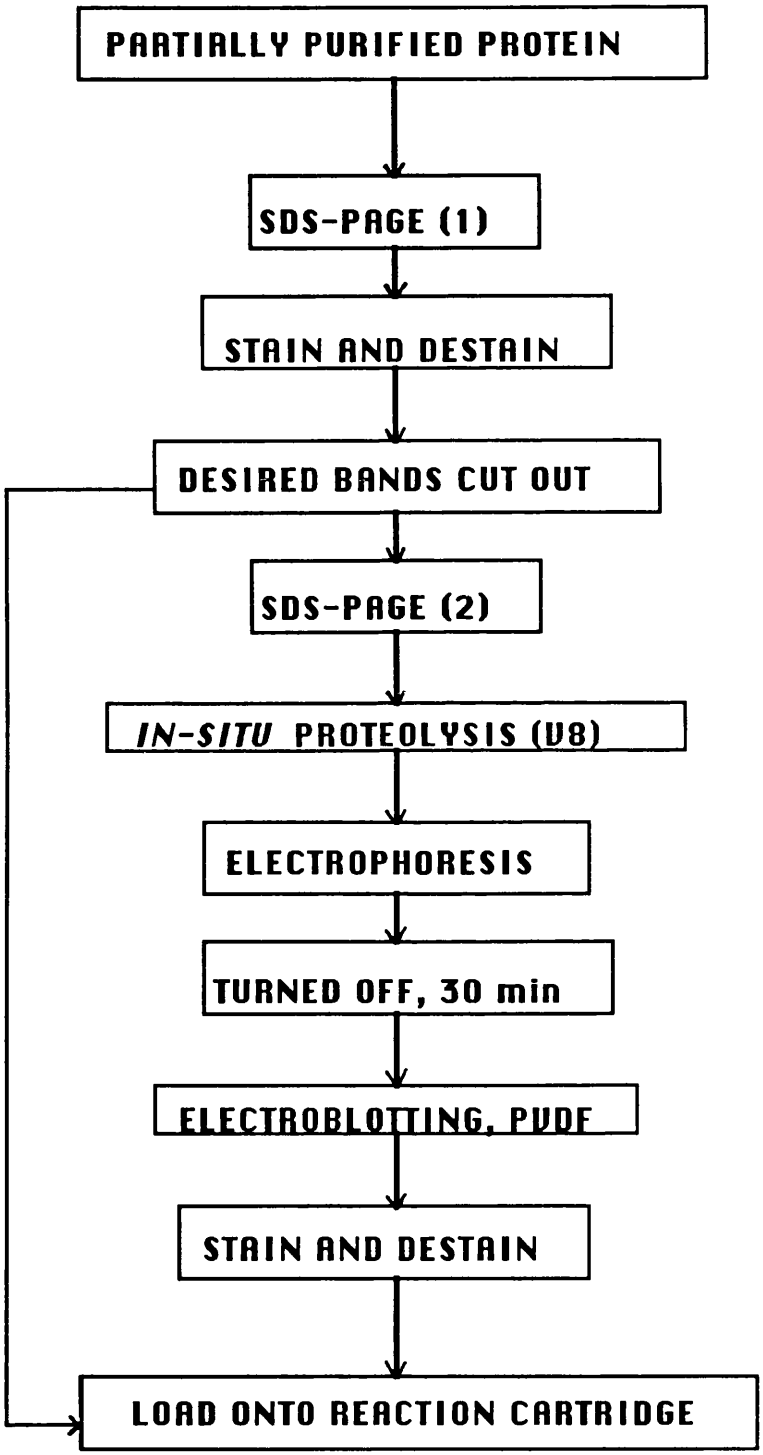


Fig. 2. 2 A flow chart for sequencing of electroblotted protein/peptide

2.12.1 Electroblotting of SDS-protein onto PVDF membrane

The method of peptide mapping by limited *in situ* proteolysis (Cleveland *et al.*, 1977) and the process of transferring protein from gels to PVDF membrane have been described previously (Matsudaira, 1987). The actual procedure used for the sequence analysis of the *Pisum sativum* DHQase-SDHase (Fig. 2.2) is given below:

[All the solutions were prepared in Millipore filtered water]

a) Electroblotting buffer: 10mM CAPS (3-[cyclohexylamino]-1-propanesulfonic acid), 10% methanol - titrated to pH 11.0 with 2M NaOH and stored at room temperature. This buffer has a low ionic strength and is suitable for the electrophoretic transfer, permitting the high voltages required for effective transfer.

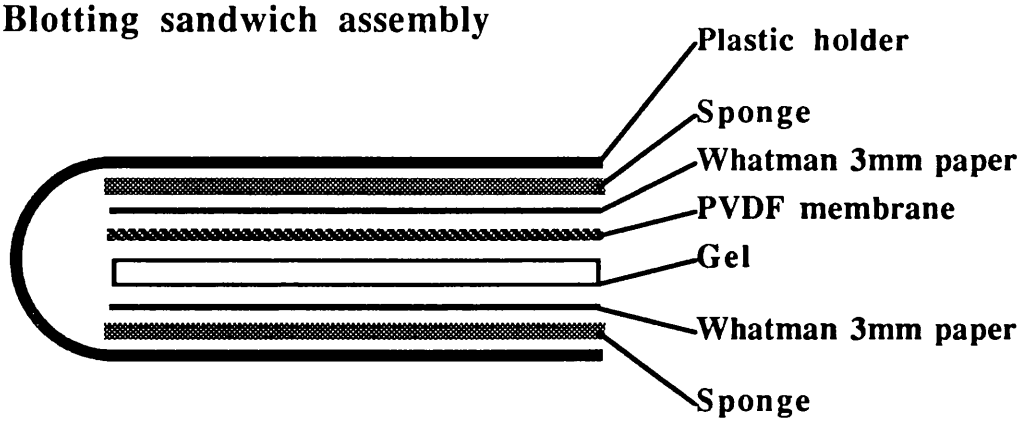
b) The PVDF membrane is cut to the size of the gel and wetted with 100% methanol for 3-5sec., rinsed with 50% methanol and placed in a petri dish containing blotting buffer.

c) After slab-gel (1.5mm thickness) electrophoresis in the presence of SDS (Laemmli, 1970) gels were removed and soaked in the electroblotting buffer for ~5 min to reduce the amount of Tris and glycine.

d) Pieces of Whatman 3mm paper were cut to the size of the gel and soaked in the blotting buffer.

e) The gel, sandwiched between a sheet of PVDF membrane and sheets of Whatman 3mm paper, was assembled in the blotting apparatus (Bio-Rad Transblot Apparatus) as shown in Fig. 2.3. One layer of pre-wetted Whatman paper was placed on the lower graphite electrode with excess of blotting buffer. Then the gel was carefully laid on top

A) Blotting sandwich assembly



B) Orientation of sandwich in tank

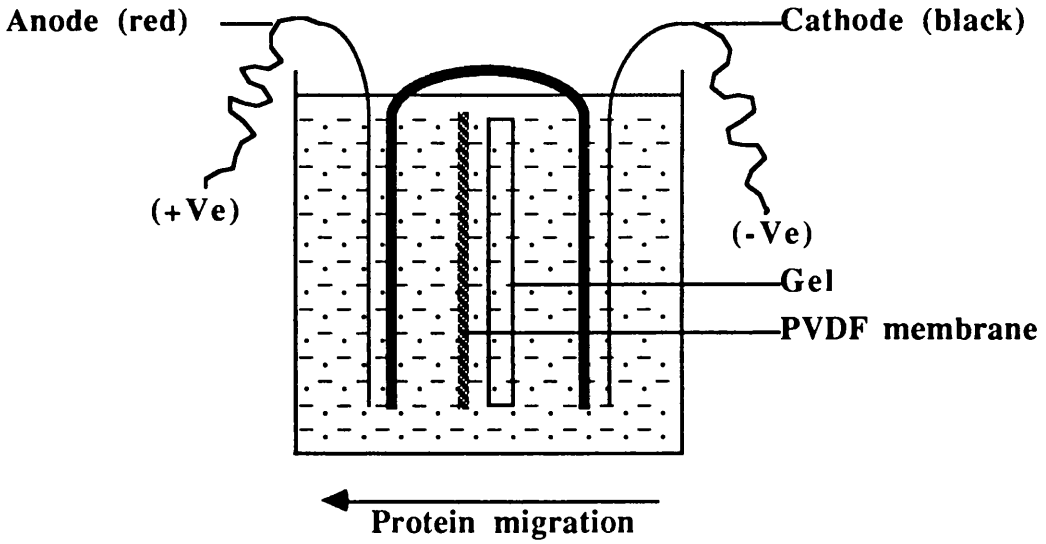


Fig. 2.3 Electroblotting of SDS-protein onto PVDF membrane
A) Assembly of the blotting 'sandwich'. B) Orientation of the 'sandwich' in the blotting tank.

ensuring that no air was trapped. One corner of the PVDF membrane was cut off to mark the orientation and then the membrane was placed on the gel excluding air. Finally, a piece of Whatman paper was placed on the top. The upper electrode was positioned and the sandwich placed in the blotting apparatus in the right orientation (*i.e.*. membrane to wards anode). The electroblotting was carried out at 50V, room temperature for 60-90min. The membrane was removed from the sandwich and rinsed with water, then dipped in 100% methanol for 5min prior to staining.

2.12.2 Staining of the blotted protein

The PVDF membrane was stained with 0.1% Coomassie Blue R-250 in 50% methanol for 2-5min, and then destained in 50% methanol, 10% acetic acid for 5-10min at room temperature (at least three changes). The membrane was finally rinsed in water for 5-10min (several changes), air-dried, and stored at -20°C.

2.12.3 *In situ* proteolysis of DHQase-SDHase with V8 protease

Mixtures of proteins can be separated on SDS PAGE and individual bands cut out for further analysis.

Almost homogeneous DHQase-SDHase was completely purified for microsequencing by SDS PAGE in a 10% gel. The bands were visualised by staining with Coomassie Blue R-250 for 5min followed by destaining for 20min (briefly to avoid possible acid hydrolysis during staining and destaining). The band corresponding to DHQase-SDHase was excised with a scalpel and neutralized in two changes of buffer containing 125mM-Tris/HCl, pH

6.8, 0.1% SDS, 1mM-DTT (with gentle swirling). The band was stored at -20°C in the same buffer until ready to use.

Gel slices were loaded onto a second 18% SDS gel, and the proteins were digested with V8 proteinase (Glu-C, Boehringer) during electrophoresis as described by Cleveland *et al.* (1977).

Gels were aged at least 20hrs before use and pre-run for 45min by adding 0.1mM sodium thioglycollate (free radical scavenger) to the upper tank buffer. This was done as a precautionary measure to avoid any free radical that might damage the protein. Gels were 1.5mm thick and cast with longer (5cm) stacking gels and wider sample wells.

The sample wells were filled with Cleveland buffer (125mM-Tris/HCl, pH 6.8, 0.1% SDS, 1mM-DTT) and each gel slice was pushed to the bottom of a well with a spatula. Spaces around the slices were filled by overlaying each slice with 20µl Cleveland buffer containing 10% glycerol. Finally, 10µl Cleveland buffer containing 1µg V8, 10% glycerol and a tint of bromophenol blue was added (the dye allows the progress of the sample in the gel to be followed). The gel was run at 40mA until the bromophenol blue dye front reached the interface of the stacking and resolving gel. The current was turned off for 30min to allow proteolysis to occur and then electrophoresis was performed as usual for a Laemmli gel. After electrophoresis the protein bands were transferred electrophoretically onto a PVDF membrane as described in Section 2.12.1 and then the membrane was stained to visualize the bands as described in Section 2.12.2.

2.12.4 Sequencing of the blotted protein

The membrane was centered on the Teflon seal and placed in the cartridge block of the sequenator (specially designed for PVDF bound samples) and sequenced on an Applied Biosystems model 477A pulsed-liquid sequenator equipped with on-line PTH analysis by Dr. Brian Dunbar of Aberdeen University.

2.13 Amino acid composition analysis of DHQase-SDHase

Amino acid compositions were determined by using an Applied Biosystems Model 420 amino acid analyser. Norleucine was included as an internal standard, the sample was automatically acid hydrolysed and converted to PITC amino acid derivatives which were analysed following reverse phase chromatography.

2.14 Plant material for RNA extraction

Plants were grown in a green house as described in Chapter V (Section 5.2).

2.15 Preparation of *P. sativum* total RNA

Total RNA was extracted from freshly harvested 11 day-old-pea shoots by a phenol extraction and LiCl-precipitation method (Sambrook *et al.*, 1989).

2.15.1 RNase-free technique

The major problem in working with RNA is contamination of the sample with RNase. RNases are very stable enzymes which do not require cofactors for activity and, if appropriate precautions are not followed, will degrade the RNA before it is isolated. Therefore, RNase-free conditions must be maintained throughout the purification to assure the integrity of the RNA, since contamination of samples with even small quantities of RNase can lead to degradation.

Precautions for working with RNA

Glassware: All glassware used in the preparation of RNA was treated with 0.1% (w/v) DEPC before baking for 16-24h at $\geq$ 200°C.

Plasticware: Sterile, disposable plasticware straight out of the package is considered RNase-free. New, non-sterile plasticware was autoclaved at 20 p.s.i. for 15 min.

Solutions: All solutions used in the preparation of RNA were made RNase-free by treatment with DEPC. To treat solutions 0.1ml DEPC was added per 100ml of solution and after vigorous mixing the solutions were allowed to stand overnight at room temperature and then autoclaved as for the plasticware. Since DEPC reacts with Tris, the solutions containing Tris were prepared in DEPC-treated water and then autoclaved.

Note: Since DEPC is a suspected carcinogen, the above procedures were carried out in a fume hood.

Other precautions: Since human hands are a major source of exogenous RNase, gloves were worn during the handling of solutions, pipettes, etc., to prevent contamination.

2.15.2 Extraction of RNA

5-10g of plant tissue was frozen in liquid N_2 and then ground to a fine powder using a morter and pestle which has been sterilised by autoclaving. This operation took ~15min and the liquid N_2 was continually topped up. Once a fine powder was obtained 25ml of phenol-cresol solution was added along with 50ml of homogenising buffer.

Homogenising Buffer:	400mM NaCl
	50mM Tris-HCl (pH 9.0)
	5mM EDTA
	1% SDS
	10mM DDT
	1mM ATT (Aurintricarboxylic acid)
Phenol-Cresol Solution:	100g phenol
	0.1g 8-hydroxyquinoline
	10ml m-cresol
	Homogenising buffer added to saturation.
Phenol-Chloroform Solution:	100g phenol
	0.1g 8-hydroxyquinoline
	50ml chloroform
	sterile doubled-distilled water added to saturation

The plant material and solutions within the mortar were allowed to thaw whilst continuing to grind until a paste formed. The paste was then transferred to 4 x 30ml Corex® tubes and centrifuged at 10,000 rpm, 4°C for 15min. The supernatant was

transfered to 4 x 50ml falcon tubes containing an equal volume of phenol-chloroform, extracted thoroughly and centrifuged at 3000 rpm, 4°C for 30min. This last step was repeated before adding 12M-LiCl to the supernatant to give a final concentration of 2M. The tubes were incubated at 4°C overnight and then centrifuged as previously described. The supernatant was discarded and the pellet resuspended in a minimal volume of 5mM-Tris/HCl (pH 7.5) and reprecipitated with LiCl as previously described. The tubes were centrifuged once more as previously described and the pellet remaining was washed with ethanol (70% v/v) and left on ice for ten minutes. The tubes were centrifuged once more and the pellet dried under vacuum before being resuspended in a small volume (100-500µl) of DEPC-treated water. The concentration of RNA was then determined as described in Section 2.2.6. The purity of the preparation was estimated by measuring the O.D. at 260 nm and 280 nm. A pure preparation was taken as having an A260/A280 ratio of 1.8-2.0. The integrity of the RNA was examined on a 2% agarose mini-gel. All RNA preparations were stored at -80°C.

2.16 Digestion of DNA with restriction enzymes

The methods used were as described in Sambrook *et al.* (1989). Restriction digests were carried out using the Gibco-BRL React buffers which were provided with each batch of enzyme. There are ten different React buffers with a range of salt concentrations, each one suitable for a range of enzymes. DNA to be digested was incubated for 1-6h at 37°C in 1x appropriate buffer, 0.25mg/ml RNase (if necessary) and 1-20U/µg DNA of restriction endonuclease. Analytical digests were carried out in a volume of 10 or 20µl whereas preparative digests were carried out in larger volumes. When DNA was digested with two restriction enzymes, the endonuclease requiring the lower salt buffer was used first. After the recommended duration of digestion the salt concentration was adjusted and the second enzyme added.

2.17 Agarose gel electrophoresis of DNA

DNA was separated at room temperature on horizontal submerged agarose gels as described by Sambrook *et al.* (1989). The buffer system used was the Tris borate (TBE).

Agarose gels of appropriate concentration (1-3%) were made by adding the correct amount of agarose to 1 x TBE (0.09 M-Tris-borate, 2mM-EDTA) and heating in a microwave oven until the solution just reached boiling point. The melted agarose solution was allowed to cool to 60°C and ethidium bromide added to a final concentration of 0.5µg/ml. The solution was poured into the electrophoresis apparatus and allowed to set for approximately 30min. Electrophoresis running buffer was 1 x TBE. Prior to loading, one-tenth volume of loading buffer (125mM-EDTA, 50% (v/v) glycerol, 0.1% (w/v) bromophenol blue) was added to the DNA samples and the solutions mixed. Samples were loaded with an automatic pipette. Gels were run at 40-100mA until the bromophenol blue had travelled approximately 2/3 down the gel.

2.18 PCR amplification of cDNA

2.18.1 Synthesis of single-stranded cDNA

The first strand cDNA was synthesised (Gubler and Hoffman 1983) using a Pharmacia First-Strand cDNA Synthesis Kit (Pharmacia). The standard 33µl first strand synthesis reaction contained ~10µg of total RNA, 11µl Bulk First-Strand cDNA reaction Mixes (cloned, FPLCpure® Murine Reverse Transcriptase, RNAguard, RNase/DNase-free BSA, dATP, dCTP, dGTP, and dTTP, in aqueous buffer), 1µl DTT, 1µl NotI-d(T)18 bifunctional primer (1:25 diluted) or 1µl pd(N)6 primer or ~0.4µg specific primer.

The RNA in RNase-free water was incubated at 65°C for 10min followed by rapid

cooling on ice before addition to the reaction mixture. Synthesis was carried out at at 37°C for 1h after which the sample was heat denatured at 90°C for 5min, cooled on ice and directly used for PCR or stored at -20°C.

2.18.2 Synthesis of degenerate PCR primers

Oligonucleotides were synthesised by a solid-phase phosphotriester method on an Applied Biosystems Model 280A DNA synthesiser by Dr. V.B. Math. They were used for the purpose of PCR amplification of template-DNA and are discussed in Chapter VI.

2.18.3 PCR conditions

The PCR was carried out in a total volume of 100µl in a 0.4ml sterile Eppendorf® tube. The reaction mixture consisted of 10mM-Tris/HCl, pH 8.3, 3.5mM-$MgCl_2$, 0.01% Triton X-100, each dNTP at 0.2mM, each primer at 2.0 µM and template DNA. 2.5 units of *Taq* polymerase were added and the mixture overlaid with 100µl paraffin oil.

PCR was performed essentially as described by Saiki *et al.* (1988) in a DNA thermal cycler (Perkin-Elmer Cetus Corporation). PCR reaction conditions were

1	95°C	3min
2a	94°C	1min
2b	45 - 60°C	1 - 2min
2c	72°C	1 - 5min
3	70°C	7min
4	4°C	hold

2a, 2b and 2c were repeated 25 to 40 times.

2.18.4 Recovery and analysis of PCR products

The PCR product was cleaned (to remove paraffin oil that was used as an overlay) by extracting with 200µl $CHCl_3$ saturated with water; 10% of the aquous layer was loaded onto a gel for analysis of the products.

Cleaned PCR product was digested with appropriate restriction enzymes followed ethanol precipitation of DNA. The precipitate was dissolved in a suitable volume of TE and run on l.m.p. agarose gels.

The DNA band required was located by illumination of the gel (after ethidium bromide staining) with short-wave UV light and excised with a scalpel and placed in a 1.5ml microfuge tube and 300µl TE added. The gel slice was melted at 65°C for 10min with agitation, transferred to 37°C for 5min and then extracted with an equal volume of prewarmed (37°C) phenol saturated with TE. The DNA was further purified by another two rounds of phenol extraction followed by 3 extractions with $CHCl_3$ saturated with water. DNA was recovered by ehtanol precipitation as detailed in Section 2.19. DNA purified in this way was sufficiently pure for cloning.

2.19 Extraction, purification and ethanol precipitation of DNA samples

After the various enzymatic treatments of the DNA, protein and other contaminants, which could inhibit further reaction, were removed from the DNA solutions by phenol/chloroform extraction. An equal volume of TE-saturated phenol/chloroform (1:1 v/v) was added to samples which were then mixed by vortexing and centrifuged in a microfuge for 5min. The upper aqueous phase was removed to a fresh microfuge tube and

the process repeated. Finally, traces of phenol were removed similarly by extraction with an equal volume of chloroform. The final aqueous DNA pool was made 0.25M-sodium acetate by adding 4M-sodium acetate, pH 7.5 and 2.15 volumes of ice-cold ethanol was added. The mixture was cooled either at 0°C overnight, or at -20°C for 2h, or at -70°C for 15min and the precipitated DNA recovered by centrifugation (10,000 rpm, 4°C, 10 min) in a microfuge. DNA was air-dried before being resuspended in either TE or sterile deionised water.

2.20 Subcloning of PCR products

PCR products were subcloned into M13 mp18 and mp19 for sequence analysis.

2.20.1 Cloning into M13

M13 is a male specific filamentous bacteriophage of *E. coli*. The virus particles contain single-stranded, circular DNA which following infection, serves as a template for the synthesis of the complementary strand. The double-stranded form of the viral DNA present in infected cells is called the replicative form (RF-form). Several modifications have been introduced to phage and host cells to produce an efficient vector/host system (Messing, 1983), making it an ideal source of single-stranded DNA to be sequenced by the dideoxy sequencing method (Sanger *et al.*, 1977). The double-stranded RF form was modified to serve as an efficient cloning vector, by introducing a multiple cloning region (polylinker) and colour selection for the identification of recombinants using β-galactosidase activity as the basis of selection (Messing, 1983; Norrander *et al.*, 1983; Yanisch-Perron *et al.*, 1985). The vector carries a segment of *E. coli* DNA that contains the regulatory sequences and the coding information for the first 146 amino acids of the β-galactosidase gene (*lacZ*). Vectors of this type are used in host cells (eg. *E. coli* TG1)

that code for the carboxy-terminal portion of β-galactosidase, therefore though neither the host-encoded nor the plasmid-encoded fragments are themselves active, they can associate to form active enzyme. A polycloning site has been inserted, in-frame, within the coding region of the plasmid which does not affect the complementation, however insertion of additional DNA into the polycloning site generally destroys the complementation. Active β-galactosidase cleaves the chromogenic substrate X-Gal to produce a blue chromophore, when transformed cells are grown in the presence of the non-metabolizable *lac* operon inducer IPTG. However, in recombinant plasmids the ability for complementation is lost, the enzyme is inactive, and consequently the colonies appear white. False positive white colonies appear at low frequency, probably due to incorrect self-ligation of the vector (Yanisch-Perron *et al.*, 1985).

2.20.2 Ligations

The insert DNA was digested with the appropriate restriction enzymes, purified, extracted with organic solvents and ethanol precipitated. Bacteriophage vector M13 DNA was supplied linearised by digestion with the appropriate restriction enzymes.

Depending on the size of insert, ligation mixtures contained 20ng vector and between 50 and 100 ng insert in a final volume of 20μl. Ligation mixtures were pre-incubated at 60°C for ≤1min to melt any cohesive termini that had reannealed before addition of 1U T4 DNA ligase. The ligations [in 66mM Tris/HCl pH 7.6, 6.6mM $MgCl_2$, 0.5mM-ATP, 10mM DTT using 1U (Weiss *et al.*, 1968) of bacteriophage T4 DNA ligase] were performed in two stages:

a) Ligation at 14°C for 2h and 20min and

b) Ligation at 22°C for 1h and 20min.

2.21. Transformation of *E. coli* with bacteriophage M13 DNA

2.21.1 Preparation of competent cells

A single colony of *E. coli* TG1 from a stock minimal media plate was used to inoculate 10 ml of 2 x YT which was incubated overnight at 37°C. 2ml of overnight culture was used to inoculate 40 ml of 2 x YT which was grown at 37°C for around 2-3 hours with constant shaking to an OD at 600 nm of 0.4 - 0.6. The cells were cooled on ice for 10min and gently harvested (Corex®, 7000 rpm, 2min). Competent cells were prepared by the $CaCl_2$ method (Cohen *et al.*, 1972). First the cells were resuspended in one half of the original culture volume (20ml) of ice-cold, sterile 50 mM-$CaCl_2$ and placed in an ice bath for 20min. Following centrifugation (Corex®, 7000 rpm, 2 min) the cells were resuspended in 1/10th original culture volume of 50mM-$CaCl_2$. These cells were kept on ice and used the same day.

2.21.2 Transformation of competent cells

Transformations were carried out in sterile 1.5 ml microfuge tubes. An aliquot of ligation mix containing up to 20-30ng phage DNA was added to 300µl aliquots of competent cells and the mixture was incubated on ice for at least 40min. The DNA/cell mix was then heat shocked at 42°C for 3min and then put back onto ice. Mixtures containing 40µl 100mM-IPTG + 40µl 2% X-Gal + 200µl lawn cells were added to the tube followed by addition of 3ml molten H-top agar (kept at 42°C). These were mixed by rolling and poured immediately onto a pre-warmed (37°C) H-plate and left at room temperature (~30min) to set. Plates were then transferred to 37°C for overnight growth.

2.22 Preparation of single-stranded template DNA for sequencing

A 10ml 2 x YT overnight culture of *E. coli* TG1 was used to inoculate fresh 2 x YT medium at a ratio of 1ml:100ml. A single plaque was lifted using a sterile pasteur pipette and inoculated into 1.5ml of the low density *E. coli* TG1 culture in a Bijou bottle (7.5ml). The culture was shaken (37°C) for 6h and then centrifuged for 10min in a microfuge. The supernatant was carefully transferred to a fresh Eppendorf tube and recentriguged (5min) to remove any residual cells. The supernatant was added to 200µl PEG/NaCl, kept at room temperature [(20% w/v) polyethylene glycol 6000 in 2.5M NaCl)], mixed thoroughly and left for 15min at room temperature. Precipitated intact phage particles were harvested by centrifugation (5min, microfuge), and residual PEG/NaCl removed by a second centrifugation step (2min, microfuge). The viral pellet was resuspended in 100µl TE buffer, 50µl phenol saturated with TE added, the mixture vortexed for ≥20sec and left at room temperature for 15min. The mixture was revortexed for 20sec and then centrifuged for 3min. The aqueous layer was transferred to a fresh microfuge tube and extracted with 1ml chloroform. 7.5µl of 4M-sodium acetate pH 7.5 and 250µl ice-cold ethanol were added to the chloroform-extracted solution and the DNA was precipitated at -20°C overnight. The DNA was recovered by centrifugation (10min) and the template DNA resuspended in 30µl TE buffer. The integrity and purity of the template preparation was examined by running an aliquot on a 1% agarose gel. The template DNA was stored at -20°C.

2.23 Dideoxy DNA sequencing

PCR products were subcloned into M13 and sequenced using the dideoxy chain-termination method (Sanger *et al.*, 1980).

The Sequenase® Version 2.0 enzyme and kit obtained from United States Biochemicals (Cleveland, Ohio, USA) was used. Single-stranded template DNA was prepared as in Section 2.22.

2.23.1 Annealing of template and primer

Annealing reaction was carried in a total volume of 10µl in a centrifuge tube (1.5ml). The reaction mixtures contained 1µl (3ng) sequencing primer, 2µl reaction buffer, and ~1µg single-stranded template DNA. The volume was made up with deionised water and mixed thoroughly by pipetting 'up' and 'down'. The tube was warmed to 65°C for 2min, then the temperature of the tube was allowed to drop to $\leq$ 30°C over a period of 40min after which the tube was transferred onto ice. The annealed template was used within 4h.

2.23.2 Labelling and termination reactions

To the annealed template-primer was added: 1.0µl 0.1M DTT, 2.0µl diluted labelling mix (usually dITP mix was used to eliminate the compressions in GC rich region; 3µM dITP, 1.5µM dCTP, 1.5µM dTTP), 0.5µl (5µCi) α-^{35}S-dATP (1000 Ci/mmol), 2.0µl diluted Sequenase® 2.0 (1 : 8 dilution in enzyme dilution buffer and was used within 60min). The contents were mixed and incubated for 3min at 18°C. Four microfuge tubes were labelled G, A, T and C. Into each tube was placed 2.5µl of the appropriate termination mix and the tubes prewarmed for 1min at 37°C. On completion of the labelling reaction, 3.5µl of the reaction mix was placed in each of the four tubes. Termination reactions proceeded for 5 min at 37°C after which time the reaction was stopped by adding 4µl stop solution, mixed thoroughly and the tubes were stored at -20°C prior to electrophoresis. Samples were denatured by heating to 85-95°C for 2min before loading onto the gel.

2.23.3 Sequencing gel electrophoresis

Gels of dimension 20cm x 40cm x 0.4mm were prepared using 50ml of de-gassed linear sequencing mix (6% acrylamide:bisacrylamide (38:2), 7M urea, 1 x TBE), to which was added 300µl 10% ammonium persulphate and 50µl TEMED. Aluminium plates were placed either side of the gel to ensure even heat distribution during electrophoresis. The gels were pre-run for 30min at a constant current of 30-35mA. The wells were rinsed out to remove urea that had diffused into the well, before loading the samples. Approximately 3µl sample was loaded per well with Sharkstooth combs. Gels were run (30-35mA) for periods of $1^1/2$, 3 and $4^1/2$h for reading the maximum number of nucleotides. After electrophoresis the gel plates were removed from the sequencing apparatus and separated. The gel was fixed (also to remove urea) in 5% acetic acid (v/v), 15% methanol (v/v) for at least 20min. The gel was transferred onto Whatman 3MM paper and dried on a slab gel-drier at 80°C for 30min. Gels were autoradiographed for 20-48h without intensifying screens.

2.24 Computer programs for the analysis of DNA and protein sequences

The University of Wisconsin Genetics Computer Group (UWCGC) package of programs (Devereux *et al.*, 1984) mounted on the University of Glasgow VAX Cluster (VMS3), was used for the manipulation and comparison of DNA sequences. This package contains programs for the comparison of DNA sequences with those in GenBank and EMBL (European Molecular Biology Laboratory, Heidelberg, Germany) databases which were also mounted on VMS3.

2.24.1 UWCGC programs

WordSearch :-searches for similarity between a new sequence and any group of sequences.

Segments :-aligns and displays the segments of similarity found by WordSearch.

BestFit :-Finds the best region of similarity between two sequences, and inserts gaps if necessary to obtain optimal alignment.

The final alignment of the *P. sativum* DHQase-SHDase sequence with the other DHQases and SDHases sequences was produced by optimally matching the *P. sativum* DHQase-SDHase sequence to already generated multiple alignments of the other DHQase and SDHase sequences.

The multiple alignments of the DHQase and SDHase sequences were generated using a novel interactive multiple alignment program, *Mix'n'Match* (Bell *et al.*, 1993).

CHAPTER III

The essential histidine residue at the active site of *Escherichia coli* dehydroquinase

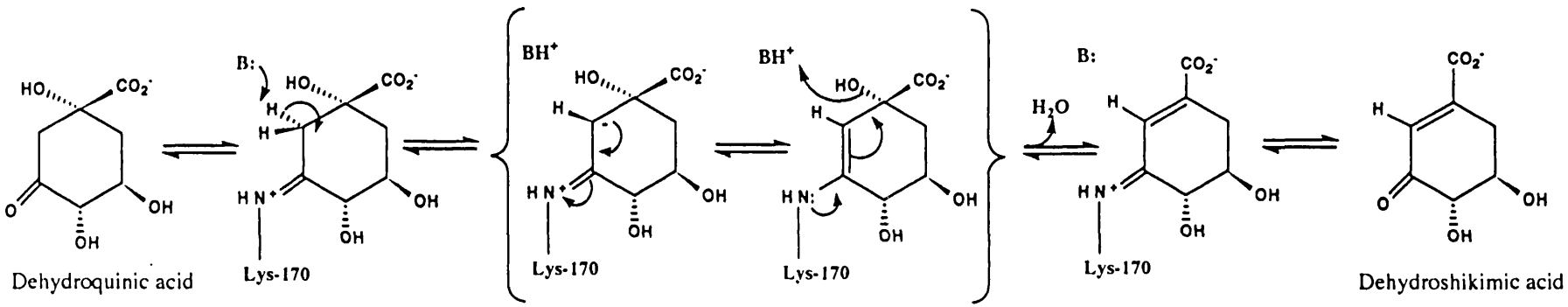


Fig. 3.1 Mechanism proposed for *E. coli* type I 3-dehydroquinase

(***B*** represents a general base thought to participate in the mechanism)

3.1 Introduction

Chemical modification is one of the most commonly used methods for identifying the residues of functional importance in proteins. If modification of a particular amino acid residue with group specific reagents results in loss of catalytic activity, then that residue may have a role in catalysis or binding (Means and Feeney, 1973). However, there are always problems in the interpretation of chemical modification results. For example, even if loss of enzyme activity is associated with modification of a specific residue, this does not confirm that the residue is located at or near the active site. The loss of activity could be due to steric hindrance of the active site or to a conformational change brought about by the covalent attachment of the modifying group. Thus, although chemical modification is an effective tool, it is necessary to confirm the function assigned to the modified residue by other means. A good approach is site-directed mutagenesis. The two techniques are complementary, since, to avoid testing large numbers of mutants one needs information about which residues of the protein are likely to be involved in its function. Clearly it is an enormous advantage if the three dimensional structure of the protein is available.

This chapter describes chemical modification experiments that were carried out on *E. coli* dehydroquinase (DHQase).

3.2 Mechanism of type I DHQases - is a general base involved?

The mechanism proposed for *E.coli* DHQase [(Walsh, 1979), see Fig.3.1] involves formation of a Schiff's base between the substrate and a lysine residue of the enzyme, followed by abstraction of the pro-R hydrogen at the C-2 position by a general base. The imine is then in a position to act as an electron sink which can stabilize the resulting

carbanion. The hydroxyl at C-1 is then eliminated (possibly by general acid catalysis) to generate the product (as a Schiff base intermediate) which then dissociates from the enzyme.

Central to this hypothesis is the involvement of a general base/acid. Evidence to support the idea of a general base comes, principally, from the pH/logV_{max} profile for the enzyme which shows that a group with a pK_a of 6.1 is involved in the mechanism (Chaudhuri *et al.*, 1986). Preliminary data on diethylpyrocarbonate (DEPC) inactivation of *E. coli* DHQase has suggested that there may be a reactive histidine residue at the active site (Chaudhuri *et al.*, 1986). The present investigation confirms and extends this work.

3.3 Chemical modification with DEPC

3.3.1 Background

DEPC has proven to be the most frequently used reagent for the selective modification of histidine residues (Lundblad and Noyes, 1984). Prior to the development of DEPC, photo-oxidation methods were used for histidine modification, but problems persisted with specificity; methionine, tryptophan, tyrosine, serine and threonine are all sensitive to photo-oxidation. DEPC shows good specifity at or near neutral pH. The reaction yields N-carbethoxyhistidine (Fig. 3.2) due to reaction with one of the imidazole nitrogens and can be reversed by addition of hydroxylamine, resulting in the recovery of histidine. The formation of carbethoxyhistidine leads to an increase in absorbance at 237-242nm and from the molar absorption difference ($\Delta\varepsilon_{240}$ = 3200$M^{-1}cm^{-1}$; Miles, 1977), the number of residues modified can be calculated.

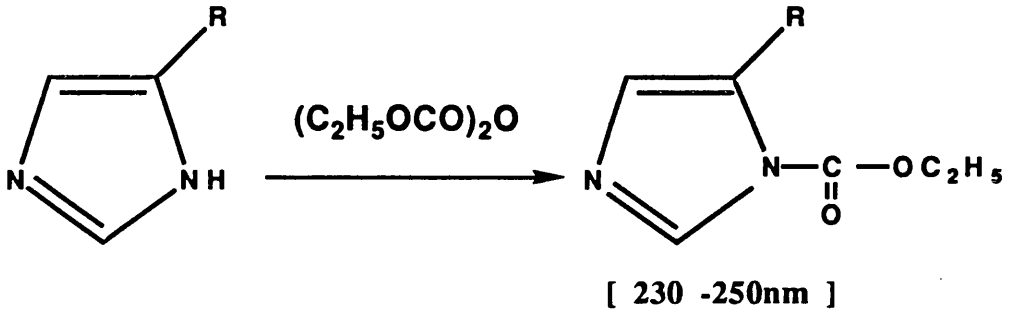


Fig. 3.2 Reaction of DEPC with the imidazole side chain of histidine

3.3.2 Kinetics of DEPC inactivation of DHQase

Incubation of *E. coli* DHQase with DEPC at 25°C in 50mM-potassium phosphate buffer, pH 6.0, results in the rapid loss of enzymatic activity. In control experiments no inactivation was detected. The fraction of activity remaining was calculated as the percentage of activity remaining (A/A_0, where A is the specific activity at each time and A_0 is the initial specific activity *i.e.* at 0 time) at time (t). The plot of the logarithm of A/A_0 versus t at various concentrations of DEPC shows pseudo first-order kinetics (Fig. 3.3.A). It is evident that the rate of inactivation depends on the DEPC concentration. From the pseudo first-order plot, the half-life period ($t_{1/2}$) of inactivation was determined at various concentrations of DEPC (Table 3.1).

Table 3.1

Half-life **($t_{1/2}$) and first-order rate constant (k_{obs}) for the inactivation of DHQase at different DEPC concentrations**

[DEPC] mM	$t_{1/2}$ (min)	k_{obs}/min
0.35	29.30	0.024
0.50	10.20	0.068
0.65	7.25	0.095
1.00	4.75	0.146
1.30	3.125	0.222
2.70	1.75	0.396

The *half-life period* ($t_{1/2}$) for the inactivation process is the time after which the activity has decreased to half of its original activity, *i.e.* $A = A_0/2$. So the half-life period $t_{1/2}$ may be expressed:

$$t_{1/2} = \ln 2/k_{obs} = 2.303 \log 2/k_{obs} = 0.693/k_{obs}$$

From this relationship the pseudo first-order rate constant for inactivation, $k_{obs} = 0.693/t_{1/2}$, was calculated (Table 3.1).

The relationship between the observed first-order rate constants (k_{obs}) and the concentrations of DEPC may be described by the equation:

$$k_{obs} = k[\text{DEPC}]^n \qquad (1)$$

where n is the reaction order with respect to DEPC concentration, k_{obs} is the pseudo first-order rate constant, and k is the second-order rate constant with respect to DEPC concentration (Levy *et al.*, 1963).

A plot of the first-order rate constants (k_{obs}) as a function of DEPC concentration (Fig. 3.3.B) shows that this inactivation is a bimolecular process which is not dependent on the formation of a reversible enzyme-DEPC complex prior to inactivation (Kitz and Wilson, 1962; Church *et al.*, 1985). The second-order rate constant (k) for the inactivation of dehyroquinase by DEPC calculated from these data is 148.5 $M^{-1}min^{-1}$.

A kinetic analysis was used to determine the number of essential histidine residues (Takeuchi *et al.*, 1986; Levy *et al.*, 1963) in DHQase. The inactivation process of

dehydroquinase (DHQ) and DEPC occurs according to the following expression to produce DHQ-DEPC complex.

$$DHQ + n\text{DEPC} \longrightarrow \text{DHQ-DEPC}n \qquad (2)$$

where DHQ-DEPC*n* represents inactivated enzyme.The rate of inactivation can be described by

$$-d[\text{DEPC}]/dt = k[\text{DHQ}][\text{DEPC}]^n \qquad (3)$$

where *n* is the kinetic order of the reaction or minimal number of moles of DEPC required to react with essential histidine residues of DHQase to cause inactivation (Levy *et al.*, 1963). If [DEPC] >> [DHQ], then equation (3) can be rearranged and integrated to yield

$$\log k_{obs} = n \log[\text{DEPC}] + \log k \qquad (4)$$

The reciprocal of the half-life ($1/t_{1/2}$) may be substituted for k_{obs};

$$\log(1000/t_{1/2}) = n \log[\text{DEPC}] + \log k \qquad (5)$$

in this way, a constant is introduced that has no effect on the slope. The plot of $\log(1000/t_{1/2})$ versus log[DEPC] (Levy *et al.*, 1963) yielded a straight line with a slope near unity (1.04) (Fig. 3.3.C). This indicates that the covalent modification of a single residue of histidine was required to inactivate DHQase.

Fig. 3.3 Kinetics of inactivation of DHQase with DEPC

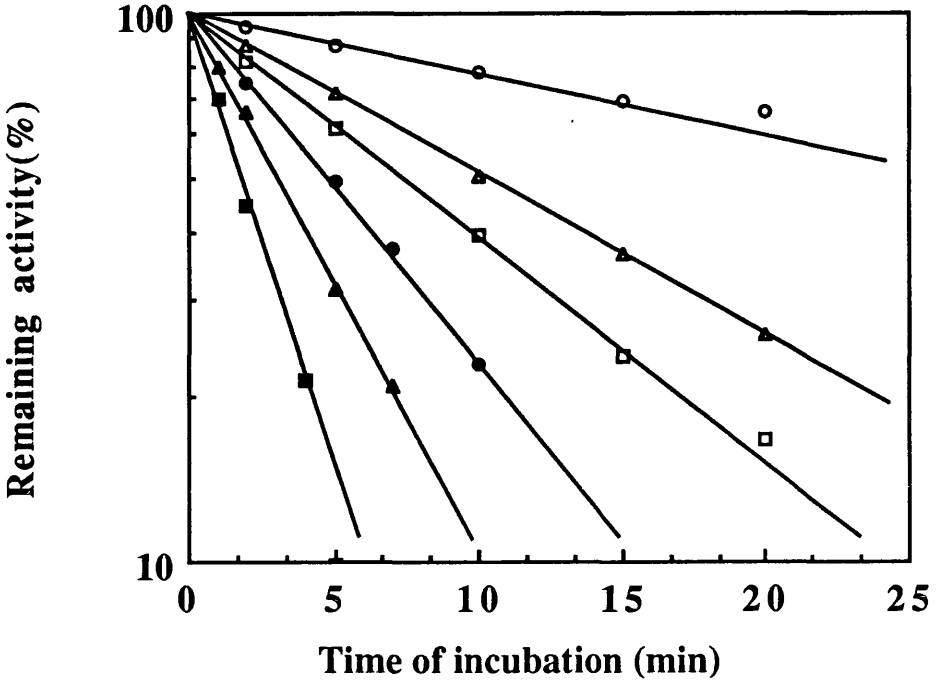
A. Pseudo first-order plots for inactivation

DHQase (~2μM) was incubated with increasing concentrations of DEPC in 50mM-potassium phosphate buffer, pH 6.0, at 25°C. At time intervals, aliquots were removed for measurements of the remaining activity as described in "Materials and Methods" (Section 2.7.2). The concentrations of DEPC used were, 0.35mM (O), 0.50mM (Δ), 0.65mM (□), 1.00mM (●), 1.3mM (▲) and 2.7mM (■).

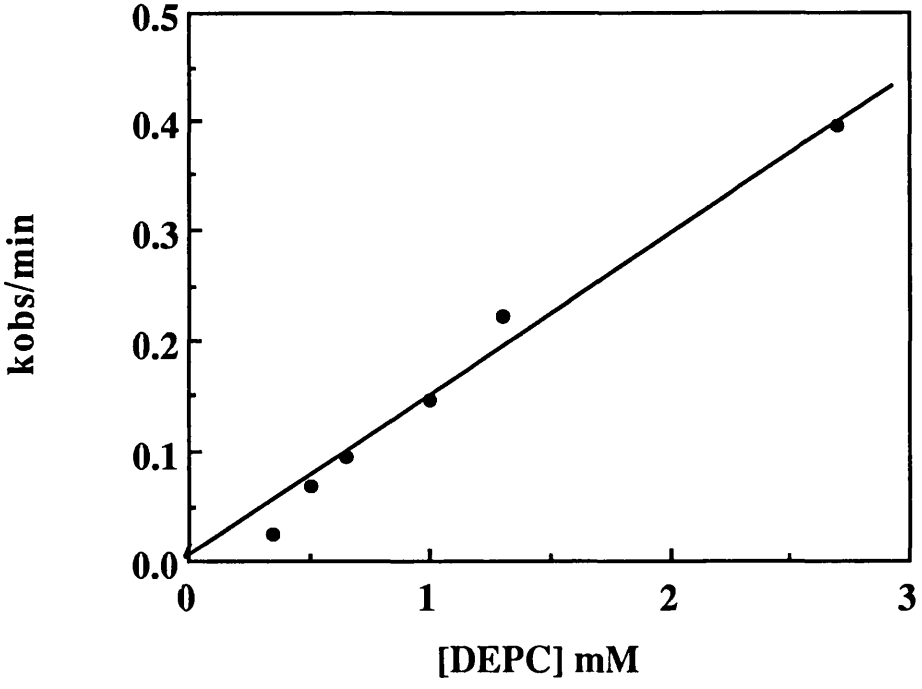
B. Determination of the second-order rate constant of inactivation

Pseudo first-order rate constants (k_{obs}) calculated in part *A* were proportional to DEPC concentration. The slope of this curve, calculated to be 148.50 $M^{-1}min^{-1}$, is the second-order rate constant (k) of inactivation.

A


100
10
Remaining activity(%)
0
5
10
15
20
25
Time of incubation (min)

B


0.5
0.4
0.3
0.2
0.1
0.0
kobs/min
0
1
2
3
[DEPC] mM

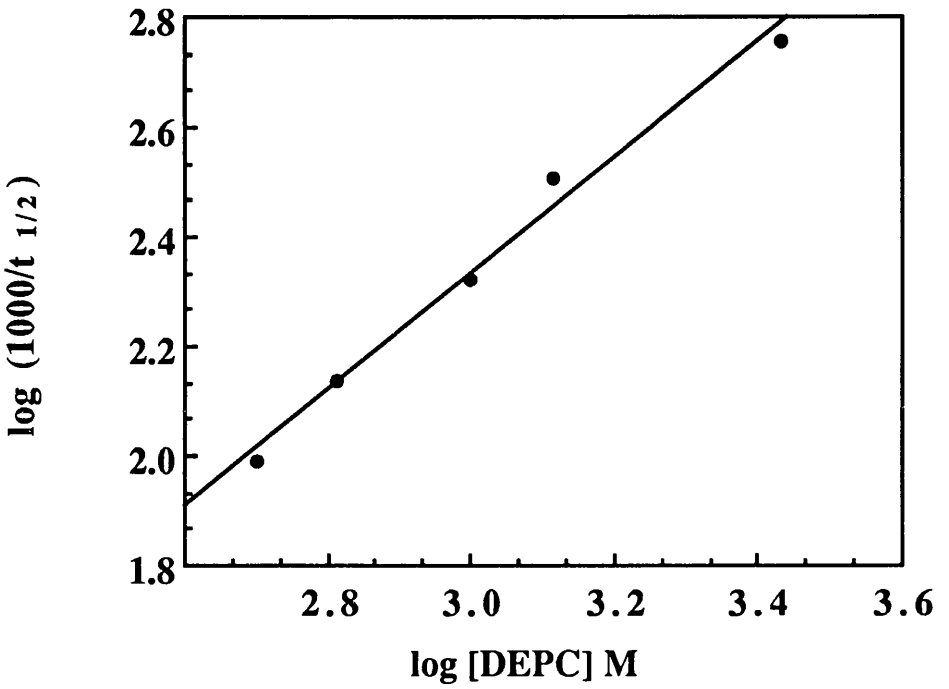


Fig.3.3 C The first-order rate of inactivation of DHQase with respect to DEPC concentration

Data were collected from Fig. 3.3 ***B*** and plotted as described in the text (Section 3.3.2). The slope of this line was 1.04, which indicates that modification of a single mole of histidine is required to cause inactivation of DHQase.

3.3.3 Substrate protection against DEPC inactivation

Substrate protection of an enzyme against inactivation would mean that the vulnerable amino acid residues are no longer accessible to the modifying reagent. Such protection strengthens the argument that the residues are located at or near the active site (Cohen, 1970).

DHQase activity could be protected against DEPC inactivation by the prior addition of ammonium dehydroquinate (Fig. 3.4.A). Protection (%) is expressed as

$$[k_{obs}(\text{unprotected}) - k_{obs}(\text{protected})/k_{obs}(\text{unprotected})] \times 100$$

Protection occurred at very high substrate concentrations, which exceeded the K_m for dehydroquinate (16μM at pH 7.0). The K_S value was determined as 325μM at pH 6.0 (Fig.3.4.B). The protective effect of dehydroquinate against 0.7mM DEPC inactivation is evident also from the data shown in Table 3.2.

Table 3.2

Effect of substrate concentration on DHQase inactivation by DEPC

Substrate	$t_{1/2}$ (min)	k_{obs}/min	% Protection
No substrate	6.0	0.115	No protection
100μM	9.4	0.074	36%
200μM	11.2	0.062	46%
600μM	19.25	0.036	69%
700μM	22.60	0.030	74%

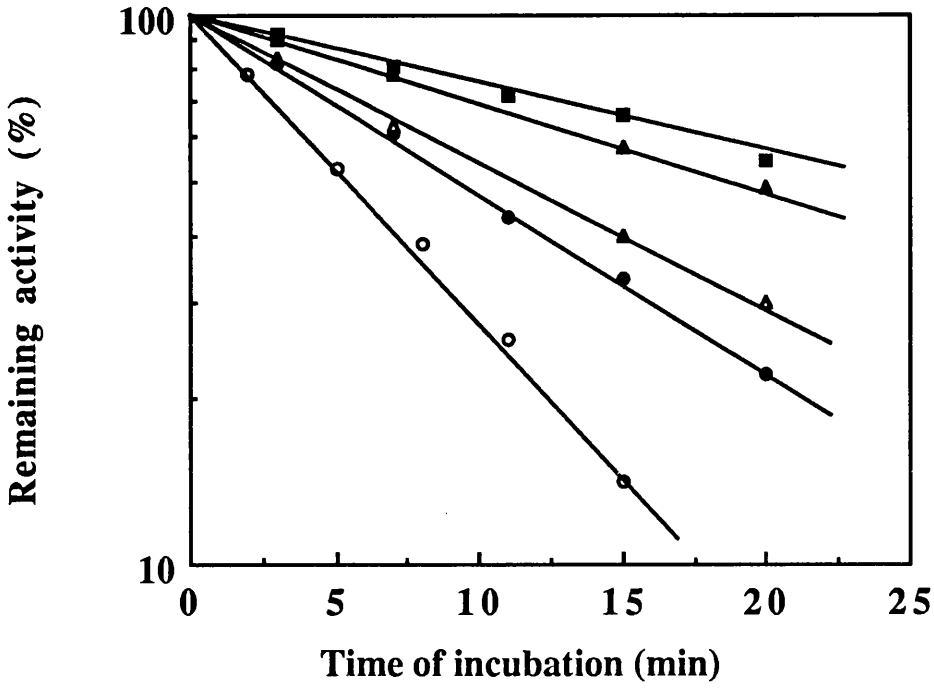
Fig. 3.4 Substrate protection against inactivation of DHQase by DEPC

A. Pseudo first-order rate plots are shown for the inactivation of DHQase (2μM) by DEPC (0.7mM) in the presence of equilibrium mixtures of substrate and product (K_{eq} = 5; Mitsuhashi and Davis, 1954). Starting concentrations of ammonium dehydroquinate were, 0μM (O), 100μM (●), 200μM (Δ), 600μM (▲) and 700μM (■).

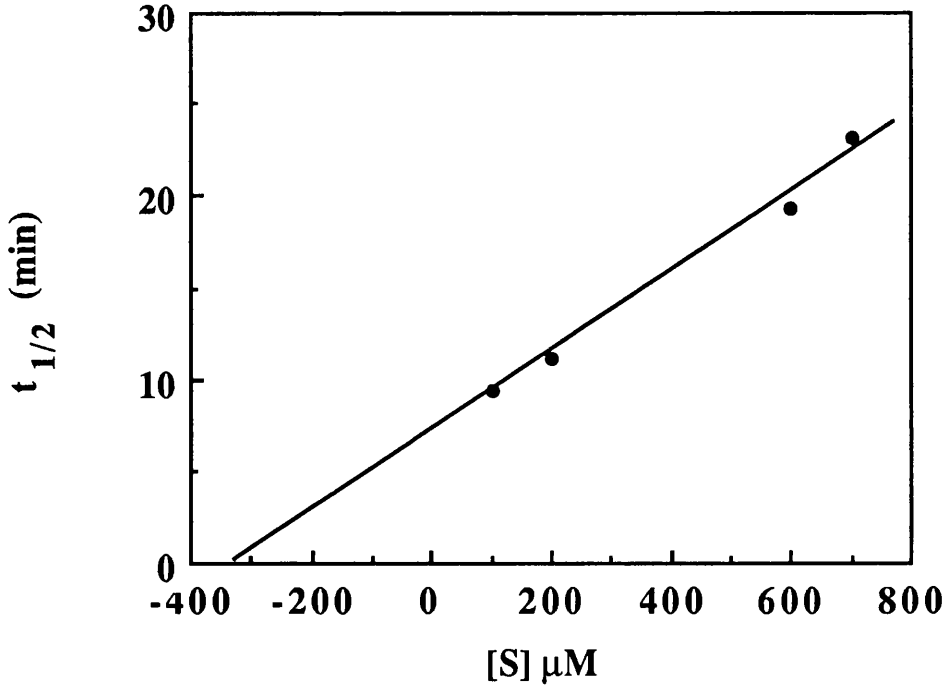
B. K_S for ammonium dehydroquinate

The *half-life* of the enzyme calculated in part *A* were plotted against the concentration of ammonium dehdroquinate. From the intercept K_S value calculated to be 325μM.

A


100
10
0
5
10
15
20
25
Remaining activity (%)
Time of incubation (min)

B


30
20
10
0
-400
-200
0
200
400
600
800
$t_{1/2}$ (min)
[S] μM

3.3.4 Characterization of DEPC-modified DHQase

In addition to histidine residues, it has been demonstrated that DEPC can react with cysteine, lysine, and tyrosine residues in proteins (Muhlrad *et al.*, 1967). Data are presented in the following paragraphs which rule out inactivation due to modification of residues other than histidine.

The difference spectrum of the DEPC-modified enzyme and untreated DHQase shows an absorption maximum at 240nm (Fig.3.5.A) which is characteristic of the carbethoxylation of histidine residues (Ovadi *et al.*, 1967; Miles, 1977). As there is no cysteine residue in DHQase, any contribution to the absorption spectrum at 230nm due to cysteine modification can be discounted (Melchior and Fahrney, 1970). The complete absence of spectral change above 270nm, ruled out the possible formation of *o*-carbethoxytyrosine, which is characterised by a dramatic decrease in the absorption spectrum at 278nm ($\Delta\varepsilon_{278}$ = 13,000 $M^{-1}cm^{-1}$; Burstein *et al.*, 1974).

3.3.5 Inactivation and carbethoxylation reaction

Figure 3.5.B shows that the enzyme (19.2μM) is rapidly inactivated by DEPC (0.50mM), with a half-time of about 10min. Concomitantly, a rise in A_{240} indicates the carbethoxylation of histidine residues. Loss of activity correlates with the modification of six histidine residues per subunit (Fig.3.5.C), calculated from the extinction coefficient for carbethoxyhistidine ($\Delta\varepsilon_{240}$ = 3200 $M^{-1}cm^{-1}$; Miles, 1977). These results suggested that both essential and nonessential histidine residues were being modified by DEPC.

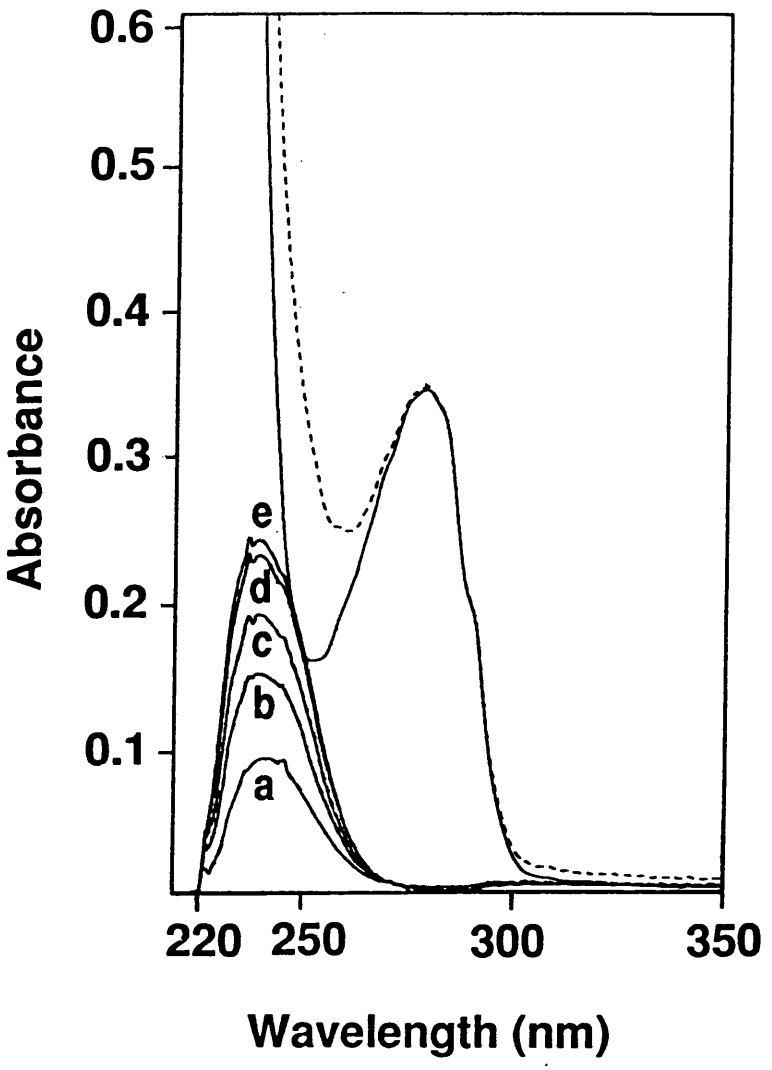


Fig. 3.5 A Difference spectra for modification of DHQase by DEPC illustrating the signal, representative of N-carbethoxyhistidine, appearing at 240nm

The curves were obtained at various time points during the inactivation; **a** (2min), **b** (5min), **c** (10min), **d** (25min) and **e** (40min). A_{280} absorbance of unmodified DHQase (——) and after 40min of modification (·····) are also shown.

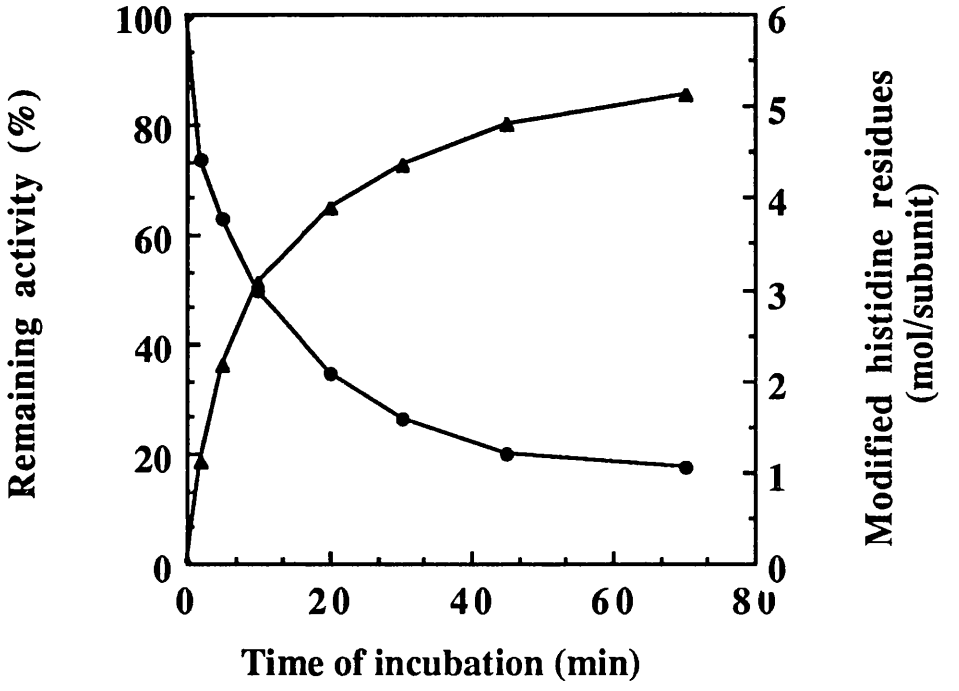
Fig.3.5 B The correlation between the extent of inactivation and the extent of modification

DHQase (18.44μM) was incubated with 0.5mM-DEPC in 50mM-potassium phosphate buffer, pH 6.0, at 25°C. At times the fraction of remaining activity (●) was measured and difference spectrum was obtained. The number of modified histidine residues(▲) was determined from the increase in absorbance at 240nm by using molar absorption coefficient $\varepsilon_{240} = 3200M^{-1}cm^{-1}$ (Miles, 1977).

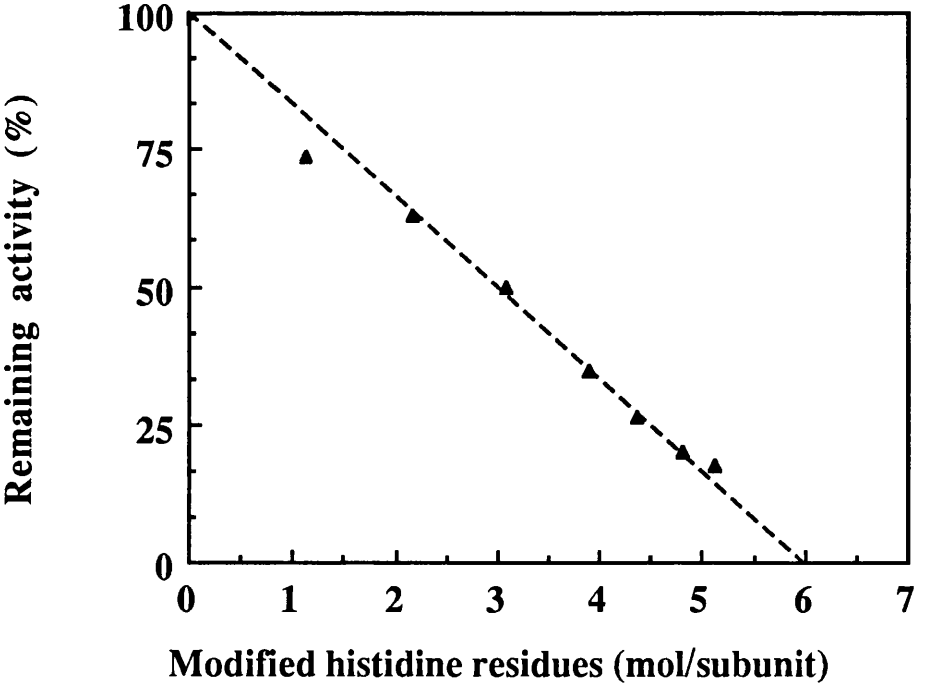
C. The relationship between the remaining activity and the number of histidine residues modified by DEPC

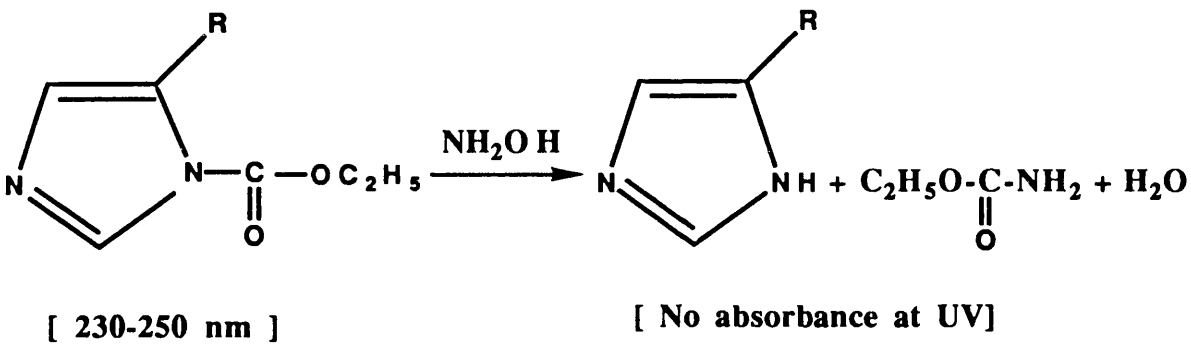
Extrapolation of remaining activity correlates the modification of six histidine residue per subunit. Data were taken from ***B***.

B


Remaining activity (%)
100
80
60
40
20
0
0
20
40
60
80
Time of incubation (min)
6
5
4
3
2
1
0
Modified histidine residues (mol/subunit)

C


Remaining activity (%)
100
75
50
25
0
0
1
2
3
4
5
6
7
Modified histidine residues (mol/subunit)



[230-250 nm]

[No absorbance at UV]

Fig. 3.6A Reaction of hydroxylamine with carbethoxyhistidine

3.3.6 Reactivation of DEPC-modified enzyme with hydroxylamine

The reversal of the DEPC modification with hydroxylamine, and recovery of an altered biological function provides insight into the chemistry of the reaction of DEPC with histidine residues (Fig.3.6.A). This is additional useful evidence to establish the possible function of histidine residues.

Treatment of DEPC-modified DHQase with hydroxylamine regenerated DHQase activity (Table 3.3), a result which is indicative of the reversible carbethoxylation of histidine residues (Miles, 1977).

Table 3.3

Reversal of DEPC inactivation upon hydroxylamine treatment

Time	% Recovery
1h	30%
2h	51%
4h	60%
6h	78%
20h	75%

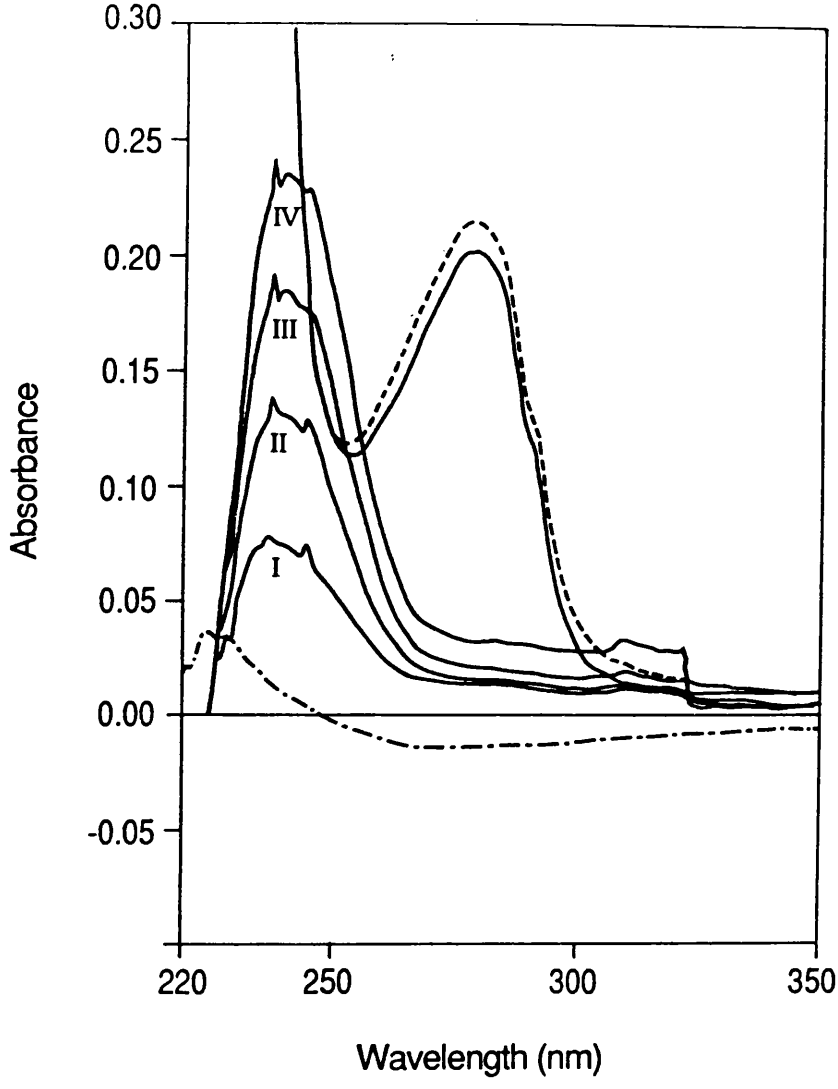
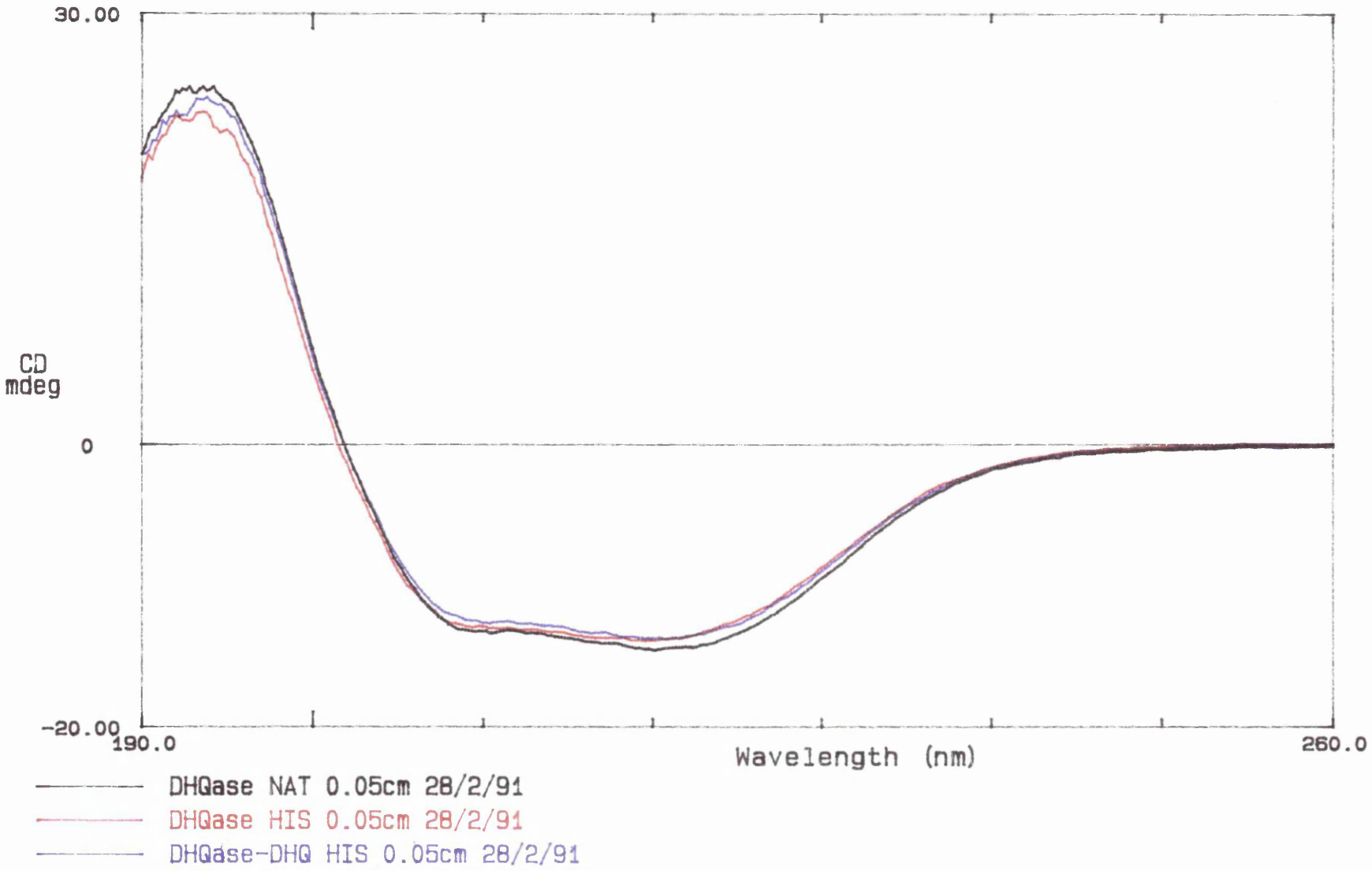


Fig. 3.6 B The effect of hydroxylamine on carbethoxylation at 240nm

DHQase (~18μM) was incubated with 0.5mM-DEPC in 50mM-potassium phosphate buffer, pH 6.0 at 25°C, and the changes in absorbance (220 - 350nm) were monitored with time against the native enzyme. The difference spectra I, II, III and IV were recoded at 2min, 5min, 10min and 20min respectively, indicated the carbethoxylation reaction. After 30min (at this stage enzyme is >90% inactive), hydroxylamine (500mM) was added to both DEPC-modified and native enzyme and left at 4°C for 24h. A difference spectrum (–·—·–) determined after 24h hydroxylamine treatment showed a marked decrease in A_{240}, indicated the decarbethoxylation reaction (see Fig. 3.6 A). Spectra of DEPC-modified enzyme (– – –) and native enzyme (——) after hydroxylamine treatment are also shown.

Approximately 75% of the activity was restored by hydroxylamine with the removal of all the carbethoxylated histidine residues, as detected by the decrease in the difference spectrum at 240nm (Fig.3.6.B). According to Miles (1977), failure of hydroxylamine to fully reactivate a modified enzyme implies either reaction of DEPC with residues other than histidine or reaction of two equivalents of DEPC/equivalent of histidine followed by a Bamberger reaction to open the imidazole ring (Dann and Britton, 1974). For DHQase, the reaction of one histidine residue with two equivalents of DEPC is unlikely because there is no marked increase in absorbance in the 220-230nm region during inactivation, characteristic of the Bamberger reaction (Loosemore and Patt, 1976). Other residues in proteins such as lysines can react with DEPC. An essential lysine for DHQase has been identified (Chaudhuri *et al.*, 1991), so the partial reactivation of the DEPC-modified enzyme by hydroxylamine may due to the reaction of this amino group with DEPC. The carbethoxylysyl bond is stable and is not cleaved by hydroxylamine. However, the K_m value of the reactivated enzyme was very similar to the value for native, underivatized DHQase (19μM compared with 16μM), further supporting the notion that inactivation of the enzyme is solely due to the reversible modification of histidine residues by DEPC. The activity recovered after hydroxylamine treatment is consistent with the material being a mixture of 75% fully active, native enzyme and 25% completely inactive enzyme. Alternatively the incomplete recovery of activity may be due to irreversible denaturation of the enzyme by hydroxylamine; this is consistent with similar difficulties which other workers have experienced in restoring enzyme activities after DEPC modification (Badet and Badet, 1992).


30.00
CD
mdeg
0
-20.00
190.0
260.0
Wavelength (nm)
DHQase NAT 0.05cm 28/2/91
DHQase HIS 0.05cm 28/2/91
DHQase-DHQ HIS 0.05cm 28/2/91

3.3.7 Circular dichroism (CD) measurement

It is unlikely that any chemical modification can be accomplished without some change in the conformation of the protein. If conformational changes occur, they may be responsible for the observed changes in the biological property being considered. Therefore, it is important to know whether or not treatment has any effect on protein structure.

The effect of DEPC treatment on the structure of DHQase was studied using CD to detect any change of secondary structure upon modification. The CD spectra of the inactivated enzyme, which had been incubated with 1mM DEPC for 25min at pH 7.0, were almost the same as those of the native and substrate protected enzyme. This established that the DEPC treatment of DHQase does not effect the CD spectrum, confirming that there is no major conformational change upon modification (Fig. 3.7).

Fig. 3.7 Secondary structure of native and DEPC-modified DHQase

DEPC-modified protein in the presence and absence of substrate were prepared in 50mM-potassium phosphate buffer, pH 7.0 as described in "Materials and Methods" (Section 2.8.2). The same buffer, in the absence of enzyme was used as reference. Far-UV CD spectra on native enzyme (black line), DEPC-modified enzyme (red line) and DEPC-modified enzyme in the presence of substrate (blue line).

3.3.8 Number of essential histidine residues

Inactivation of DHQase by DEPC was linearly related to the extent of histidine residue modification and, on extrapolation of the plot to 0% residual activity, correlated with the modification of six histidine residues per subunit (Fig 3.5.C). DHQase has six histidines per subunit and, it appears that all of them can be modified by DEPC. The problem is in differentiating essential residues from non-essential residues. As reported by Tsou (1962) and Horiike and McCormick (1979), simple extrapolation of the plot to 0% activity rarely provides the number of essential residues. This problem can be addressed, however, by the graphical method of Tsou (1962) which uses the relationship,

$$a^{1/i} = (p - m) / p \qquad (6)$$

where a is the remaining activity when m groups have reacted, p is the total number of groups modified, and i is the number of critical residue(s) for catalytic activity. The number of essential residue(s) is that value of i, for a given value of p, which gives a straight line when $a^{1/i}$ is plotted against m. Non-essential residues must react either much faster or much slower than the essential residue for the method to be valid. A straight line shows a satisfactory fit to the observed experimental data (solid triangles) for the inactivation of DHQase with DEPC when $p = 6$ and $i = 1$ (Fig. 3.8). The other lines shown in Fig. 3.8 (open symbols) are theoretical curves fitted to the same data also for $p = 6$ but assuming that more than one critical residue ($i > 1$) is responsible for the inactivation. The analysis indicates that the nonessential hidtidine residues react faster than the essential one.

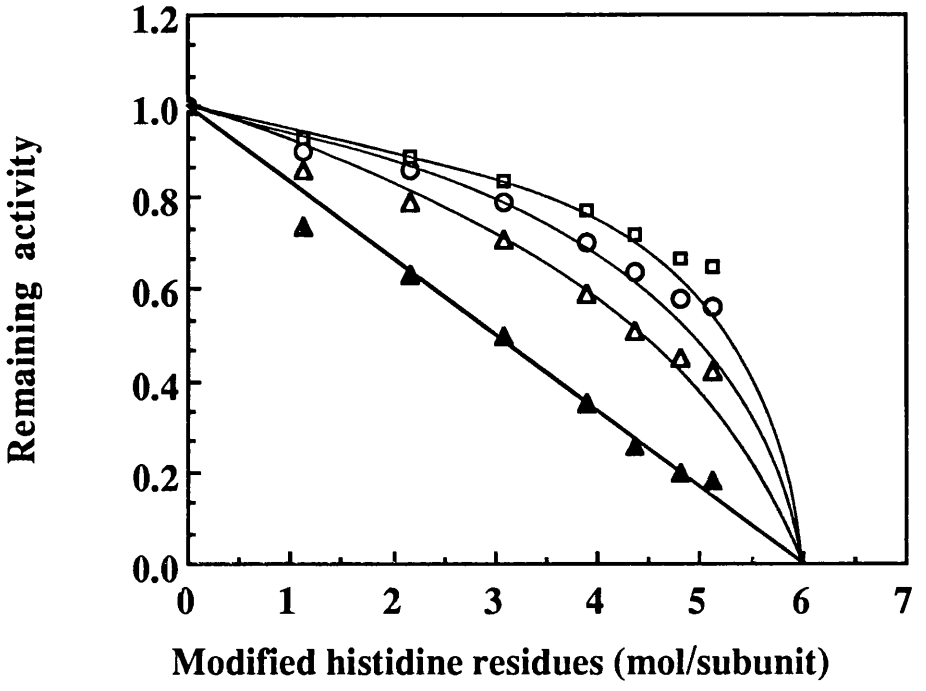


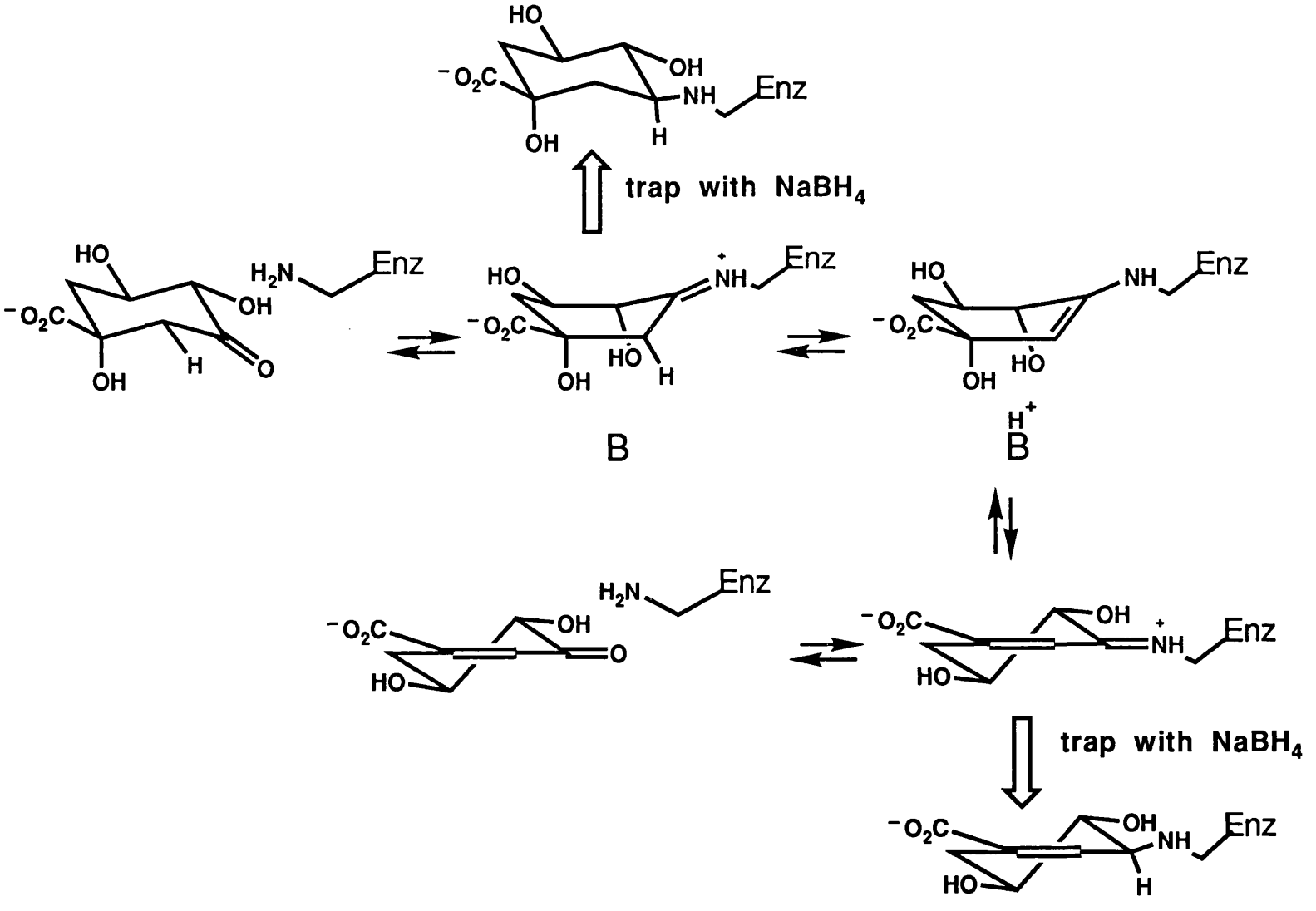
Fig. 3.8 Analysis of the DEPC-modified DHQase data by the method of Tsou

Experimental conditions were the same as those described in Fig. 3.5. The data are plotted assuming, in Equation (6) (Section 3.3.8), $p = 6$ and $i = 1$ (▲), 2 (Δ), 3 (O) and 4 (□). The *solid* symbols represent the actual data and the *open* symbols the same data fitted to the different values of i.

3.3.9 Number of residues protected by substrate

If the modification of only one histidine residue is responsible for the susceptibility towards DEPC inactivation then it might be expected that incubating the enzyme with an equilibrium mixture of substrate and product should protect that histidine from modification, and this should correlate with the protection of activity and a reduced absorbance change at 240nm. A technical problem arises in attempting this experiment, however, because the u.v. absorption maximum of the product (234nm) at the concentrations used in protection experiments completely masks the signal at 240nm. This problem can be circumvented by comparing the stoichiometry of histidine modification in the unliganded case with enzyme to which the product Schiff base intermediate has been covalently attached by reduction with sodium borohydride (Fig. 3.9.A) (Chaudhuri *et al.*, 1991; Shneier *et al.*, 1991; Kleanthous *et al.*, 1991) and the excess unbound ligand removed by dialysis. In this experiment the active site residues should be extensively protected while the remaining histidines should still be available for modification by DEPC. The results from such an experiment are shown in Fig. 3.9.B. These data show clearly that irreversible occupancy of the active site (following trapping of the product Schiff base) protects approximately one histidine residue from DEPC modification, consistent with the preceding data that there is only one essential histidine residue at the active site of DHQase.

Fig. 3.9A Trapping of Schiff base intermediates with sodium borohydride
Large arrows signify the trapping of imine intermediates with sodium borohydride



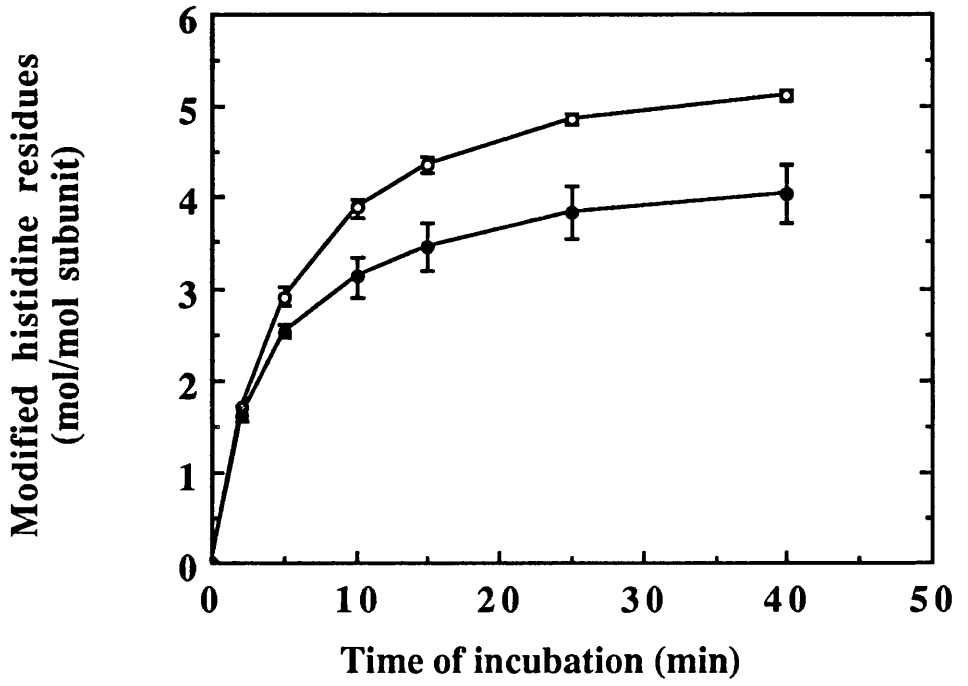


Fig.3.9B Stoichiometry of histidine modification by DEPC of native (O) and substrate-trapped DHQase (●)

See Fig. 3.9 *A* and "Materials and Methods" (Section 2.7.8) for details. Difference spectroscopy at 240nm was used to monitor the extent of histidine modification in duplicate samples of each protein. The *error bars* are the average values from duplicate determinations. The substrate-trapped enzyme is protein to which the Schiff's base product has been irreversibly attached to the enzyme active site by reduction with sodium borohydride (Shneier *et al.*, 1991).

3.3.10 pH dependence of DEPC inactivation

DEPC reacts only with the unprotonated form of the imidazole ring of histidine residues in proteins (Holbrook and Ingram, 1973) and thus the pH dependence of inactivation can provide information about the nature of the pK_a's of essential histidine residues in proteins. The apparent pK_a of the reacting active site histidine residue of DHQase was estimated by following the pH dependence of the first order rate constant for DEPC inactivation in phosphate buffer. The relation between the first order rate constant and pH is given by

$$k_{obs} = \frac{(k_{obs})_{max}}{(1+[H^+]/K_a)} \qquad (7)$$

which can be rearranged to its linear form,

$$k_{obs}.[H^+] = K_a.(k_{obs})_{max} - K_a.k_{obs} \qquad (8)$$

where K_a is the proton dissociation constant of the reacting group and $(k_{obs})_{max}$ is the pseudo first order rate constant of the unprotonated reacting group (Takeuchi *et al.*, 1986). A plot of $k_{obs}.[H^+]$ versus k_{obs} for the DEPC inactivation of DHQase resulted in a straight line (Fig. 3.10); $(k_{obs})_{max}$ was calculated to be $0.155 min^{-1}$ from the ordinate intercept and a pK_a of 6.18 was calculated from the slope. This value is in good agreement with the pK_a of the group essential for activity reported by Chaudhuri *et al.*, 1986. It is also a typical pK_a for a histidine residue in a protein, providing further evidence that the DEPC-inactivation is due to the modification of a histidine residue.

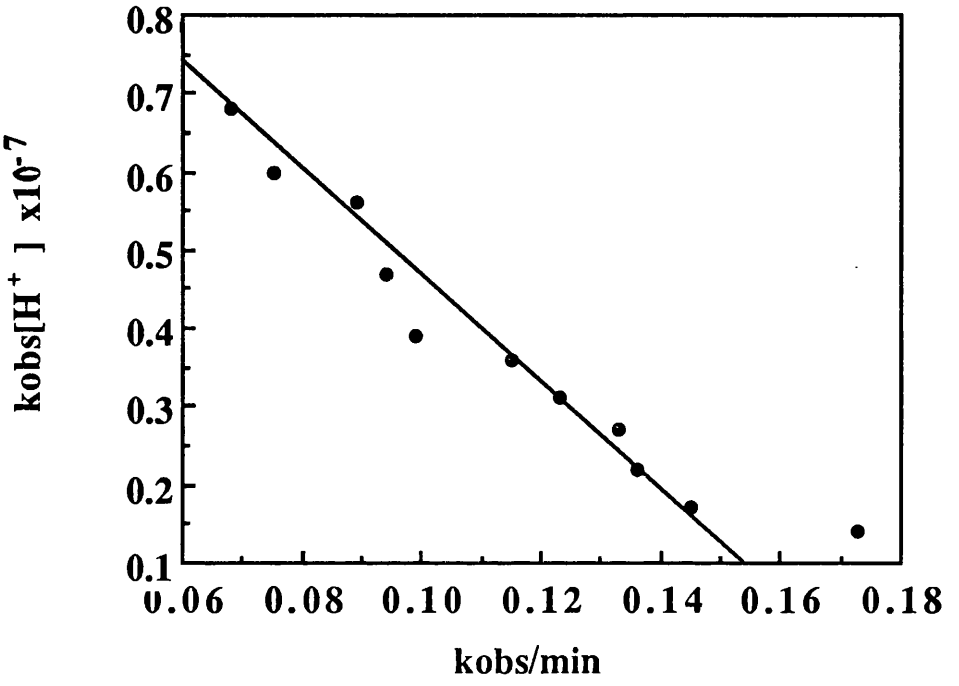


Fig. 3.10 Inactivation of DHQase by DEPC as a function of [H+]

Conditions for inhibition are the same as shown in Fig. 3.3, except the pH of each reaction varied as shown. The observed rate constants for inactivation were calculated from pseudo first-order plots at various pH values and plotted using Equation (8). The pK_a for inactivation was obtained from the slope of the line is 6.18.

3.4 Identification of the essential histidine by differential peptide mapping

As a first step in the structure-function studies, it is important to identify the active site residues in the enzyme sequence for subsequent gene manipulation. In the previous sections, chemical modification data indicate that a single active site histidine residue is involved in the mechanism of DHQase from *E. coli*. Comparison of the modification of DHQase with DEPC in the absence of substrate and with substrate-trapped enzyme indicated that the modification of one histidine correlated with the complete loss of activity (Fig. 3.8 and 3.9.B). Here I describe experiments to locate this residue in the known amino acid sequence (Duncan *et al.*, 1986; Chaudhuri *et al.*, 1991). We employed the differential peptide mapping technique successfully used by Vangrysperre *et al.* (1989) to study the DEPC mediated inactivation of D-xylose isomerase. The method relies on the fact that N-carbethoxyhistidine is sufficiently stable at pH 6-6.5 to permit non-radioactive DEPC to be used and labelled peptides are simply detected by their absorbance at 240nm. Differential peptide mapping following DEPC inactivation in the presence and absence of ligand was used to detect regions in the peptide map which are specifically protected by the ligand.

3.5 Peptide mapping

3.5.1 Proteolysis of DEPC-modified DHQase

First it was necessary to find suitable conditions for the complete proteolytic digestion of DEPC-modified DHQase.

DHQase was modified with DEPC as described in Section 2.8.2. After 30min of

inactivation (at that time the enzyme is >90% inactivated) the protein was dialysed against 50mM sodium phosphate buffer, pH 7.0. The dialysed protein was transferred to acid washed, round bottom flasks containing small magnetic stirrers and digested with either substilisin, trypsin or chymotrypsin (2% of the mass of substrate) at 30°C with constant stirring. Aliquots were withdrawn at intervals and immediately frozen with dry-ice in ethanol and stored at -20°C. Samples were injected onto a μBondapak C18 (Waters) reverse-phase HPLC column for separation. Complete digestion was obtained in all cases within 3h under the conditions described in "Materials and Methods" (Section 2.8.3) (only chymotryptic profiles will be shown in Fig. 3.11 and 3.12). Chymotrypsin was chosen for differential peptide mapping of DEPC-modified DHQase.

3.5.2 Differential peptide mapping of DEPC-modified DHQase

DHQase was modified with DEPC in the presence and absence of an equilibrium mixture of substrate and product, as described above and in the "Materials and Methods" section. The modified proteins were then digested with chymotrypsin for $2^1/_2$h and fractionated by reverse phase HPLC using a solvent system maintained at pH6.3 (solvent system I) and peptides detected at 240nm (the absorbance maximum of carbethoxylated histidine residues, Fig. 3.11) and 220nm (a common wavelength for peptide detection, Fig.3.12). Detailed inspection of the 240nm profiles which indicated the presence of many of *N*-carboethoxyhistidine containing peptides showed that only one region in the peptide map (~41min) is substantially protected by the ligand. DEPC is a hydrophobic reagent and as such it altered the hydrophobicity of modified peptides *i.e.* the hydrophobicity of the histidine containing peptides was altered by carbethoxylation (Biscoglio *et al.*, 1986). Comparison of the peptide maps of samples of DEPC-modified dehdroquinase obtained in the presence and absence of substrate protection enabled us to identify one protected peptide which was located towards the hydrophobic region of the chromatogram (~41min)

Table 3.4

Automated Edman degradation of the active site peptide

Cycle number	Residue	Yield (pmoles)
1	Ser	320
2	Asn	339
3	His	171
4	Asp	808
5	Phe	525
6	His	288
7	Lys	638
8	Thr	121
9	Pro	326
10	Glu	299
11	Ala	389
12	Glu	243
13	Glu	263
14	Ile	157
15	Ile	228
16	Ala	228
17	Arg	176
18	Leu	75

(Fig. 3.11 and 3.12). Other regions within the maps show much less pronounced changes and these, unlike the differences observed at 41min, were not reproducible between separate experiments.

3.5.3 Isolation and purification of an active site peptide

The material at 41min was collected from both the unprotected (Fig. 3.11) and protected (Fig. 3.11) samples (designated as X(-) and X(+), respectively) and rechromatographed by reverse phase HPLC using a TFA-based solvent system (solvent system II; Fig. 3.13). Solvent system II is run under conditions where the adduct would not be expected to survive (unlike solvent system I which is run under conditions in which carbethoxyhistidine is stable) but is used to fractionate further the peptide mixture represented within X(-) and X(+) (Fig. 3.11) and to desalt the samples. It is clear that the resulting peptide maps of X(-) and X(+) are very similar with the exception of a single major peak. This single differential peptide, designated X-1, was collected for amino acid sequence analysis (Table 3.4). The peptide was 18 residues in length and the sequence was determined to be:

S N H D F H K I P E A E E I I A R L

The differential peptide mapping of DEPC-modified DHQase was performed twice (on separate batches of enzyme) and a single chymotryptic peptide with the above sequence was obtained in each experiment. The peptide corresponds to residues 141-158 of the deduced amino acid sequence of *E. coli* DHQase (Chaudhuri *et al.*, 1991). Further differential peptide mapping experiments using double digests with chymotrypsin and trypsin were carried out in the same manner and again one difference was seen between the

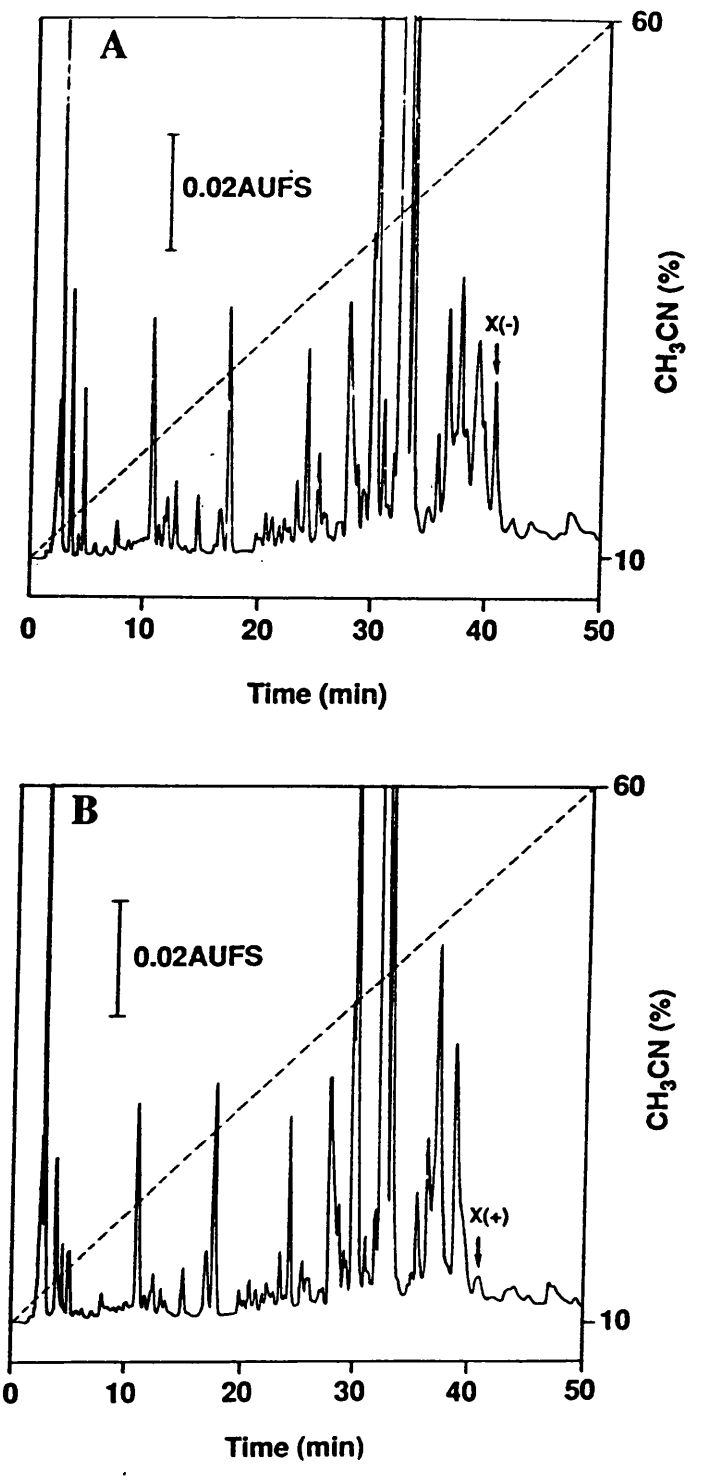


Fig. 3.11 Reverse phase HPLC profiles, monitored at 240nm using solvent system I (pH6.3), of chymotryptic digested, DEPC-modified DHQase in the absence (A), and presence (B), of substrate

Experimental details are given in "Materials and Methods" (Section 2.8). Identical quantities of protein were injected (10 nmoles). The dashed line in each experiment is the acetonitrile gradient used to elute the peptides. Detection at 240nm detects the N-carbethoxyhistidine adduct. *AUFS*, absorbance units full scale.

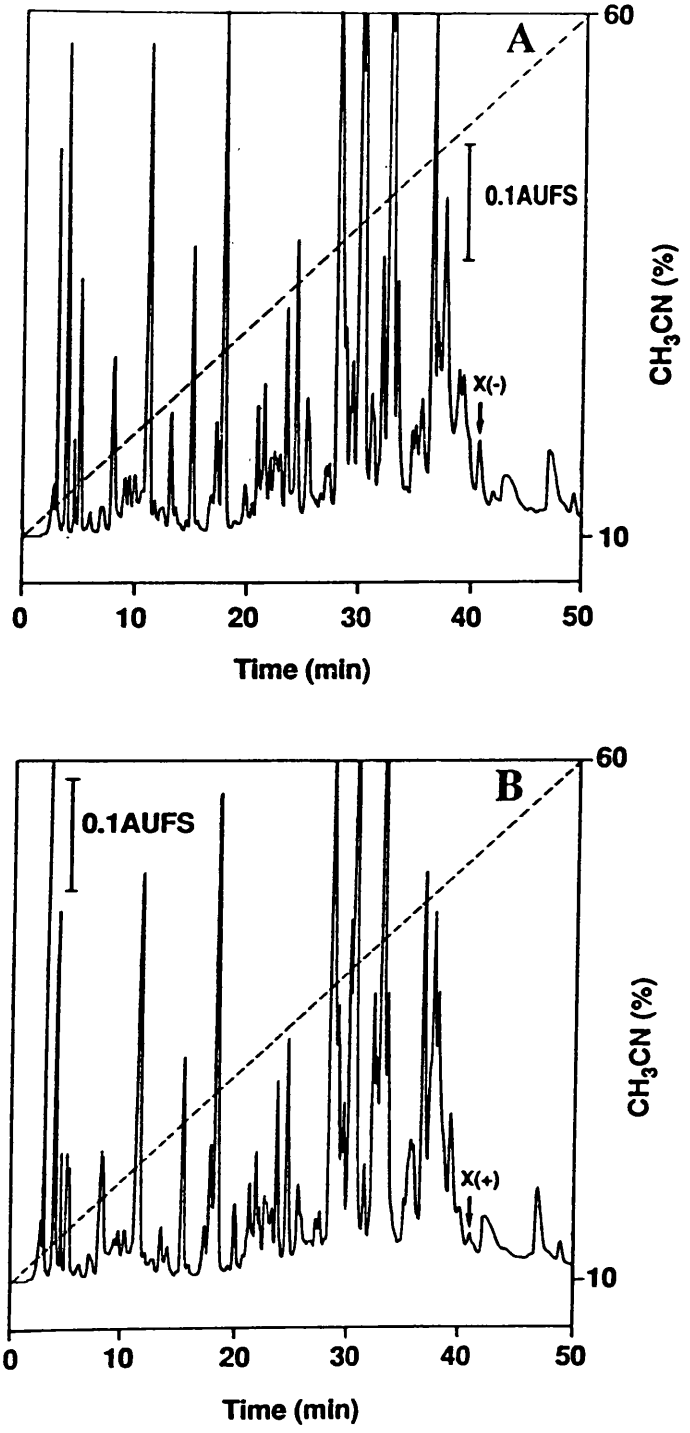


Fig. 3.12 Reverse phase HPLC profiles, monitored at 220nm using solvent system I (pH6.3), of chymotryptic digested, DEPC-modified DHQase in the absence (A), and presence (B), of substrate

Experimental details are given in "Materials and Methods" (Section 2.8). Identical quantities of protein were injected (10 nmoles). The dashed line in each experiment is the acetonitrile gradient used to elute the peptides. Detection at 220nm detects all peptides. *AUFS*, absorbance units full scale.

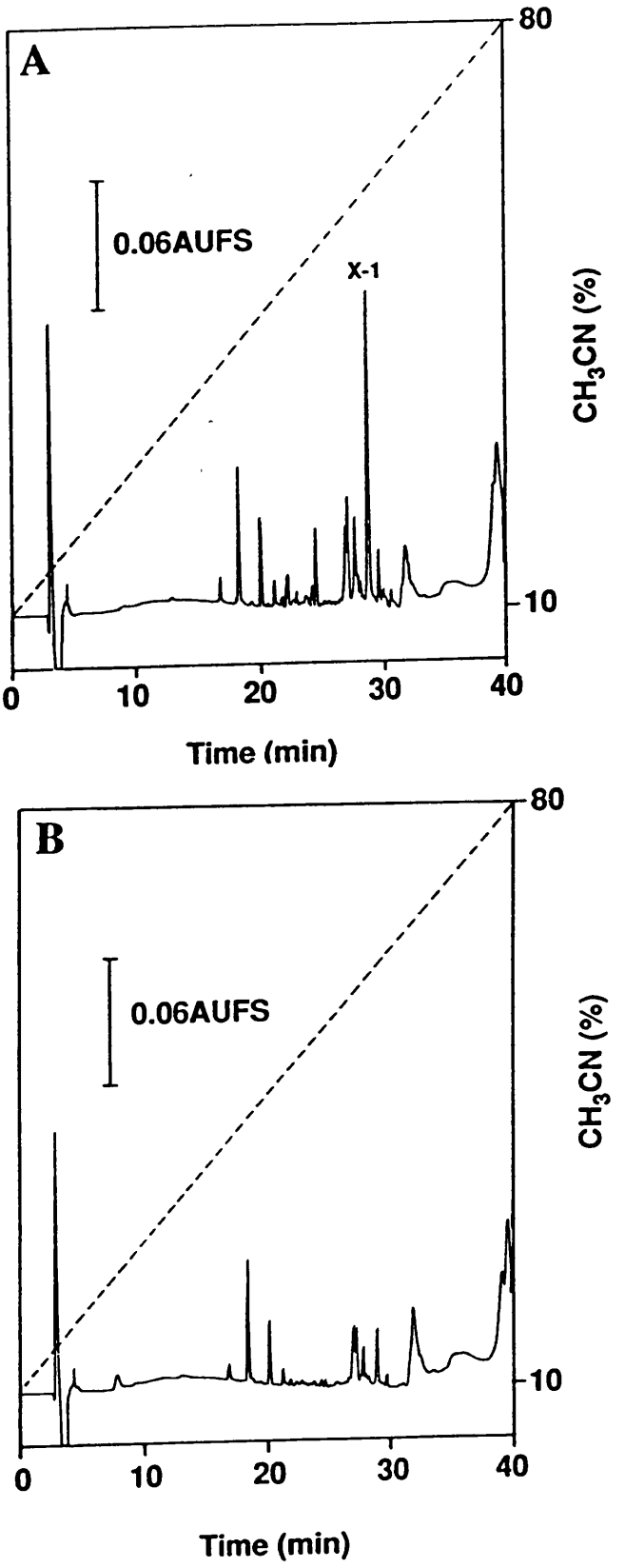


Fig. 3.13 Reverse phase HPLC profiles, monitored at 220nm and using solvent system II (pH2), of fraction X(-), (A), and fraction X(+), B
AUFS, absorbance units full scale.

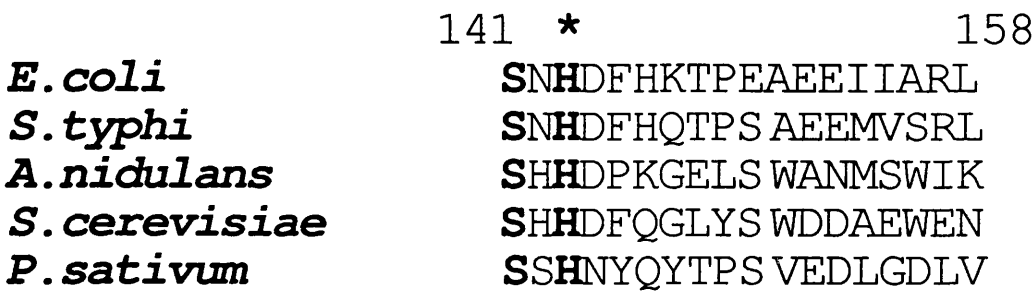


Fig. 3.14 Sequence alignments of the active site peptide of *E. coli* DHQase, identified by DEPC modification, with the corresponding regions in other type I DHQases

The numbering is according to the *E. coli* enzyme. *Bold faces* are residues conserved in all five sequences. * mark indicates His-143 in the *E. coli* enzyme. This alignment is taken from the complete alignments of all four sequences reported by Servos *et al.* (1991) and the *P. sativum* sequence reported in Chapter VI.

protected and unprotected samples. Sequence analysis of this peptide yielded a shorter version (residues 141-147) of the peptide identified above. These two separate experiments led us to conclude that we had isolated a region of the enzyme chain which contained the active site histidine. Since the peptide contained two histidine residues in close proximity to one another, separated by aspartic acid and phenylalanine, it was expected that chymotrypsin would cleave on the C-terminal side of the phenylalanine residue and so allow separation of the two histidine residues. However, even though we used chymotrypsin to obtain the peptides we did not succeed in isolating a peptide cleaved between the two histidine residues. It is not clear why chymotrypsin did not cleave on the C-terminal side of the phenylalanine residue of the peptide.

3.5.4 Active site homology

The active site peptide contains two histidine residues and it is not possible, just from this sequence, to deduce which of the two is the site of labelling by DEPC. However, comparison of this region with the same region from other type I DHQase sequences reveals that, of the two, only His-143 is conserved (Fig. 3.14). In fact, of all the six histidine residues in the *E. coli* DHQase sequence only His-143 is conserved between all the known type I sequences (Servos *et al.*, 1991; this work).

3.6 Discussion

3.6.1 The role of His-143

Two critical active site amino acid side chains have been proposed to be central to the chemical mechanism of the type I DHQases, the ε-amino group of a lysine residue and a

general base (Fig. 3.1; Walsh, 1979). The lysine residue forms a Schiff's base with the substrate (Chaudhuri *et al.*, 1991; Shneier *et al.*, 1991) which serves to stabilize the carbanion and provides an explanation for the unusual stereochemistry of the elimination reaction. The general base was postulated to participate in the mechanism because generation of the carbanion (in a stereospecific manner) requires abstraction of the pro-R proton. The active site lysine residue has been identified in the type I enzyme from *E.coli* and in the type I enzyme located on the fourth domain of the pentafunctional *arom* multifunctional enzyme of *N. crassa* (Chaudhuri *et al.*, 1991). This lysine residue is conserved in all type I enzymes (Chaudhuri *et al.*, 1991; Servos *et al.*, 1991). The evidence for the identity of the postulated general base has been based on preliminary DEPC modification data and the pH activity profile for the enzyme which is consisent with the involvement of an unprotonated active site group of pK_a 6.1 in the mechanism (Chaudhuri *et al.*, 1991).

DEPC is commonly used for modifying histidine residues in proteins (Miles, 1977; Lundblad and Noyes, 1984), but it can also react with other types of amino acid side chain (Muhlrad *et al.*, 1967; Miles, 1977). Incubation of type I DHQase from *E. coli* with DEPC resulted in a time-dependent inactivation of the enzyme and kinetic analysis revealed that this inactivation was due to a simple bimolecular reaction. The second-order rate constant for inactivation (148.5 $M^{-1}min^{-1}$) is higher than the values frequently found for proteins containing catalytically essential histidine residues (Lundblad and Noyes, 1984; Church *et al.*, 1985) and much higher than those observed for other potentially reactive side chains such as those of lysine, cysteine or tyrosine residues (Holbrook and Ingram, 1973; Wells, 1973). Cysteines can certainly be ruled out as potential sites of modification since DHQase contains no cysteine residues (Chaudhuri *et al.*, 1991). Tyrosines can also be eliminated since there was no decrease in the difference spectra (Fig. 3.5.A) at 280nm (Burstein *et al.*, 1974; Miles, 1977). Modification at histidine residues is supported by the observation that the bulk of the enzyme activity can be regenerated by hydroxylamine

treatment to yield enzyme with a similar K_m to native DHQase. CD spectra indicate that the inactivation was not due to any conformational change of DEPC-modified DHQase (Fig. 3.7).

The modification of histidine residues by DEPC is confirmed by the difference spectroscopy results at 240nm. This method directly monitors the carbethoxylation of histidine residues in proteins (Ovadi *et al.*, 1967; Miles, 1977). Complete inactivation of the enzyme is accompanied by the modification of all six histidine residues per subunit. To establish the number of essential residues the method of Tsou (1962) was employed (Fig. 3.8). This analysis suggests that the modification of a single histidine residue is responsible for the inactivation of DHQase. Kinetic analyses were utilized to estimate the order of the inactivation reaction. The results indicated that modification of a single histidine residue was responsible for inacivation of the enzyme (Fig. 3.3.C). Further evidence for the involvement of a single histidine residue was obtained by measuring the stoichiometry of modification of enzyme with and without the product Schiff's base covalently trapped at the active site. This comparison showed that a single histidine was protected by the trapped product. pH studies of the DEPC modification of DHQase reveal that the pK_a value of 6.18 for the inactivation is identical to the observed pK_a (6.1) for the catalysed reaction. We therefore conclude that there is a single, highly reactive histidine residue in the active site of *E. coli* DHQase, which is likely to be the general base in the mechanism of the enzyme.

Histidine participates in large number of enzymatic reactions by a variety of mechanisms, which include general acid/base catalysis, electrophilic catalysis, and binding of the substrate via hydrogen bonding and/or electrostatic interactions (Carter and Wells, 1988; Warshel *et al.*, 1989; Nickbarg *et al.*, 1988; DiPersio *et al.*, 1991; Miran *et al.*, 1991; Kleanthous *et al.*, 1985). Having inferred the catalytic role of the histidine residue by chemical modification, we used differential peptide mapping to identify the residue as

His-143. The mapping was done under conditions where the *N*-carbethoxyhistidine derivative is sufficiently stable to allow purification (mildly acidic conditions, ~pH6) (Vangrysperre *et al.*, 1989). The histidine peptide was located by comparing the peptide maps of enzyme modified in the absence and presence of an equilibrium mixture of substrate and product. Sequencing of this peptide has identified two histidine residues as possible candidates for the general base side. However, only one of these, His-143, is conserved in all the type I enzymes. We therefore propose that His-143 is a good candidate for the general base (depicted in Fig. 3.1) involved in the postulated mechanism of the *E. coli* type I DHQase. The identification of this residue opens up the way for further experiments such as site-directed mutagenesis aimed more directly, at probing its role in the mechanism. It is anticipated that when the crystal structure of the enzyme becomes available the proximity of His-143 to the active site will be confirmed.

CHAPTER IV

Mechanistic properties of type II dehydroquinase from *Aspergillus nidulans*

4.1 Introduction

This chapter describes some mechanistic properties of the type II DHQase from the quinate pathway of *A. nidulans* and compares them with the corresponding properties of the type I DHQase from the biosynthetic shikimate pathway of *E. coli*.

4.2 Steady-state parameters

The two types of DHQase are easily distinguished by their kinetic parameters (Table 4.1). Enzymic activity for both the type I and type II DHQases was measured as described in Section 2.4.1. Different buffers were used for each of the DHQases because, in general, buffers which are appropriate for one are inhibitors for the other. For example, potassium phosphate used for assaying the type I enzymes, is a competitive inhibitor of the type II enzyme from *A. nidulans* (K_i = 10mM; Fig. 4.1), whereas acetate and chloride are not (chloride and acetate are competitive inhibitors of the type I enzyme from *E. coli*; Chaudhuri *et al.*, 1986). As a result, 100mM potassium phosphate, pH 7.0, was used for the *E. coli* type I enzyme while 50mM Tris/acetate, pH 7.0, was used for the *A. nidulans* type II enzyme. Under these conditions, the specific activity for the type II DHQase was found to be an order of magnitude greater than that for type I, a difference which means that there is a 10-fold difference in the turnover numbers (k_{cat}) for the enzymes. Nevertheless, the specificity constants (k_{cat}/K_m) for both enzymes (at this pH and temperature) are surprisingly similar which is due to the 10-fold increase in K_m^{DHQ} for the type II enzyme as compared with the type I. The concentration of dehydroquinate used for routine assays of the type II enzymes was 1000μM against 100μM for the type I enzyme. These concentrations represent approximately seven times K_m for both the enzymes.

Table 4.1

Kinetic constants for type I and type II dehydroquinases at pH 7.0 and 25^{0}C

Parameters	Type I (*E. coli*)	Type II (*A. nidulans*)
Specific activity* (μmol min^{-1} mg^{-1})	295	4736
k_{cat} (s^{-1})	135	1300
K_m (μM)	16	150
k_{cat}/K_m (s^{-1} M^{-1})	8.7 x 10^6	8.68 x 10^6

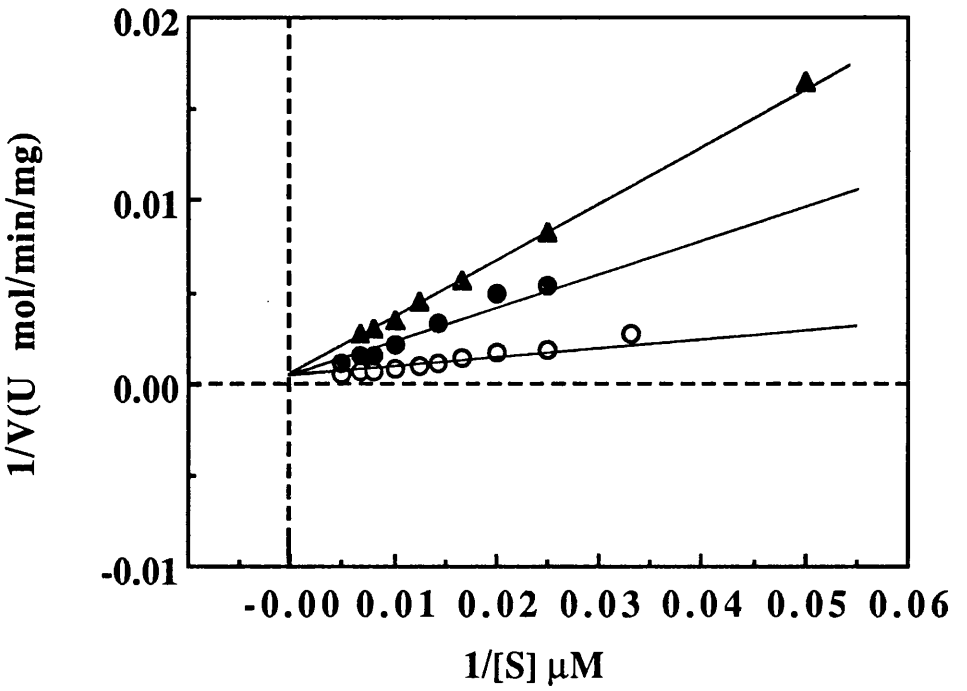
* Protein concentrations were determined by amino acid analysis
Values for type I enzyme were determined in 100mM phosphate buffer
Values for type II enzyme were determined in 50 mM Tris/Acetate

Fig. 4.1 Competitive inhibition (K_i) of the type II DHQase for potassium phosphate

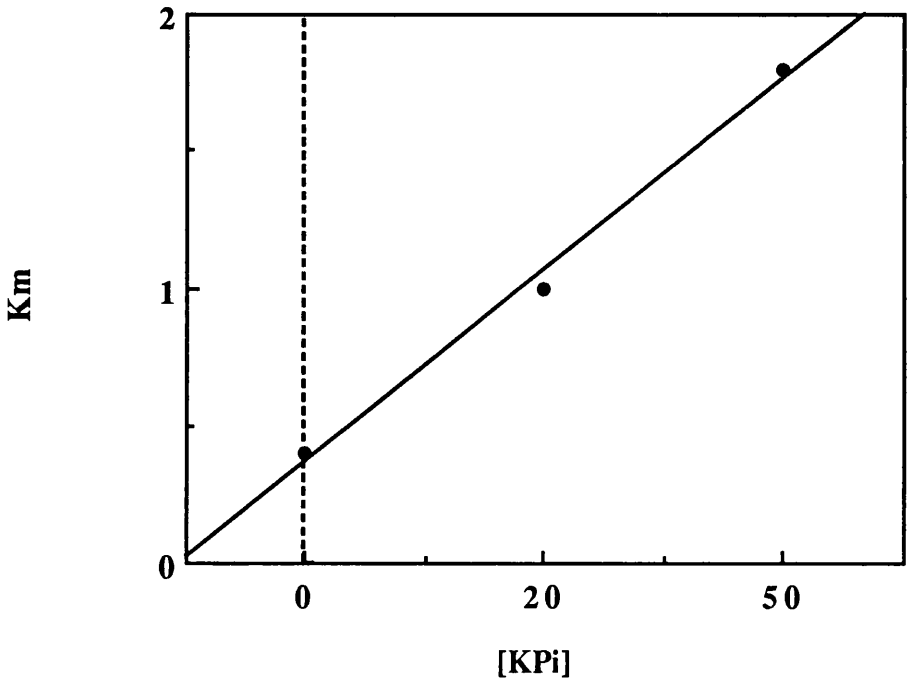
A. Double-reciprocal plots of DHQase at 0mM (O), 20mM (●) and 50mM (▲) potassium phosphate (KPi), pH 7.0 in 50mM Bis-trispropane/acetate, pH 7.0 and at 25°C.

B. Replot of K_m values as a function of [KPi].

A


0.02
0.01
0.00
-0.01
1/V(U mol/min/mg)
-0.00 0.01 0.02 0.03 0.04 0.05 0.06
1/[S] µM

B


2
1
0
Km
0
20
50
[KPi]

4.3 Chemical modification of type II DHQase with different reagents

4.3.1(a) Diethylpyrocarbonate

The type II DHQase of *A. nidulans* is very susceptible to inactivation by the histidine modifying reagent DEPC. Control experiments showed no loss of enzymatic activity. The inactivation does not obey pseudo first-order kinetics (Fig. 4.2.A). Approximate first-order inactivation rate constants can be calculated from the first time points in Fig. 4.2.A. Plotting these estimated rate constants against [DEPC] gives a linear plot passing through the origin (Fig. 4.2.B) which suggests that there is not a reversible complex of enzyme and modifier prior to the inactivation process. From the slope of the plot, a second-order rate constant of $1000M^{-1}min^{-1}$ for the inactivation was calculated.

The inactivation is substantially protected by an equilibrium mixture of substrate and product (Fig. 4.2.C), suggesting that inactivation is due to modification of active site residues.

Fig. 4.2 Inactivation of type II DHQase by DEPC

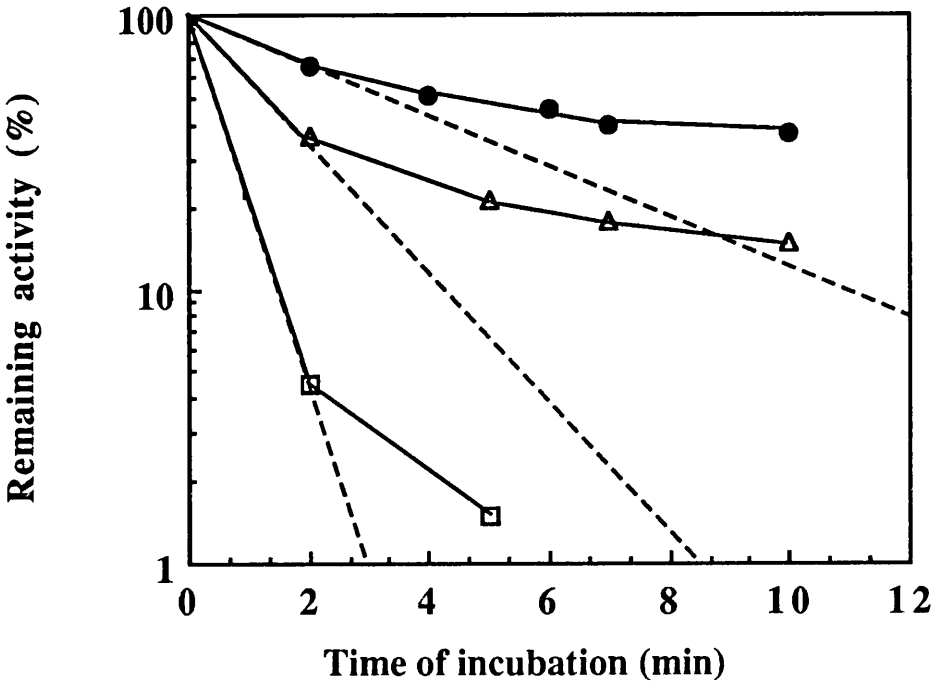
A. Time dependence of DEPC inactivation

The enzyme was incubated with different concentrations of the reagent in 50mM-bicarbonate buffer pH 8.5, at 25°C. At time intervals, aliquots were removed for measurements of the remaining activity as described in "Materials and Methods" (Section 2.9.1). The concentrations of DEPC were 0.25mM (●), 0.5mM (Δ) and 1mM (□).

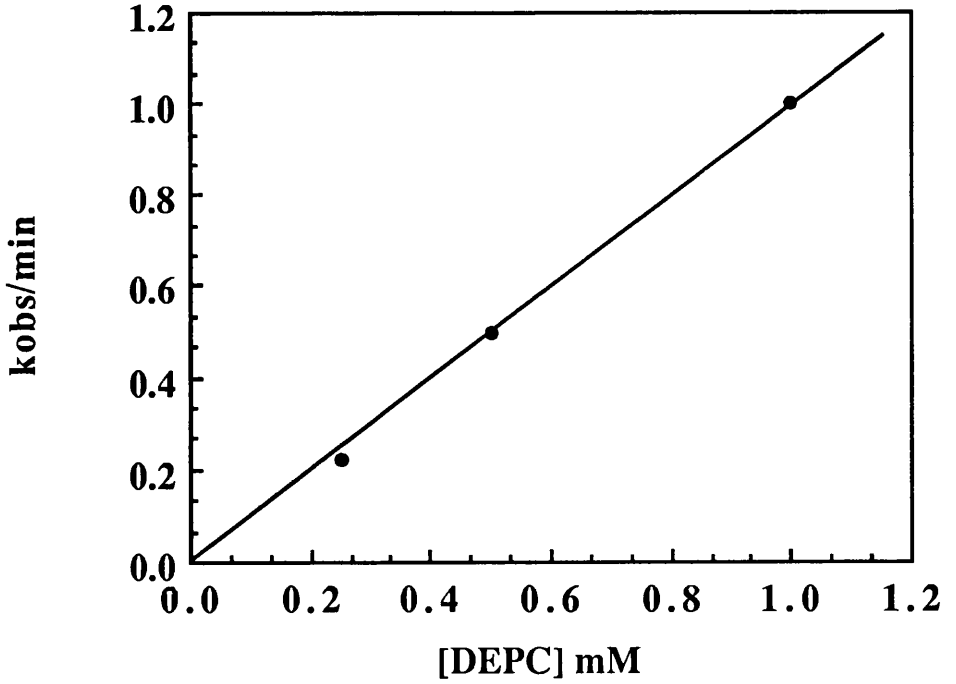
B. Determination of the second-order rate constant of inactivation

Approximate pseudo first-order rate constants for inactivation (k_{obs}) calculated (based on the first time points) in part *A* were proportional to [DEPC]. The slope of this line, calculated to be $1000M^{-1}min^{-1}$, is the second- order rate constant (k) of inactivation.

A


100
10
1
Remaining activity (%)
0
2
4
6
8
10
12
Time of incubation (min)

B


1.2
1.0
0.8
0.6
0.4
0.2
0.0
kobs/min
0.0
0.2
0.4
0.6
0.8
1.0
1.2
[DEPC] mM

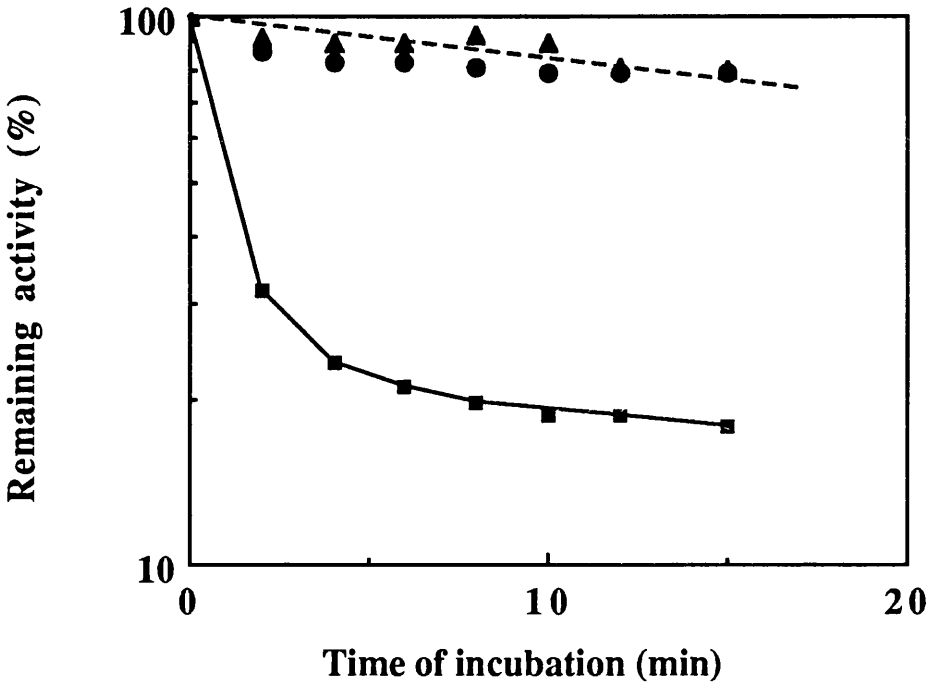


Fig. 4.2 C Substrate protection against 0.5mM DEPC inactivation

DHQase was incubated as described in Fig. 4.2 *A* in the presence and in the absence of substrate. The substrate concentrations used were 0.0mM (■), 5mM (●), and 10mM (▲).

4.3.1(b) Reactivation of DEPC-inactivated DHQase with hydroxylamine

When the DHQase (7.52μM) activity had been reduced by 80% (within 10min) with 0.5mM DEPC, 70% of the activity could subsequently be restored within 30min by 0.5M-hydroxylamine treatment as shown in Table 4.2 and Fig. 4.3. These observations suggested that the loss of activity is due to modification of histidine residues (Lundblad and Noyes, 1984) which was further supported by the observed increase in absorbance at 237nm (but not at 278nm) which accompanied DEPC modification (Fig. 4.4).

Table 4.2

Reversal of DEPC inactivation upon hydroxylamine treatment

Time	% Recovery
2min	25%
5min	29%
10min	44%
15min	52%
20min	62%
25min	57%
30min	72%
40min	67%
120min	69%
180min	75%
300min	76%

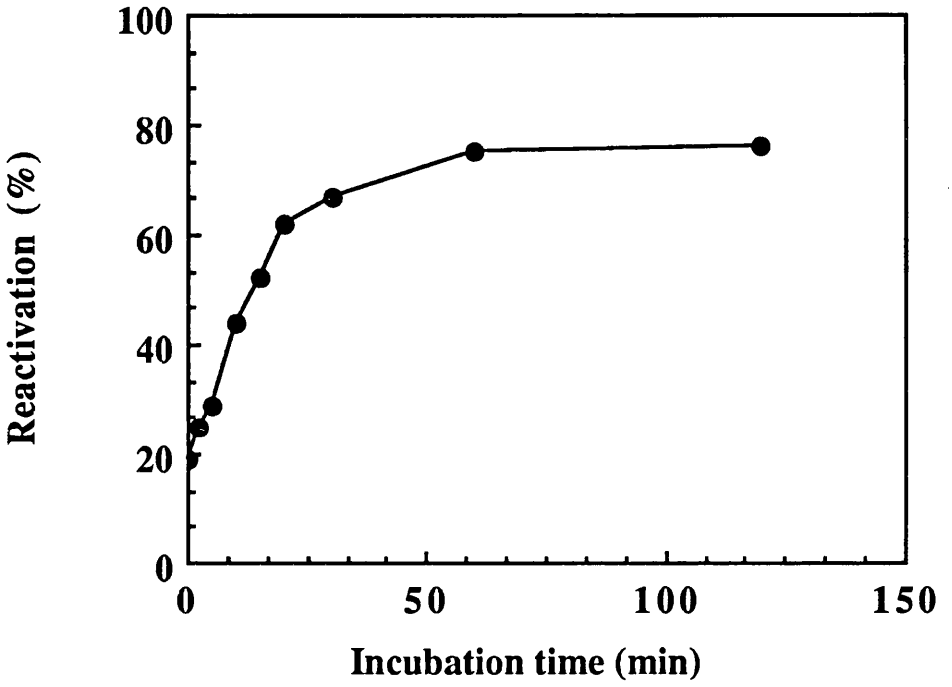


Fig. 4.3 Reactivation of DEPC-inactivated DHQase by hydroxylamine DHQase (7.52μM) was incubated with 0.5mM-DEPC in 50mM-bicarbonate buffer, pH 8.5 at 25°C. After 10min of incubation (at this stage enzyme is >80% inactive), hydroxylamine (500mM) was added to both DEPC-modified and native enzyme. % of activity recovered over the control was determined at various time intervals (also see Table 4.2).

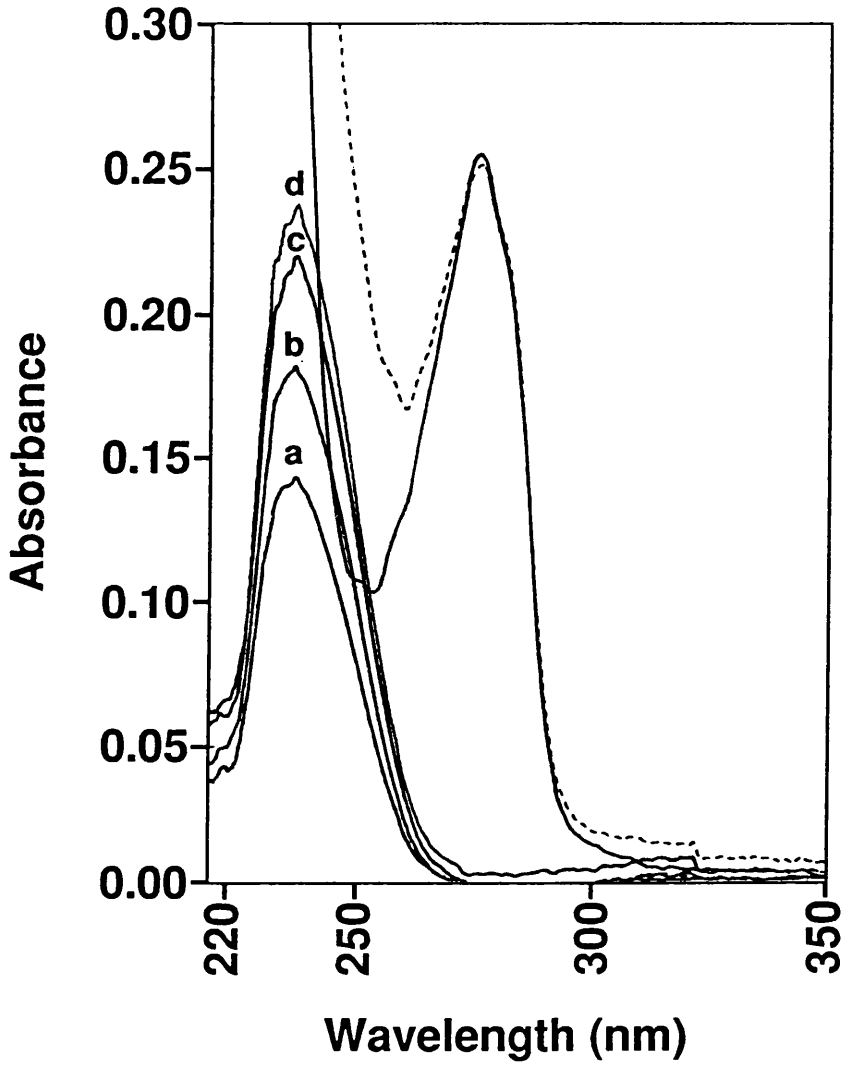


Fig. 4.4 Difference spectra for modification of DHQase by DEPC

The difference spectrum of the solution containing 0.5mM DEPC and 7.52μM enzyme in 50mM-bicarbonate buffer, pH 8.5 at 25°C, against the solution without DEPC was measured in the wavelength region from 220 to 350nm (for details see "Materials and Methods", Section 2.2.7). The increases in absorbance at 237nm in the difference spectra obtained at various times (**a**, 1min; **b**, 2min; **c**, 4min and **d**, 6min) indicated the formation of N-carbethoxyhistidine (see Fig. 3.2 in Chapter III). Absorbance traces of unmodified protein (*solid line*) and protein after 6min of modification (*dotted line*) are also shown.

4.3.1(c) The order of the inactivation reaction

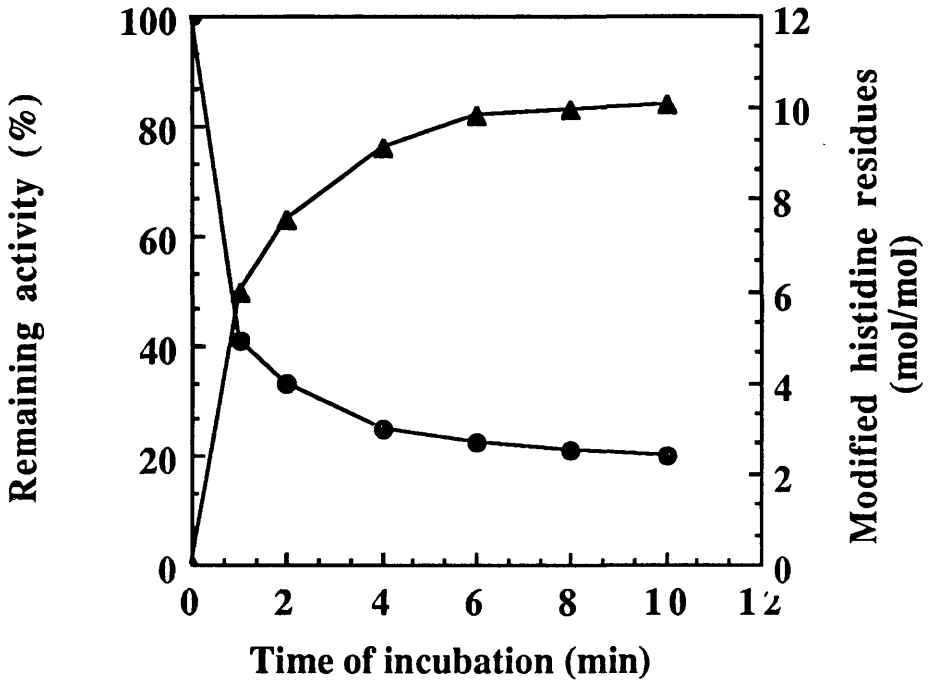
There are 10 histidine residues in the primary structure of the type II DHQase of *A. nidulans*. The number of histidine residues modified upon incubation with DEPC calculated from the increase in absorbance at 237nm ($\Delta\varepsilon_{237}$ = 3200$M^{-1}cm^{-1}$; Miles, 1977) was found to be 10 when approximately 20% of the activity remained (Fig. 4.5.A). This provides only an estimate of the stoichiometry of modification.

Fig. 4.5 reveals the extent of modification when 7.52μM DHQase was incubated with 0.5mM DEPC at pH 8.5. Examination of the data in DHQase inactivation experiments revealed that modification of the enzyme can be distinguished from its inactivation (Fig. 4.5.A). Over a suitable time interval, the extent of inactivation was compared with the extent of modification. Inactivation is expressed as the fraction of residual activity, while the extent of modification is presented as the number of residues modified ($\Delta\varepsilon_{237}$ = 3200$M^{-1}cm^{-1}$; Miles, 1977). If all of the histidine residues modified were essential for activity, the extent of modification would be expected to parallel the extent of inactivation. The results shown in Fig. 4.5.A, reveal that this was not the case. From this plot, it is evident that after 30 sec, 60% of the total modification events had been completed while the enzyme retained 40% of its original activity. After 8min there was no more increase in absorbance, indicating that the modification was nearly completed. At this stage enzyme still retained 20% of its original activity. In an similar experiment using 1mM DEPC, it was possible to inactivate the enzyme up to 90% without further increase in absorbance at 240nm (Fig. 4.5.B). Interpretation of this data indicates that there are at least two populations of histidine residues which are modified in DHQase, and that they are differentiated by their reactivity towards DEPC. The data suggest that there is one group of essential histidine residues which react faster than a second group of nonessential residues; the modification of the first group results in the inactivation of the enzyme. It was not possible to analyse these data by using Tsou's statistical method, since the method is only

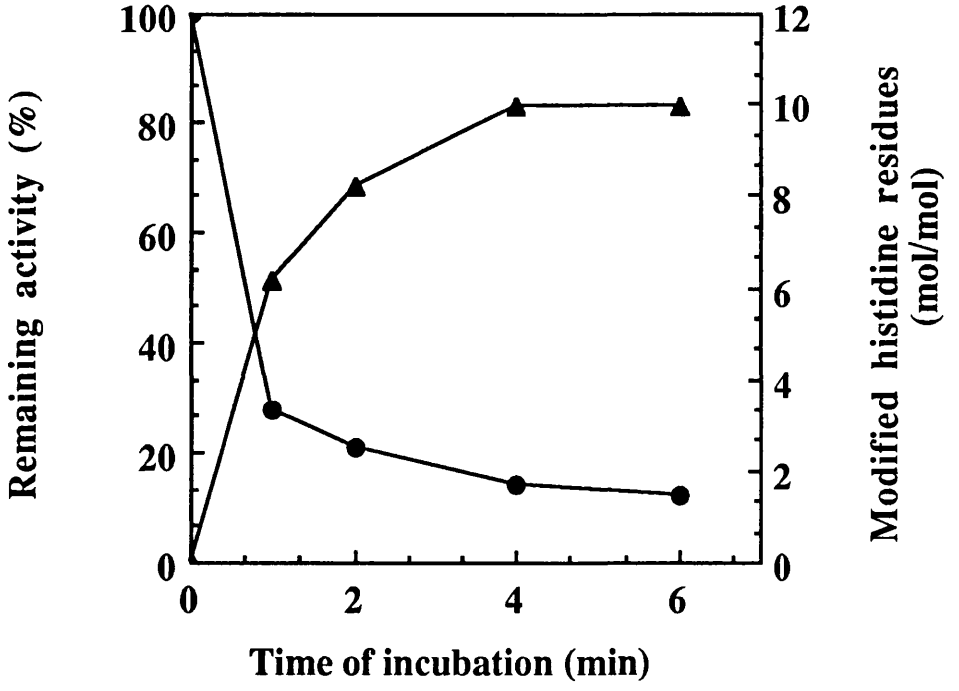
Fig. 4.5 Stoichiometry of modification and inactivation by DEPC

Inactivation experiments were conducted with 0.5mM (*A*) and 1mM (*B*) DEPC as described in Fig. 4.4. At various times the fraction of remaining activity (●) was measured. The number of modified histidine residues (▲) was calculated out as described in "Materials and Methods" (Section 2.7.7).

A


Remaining activity (%)
Modified histidine residues (mol/mol)
Time of incubation (min)
0
20
40
60
80
100
0
2
4
6
8
10
12

B


Remaining activity (%)
Modified histidine residues (mol/mol)
Time of incubation (min)
0
20
40
60
80
100
0
2
4
6
8
10
12

valid if both essential and non-essential residues react in a pseudo first-order manner (Tsou, 1962; Horiike *et al.*, 1979). However, kinetic analysis was used to determine the order of the inactivation and thereby to determine the number of essential histidine residues (Levy *et al.*, 1963), using the equation (5) (Section 3.4.2). As shown in Fig. 4.6, the plot of log ($1000/t_{1/2}$) vs. log [DEPC] (Levy *et al.*, 1963) for type II DHQase was linear. From the slope of a plot, a reaction order of 1.05 was obtained, indicating that the covalent modification of a single histidine is required to inactivate the enzyme.

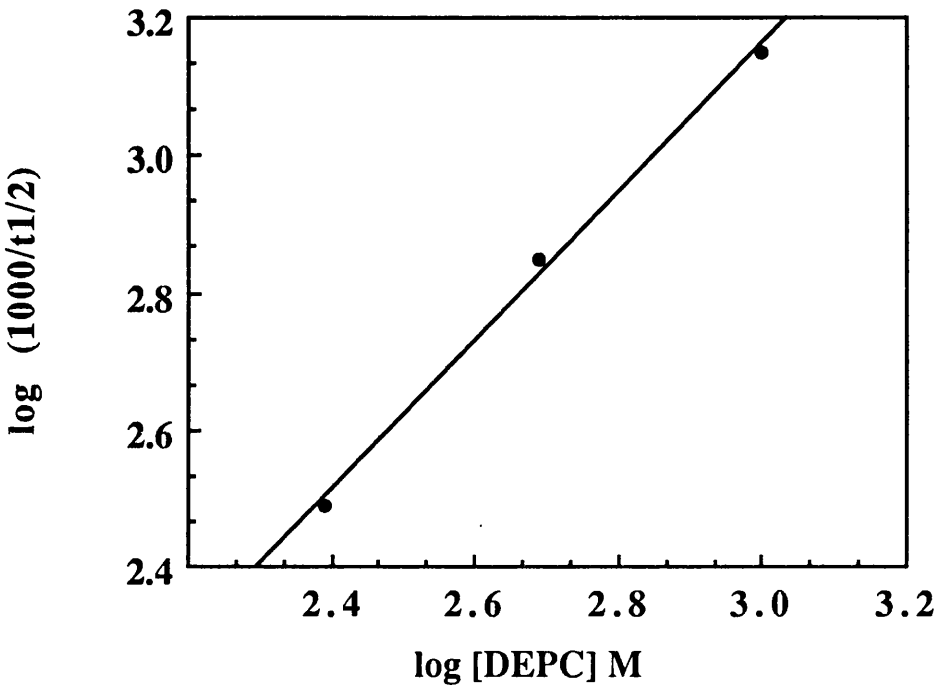


Fig. 4.6 Double-logarithmic plot of the pseudo first-order rate constant (k_{obs}) of inactivation against [DEPC]

Data were taken from the Fig. 4.2 *A* and plotted as described in Section 3.3.2. The slope, n = 1.05, is the average reaction order (*k*) for the inactivation of the enzyme by DEPC.

4.3.1(d) pH dependence of DEPC inactivation

DEPC modifies only unprotonated histidine residues, and thus the pH dependence of inactivation provides information about the pK_a of the residue modified with inactivation of the enzyme. It was found that slower inactivation occurred below pH 6.5, therefore, the pH dependence of inactivation was investigated between pH 6.5 and 7.5 in Pipes buffer. Pseudo first-order rate constants were estimated from the initial 50% of inactivation (thereafter inactivation became more complex) and plotted using equation (8) (Section 3.3.10). The pK_a value was found to be 6.95 (Fig. 4.7). The best candidate for an amino acid residue with this pK_a is histidine. Similar pK_a values for histidine residues at the active sites of other enzymes have been reported (Holbrook and Ingram, 1973; Blanke and Hager, 1989). This data suggests that the inactivation of type II DHQase by DEPC is due to the modification of a histidine residue.

4.3.2 Borohydride reduction of imine intermediates

Sodium borohydride reduces the Schiff base intermediate formed between the ligand and the primary amino group of enzyme. Reduction covalently traps the intermediates resulting in permanent inactivation of enzyme. This type of inactivation has been observed for the type I DHQases and the active site lysine involved (Lys-170) has been identified (Chaudhuri *et al.*, 1991).

Attempts to inactivate of the type II enzyme from *A. nidulans,* (as described in Section 2.9.4) in the presence of a substrate/product equilibrium mixture, with sodium borohydride in 50mM bicarbonate buffer, pH 8.5 (for details see Kleanthous *et al.*, 1992) have failed. This implies that there is no role for a Schiff's base intermediate.

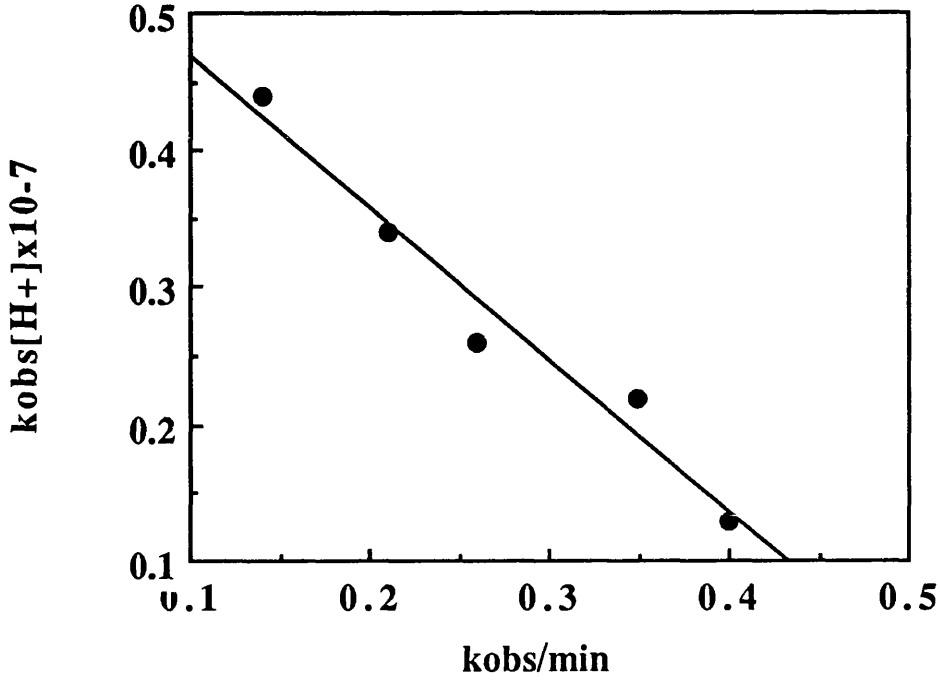


Fig. 4.7 Inactivation of DHQase by DEPC as a function of [H+]

The enzyme was incubated with DEPC at different pH values, as described in "Material and Methods" (Section 2.9.1). Pseudo first-order rate constant (k_{obs}) were obtained as in Fig. 4.2 *A* and plotted using Equation (8) in Chapter III.

4.3.3 Pyridoxal 5'-phosphate

A number of investigators have used pyridoxal 5'-phosphate (PLP) to modify lysine residues in proteins. This reagent is useful for the modification of lysine because of its selectivity of reaction.The spectral properties of the modified residue allow stoichiometry to be established. The reaction is reversible but can be made irreversible by using sodium borohydride to reduce the Schiff base initially formed on the reaction of PLP with a primary amine (Lundblad and Noyes, 1984). This reaction was performed in the dark as described in Section 2.9.3. There was no inhibition of the enzyme.

4.3.4 Phenylglyoxal

Phenylglyoxal (PG) shows a high degree of specificity in its reaction with arginine residues, both in model compounds and proteins (Takahashi, 1968; Riordan, 1979; Cheung and Fonda, 1979). The type II dehdroquinase of *A. nidulans*, was inactivated by phenylglyoxal (Fig. 4.8.A), which must be the result of arginine residue modification. Over the time period necessary to inactivate the enzyme, control samples did not undergo any loss in activity. As shown in Fig. 4.8.A, at each concentration of PG, the logarithim of the residual activity was a linear function of time, showing that inactivation was apparently first-order. A plot of pseudo first-order rate constants for inactivation (k_{obs}/min) against [PG] was linear and unusually there was an intercept on the [PG] axis (Fig. 4.8.B). The second-order rate constant (k) for inactivation at pH 9.4 and 25°C was $222M^{-1}min^{-1}$. The reaction order for inactivation with respect to PG concentration was obtained from a plot of log k_{obs} against log reagent [PG] (Fig. 4.8.C) (Levy *et al.*, 1963). The stoichiometry (reaction order) calculated from the slope of the plot yielded a value of 0.90, indicating that the modification of a single arginine results in the loss of enzyme activity. The chemical modification of arginine residues by incorporating

radiolabelled PG might be helpful in identifying the essential arginine residue.

The inactivation was partly protected by substrate dehydroquinate (Fig. 4.8.D). When 1mM PG was used considerable protection was found only by adding 20mM substrate. This may be due to the instability of the substrate at the high pH (9.4) used in the experiment.

Fig. 4.8 Effect of phenylglyoxal concentration on the rate of inactivation of type II DHQase

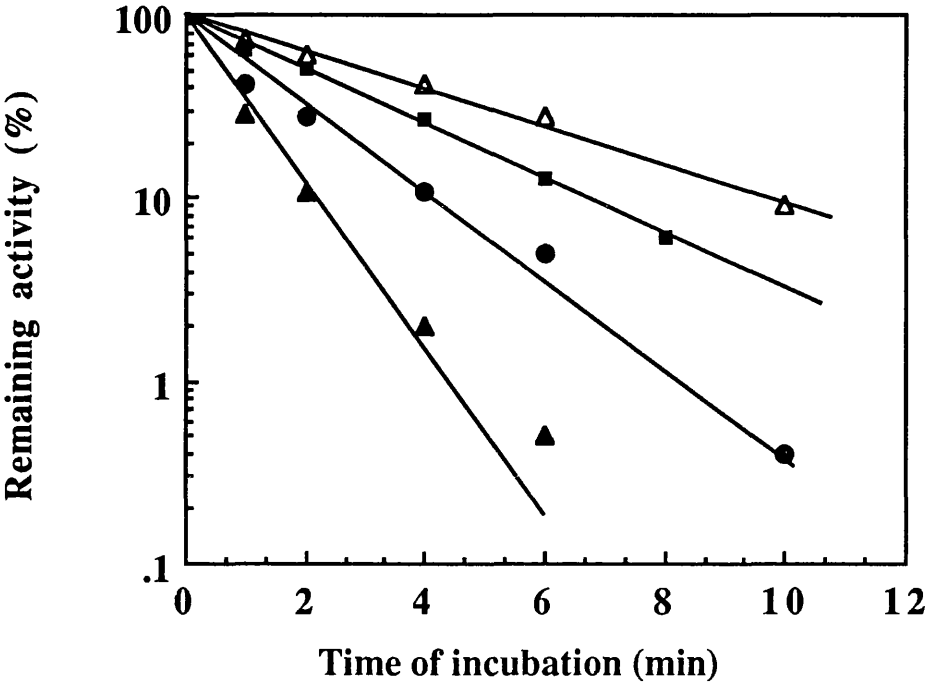
A. Time dependence of PG inactivation

The enzyme was incubated at 25°C with various concentration of PG in 100mM-bicarbonate buffer, pH 9.4 in a final volume of 1ml. At each indicated time, an aliquot was withdrawn and the remaining activity was assayed. PG concentrations were 0.5mM (Δ), 1mM (■), 2mM (●) and 4mM (▲).

B. Dependence of pseudo first-order rate constant on [PG]

Pseudo first-order rate constants were determined in part *A* as described in Section 3.2.2 and plotted versus [PG]. The slope of the line, calculated to be $222M^{-1}min^{-1}$, is the second-order rate constant (k) of inactivation at pH 9.4 and 25°C.

A


100
10
1
.1
Remaining activity (%)
0
2
4
6
8
10
12
Time of incubation (min)

B

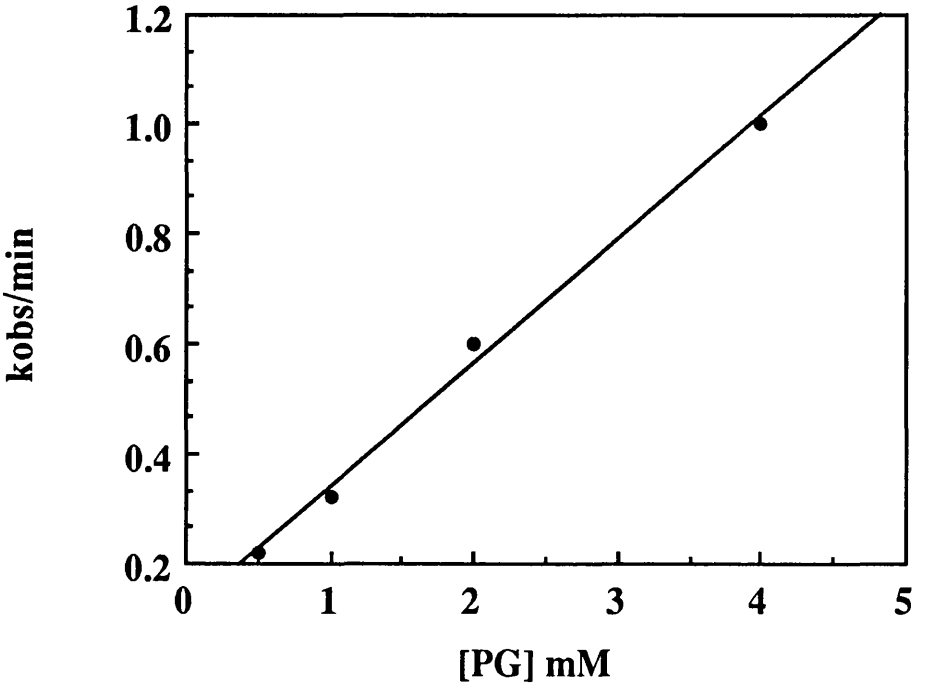

1.2
1.0
0.8
0.6
0.4
0.2
kobs/min
0
1
2
3
4
5
[PG] mM

Fig. 4.8 Effect of phenylglyoxal concentration on the rate of inactivation of type II DHQase

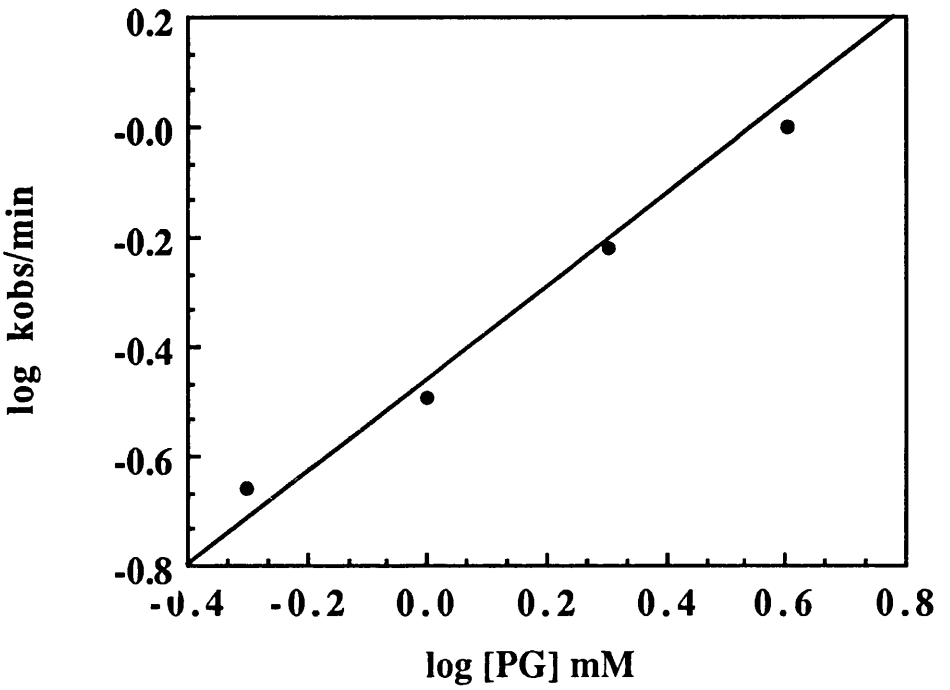
C. Double-logarithmic plot of the pseudo first-order rate constant (k_{obs}) of inactivation of DHQase against [PG]

The slope of this plot (0.90) is equal to the reaction order according to the Equation (5) in Chapter III. General conditions are those given in Fig. 4.8 *A*.

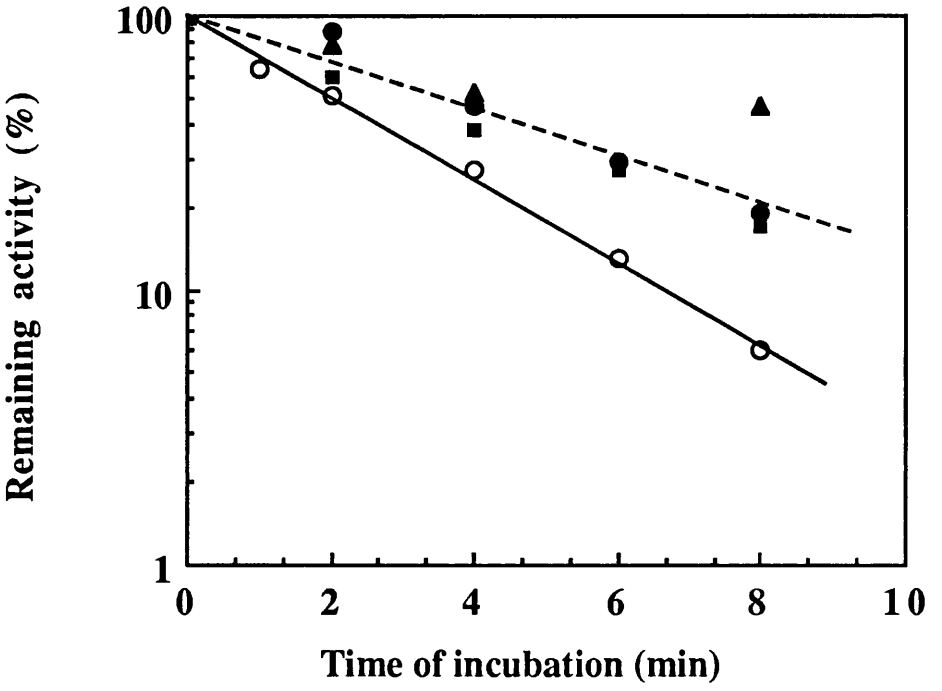
D. Substrate protection against inactivation by PG

Substrate concentrations were 0mM (O), 2mM (■), 4mM (●) and 20mM (▲). Conditions and procedure are given in Fig. 4.8 *A*.

C


0.2
-0.0
-0.2
-0.4
-0.6
-0.8
log kobs/min
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
log [PG] mM

D


100
10
1
Remaining activity (%)
0
2
4
6
8
10
Time of incubation (min)

4.4 Discussion

The inactivation of type II DHQase from *A. nidulans* by DEPC suggests that histidine residues are essential for enzyme activity. The spectral data support this conclusion and also exclude the possibility of the reaction of this reagent with tyrosine. All ten residues in the enzyme are accessible to DEPC with differing reactivities. The number of essential histidine residue(s) could not be determined. The absence of total restoration of activity of the DEPC-treated enzyme by hydroxylamine could be due simply to destruction of the enzyme during hydroxylamine treatment, or could be explained by the reaction of two DEPC molecules with one histidine residue (Miles, 1977). The rate of inactivation exhibited a pH dependence, indicating that modification of a titratable residue with a pK_a value of 6.95 was responsible for inactivation. All this data provides evidence for histidine being the site of modification by DEPC. Protection by the substrate dehydroquinate argues for the involvement of a histidine residue in the active site of the enzyme.

Borohydride in the presence of substrate and pyridoxal 5'-phosphate both of which react with amino groups, failed to inactivate the type II dehdroquinase from *A. nidulans,* indicating that the enzyme does not work through Schiff base formation.

The involvement of arginine residues, possibly in substrate binding, is suggested by the sensitivity of the enzyme to phenylglyoxal. Incubation of DHQase with this reagent resulted in a time-dependent inactivation of the enzyme. Kinetic analyses were utilized to estimate the order of inactivation and the results indicated that a single essential arginine residue was responsible for inactivation. Substrate protected poorly against inactivation; therefore, a definite assignment of arginine as an essential residue for DHQase activity is not possible at present.

4.4.1 Distinguishing type I and type II DHQases

Certain differences between the type I and type II enzymes have been noted (see Kleanthous *et al.*, 1992). Here we summarise the results of the present investigation which provide information about the differences between the two classes of enzyme.

a) Schiff base intermediate: The mechanism of type I DHQase from *E. coli* and the type I-like domain of the pentafunctional protein of *N. crassa* both involve a Schiff base intermediate (Chaudhuri *et al.*, 1991). We have attempted to inactivate the type II enzyme from *A. nidulans,* in the presence of a substrate/product equilibrium mixture, with sodium borohydride as has already been described for the *E. coli* enzyme (Chaudhuri *et al.*, 1991). Inactivation would indicate the formation of an imine intermediate in the mechanism of the type II enzyme. However, we have been unable to demonstrate substrate/product-associated borohydride inactivation of the *A. nidulans* type II DHQase as described in the Section 4.3.2. This result is further supported by the experiment with PLP which also failed to inhibit the enzyme. The failure of borohydride and PLP inactivation together with sequence comparisons (lack of a conserved lysine residue) with other type II enzymes (Garbe *et al.*, 1991) strongly supports the hypothesis that the type II enzyme from *A. nidulans* and other species does not function by a Schiff base mechanism.

***b)* pH *dependence of* V_{max}:** It is frequently observed that mechanistically important amino acid residues ionize at pH values relevant to the activity of the enzyme. The pH dependence of V_{max} for the *E. coli* type I enzyme has already been reported (Chaudhuri *et al.*,1986). The enzyme shows a simple dependence of V_{max} on pH, consistent with the presence of a single ionizing group at the active site with the pK_a of 6.1. V_{max} for the type I enzyme is maximal at pH 6.8 and remains almost unchanged up to pH 9.0. The dependence of V_{max} on pH for a type II enzyme has not been reported

previously and so it was determined for the purpose of comparison with the type I enzyme.

A problem was encountered, however, in preliminary trial experiments aimed at identifying appropriate buffer systems for this study. Most of the common buffers used in such studies (for example, those used in the study of the *E. coli* type I enzyme; Chaudhuri *et al.*, 1986) were found to be unsuitable for the type II DHQase from *A. nidulans*. These included phosphate- (a competitive inhibitor, Fig. 4.1), citrate-, maleate- (for which no enzyme activity could be detected under standard assay conditions) based buffers which were all found to be inappropriate for this study. Finally, we opted for the single broad-range buffer bistrispropane which did not alter the kinetic parameters of the enzyme (relative to Tris/Acetate under standard assay conditions). V_{max} values were obtained directly from curves fitted to the Michaelis-Menten equation and each curve was composed of 14-16 substrate concentrations, each of which was measured in triplicate. The dependence of V_{max} on pH for the type II enzyme from *A. nidulans* in bistrispropane/acetate is shown in Fig. 4.9. The results are summarised below:

(1) In contrast with the type I enzyme from *E. coli*, V_{max} is maximal for type II enzyme at pH 8.5-9.0.

(2) Whereas the pH dependence of V_{max} for the type I enzyme is consistent with a single ionizing group at the active site, this is not the case for the type II enzyme. The complex profile for the type II enzyme suggests that several ionizing groups may be involved. The data presented in Fig. 4.9 show that as the pH is increased from 6.3 to 7.5 there is a very pronounced increase in the V_{max} for the type II enzyme and then between pH 7.5-8.0 and 8.2-9.0 two 'bumps' are evident. The triangular points represent repeats at particular pH values suggesting that, for example, the increase between pH 8.4 and 8.5 is not an artifact. The exact pK_a of the effect could not be determined because of the complex nature of the rate profile. However, since rate changes occurred at pH values

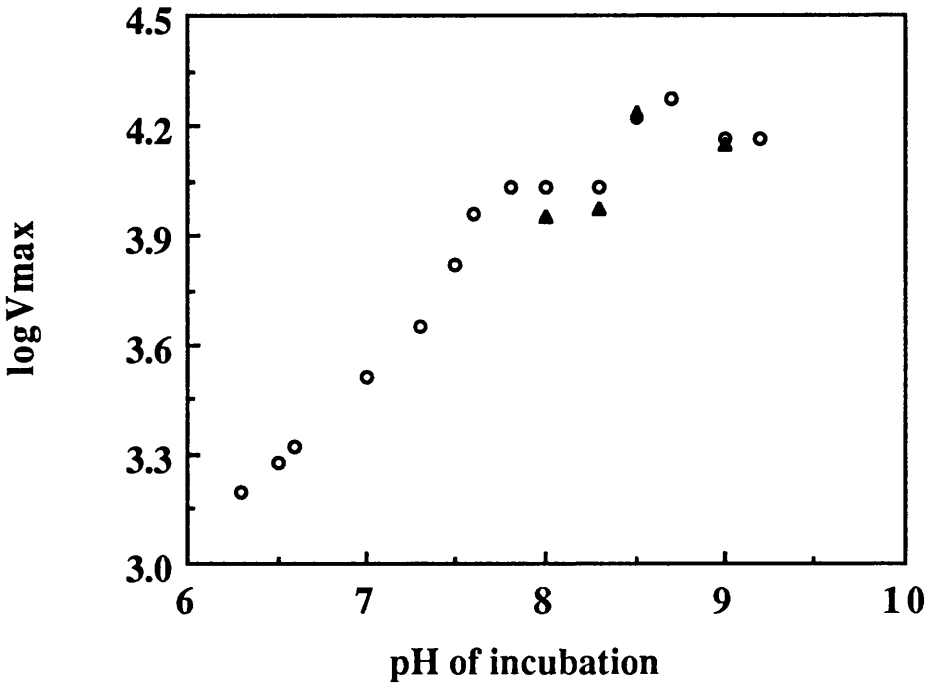


Fig. 4.9 pH dependence of V_{max} of the type II DHQase

between 6.3 to 9.2 it is unlikely that ionization of an arginine or lysine is involved since the pK_a of free arginine and lysine are usually above 10 in proteins. Although it is conceivable that the neighbouring groups in the vicinity of arginine or lysine residues could cause substantial changes in pK_a, the effect of pH on the rate is more probably caused by ionizations of other groups for example histidine or cysteine residues.

c) Histidine modification: Type I and type II DHQase are very susceptible to inactivation by the histidine modifying reagent DEPC (Fig. 3.3 and 4.2). The present results indicate that inactivation is the result of a bimolecular encounter between a single active-centre histidine residue and DEPC with a second-order rate constant of $1000M^{-1}min^{-1}$. Inactivation of type I *E. coli* enzyme by DEPC was studied and the inactivation rate constants were also linearly dependent on DEPC concentration, suggesting bimolecularity. The second-order rate constant for the reaction of the type I enzyme is ($148.5M^{-1}min^{-1}$) smaller by a factor of about 7-fold. These observations were made at pH 6.0 for the type I enzyme and pH 8.5 for the type II. Comparison of the protective effect of substrate suggests that, in both cases the inactivation is substantially protected by an equilibrium mixture of substrate and product suggesting that active-site residues have been modified. The pK_a values determined from the pH dependence of DEPC inactivation of type I and type II DHQase are 6.18 and 6.95 respectively and are qualitatively similar. These values are in good agreement with the pK_a of reactive histidine residues in other enzymes. In the case of the type I enzyme the essential histidine residue has been identified as His-143. No attempt has been made to identify the histidine(s) involved in the type II enzyme reaction. The overall conclusions from these results is that histidine is likely to be present in the active site of the type II DHQases.

d) Phenylglyoxal inactivation: In the case of the type II enzyme it was found that the enzyme was inactivated by PG. Kinetic analysis shows that there is one fast reacting arginine residue whose role has yet to be elucidated. The effect of PG on the type I enzymes has not been investigated.

4.4.2 Type I and type II DHQases are likely to function by different mechanisms

A characteristic feature of the type I DHQase catalytic mechanism (and this has been shown for both the *E. coli* enzyme and the type I domain of the *N. crassa* pentafunctional protein; Chaudhuri *et al.*, 1991) is the involvement of a Schiff base (imine) intermediate. This intermediate is formed at a conserved lysine residue (Lys-170, *E. coli*) and plays an important role as an electron sink in the mechanism. The involvement of this intermediate is easily shown by the addition of sodium borohydride to enzyme bound with substrate, which results in the inactivation of the enzyme as the active-site imine is reduced to a stable secondary amine (Chaudhuri *et al.*, 1991; Butler *et al.*, 1974). Attempts to show ligand-mediated borohydride inactivation and PLP inactivation of the type II enzyme from *A. nidulans* were unsuccessful. These results indicate that the type II DHQase does not function by a Schiff base mechanism. Sequence comparisons showed that unlike the Schiff base-forming type I enzymes, there is no conserved lysine residue in any type II enzymes. This further strengthens the earlier argument that the type II enzymes may act in a different manner. Kuriyan *et al.* (1991), from their structural comparison of *E. coli* thioredoxin reductase with glutathione reductase have shown that these two related enzymes catalysing similar reactions possess different active sites. Recently it has been shown that the type II enzyme proceeds with *anti* stereochemistry, in contrast to the *syn* elimination catalysed by the type I enzyme (Shneier *et al.*, 1992). All this data provides convincing evidence that the conversion of dehydroquinate to dehydroshikimate by the type II DHQase take place by a different mechanism from the conversion catalysed by the type I enzymes.

If the type II enzymes do not function by a Schiff base mechanism, then what alternatives can be proposed? Further work is necessary to address this question. Preliminary evidence for inactivation of the type II enzyme by the arginine specific reagent,

phenylglyoxal suggested the involvement of arginine residues. It would be interesting to see if the type II DHQase has any requirement for metal.

In conclusion, it appears that although both enzymes share the same overall reaction, their active sites do not resemble each other and their mechanisms are distinct. The precise role of histidine in the type II enzyme mechanisms remains to be established. The histidine may well be acting as a general base for proton abstraction and the data for DEPC inactivation as a function of pH suggests that a histidine residue is involved with a pK_a of 6.95. There may also be an arginine residue at the active site of the type II DHQase which could be involved in binding to the carboxyl group of the substrate, dehydroquinate.

CHAPTER V

Characterisation of *Pisum sativum* bifunctional dehydroquinase-shikimate dehydrogenase

5.1 Introduction

One of the aims of this project was to characterize the pea (*P. sativum*) DHQase, which occurs as a bifunctional polypeptide with SDHase (Mousdale *et al.*, 1987). To this end we decided to purify microgram quantities of this protein to obtain primary sequence data, to allow the basic properties of the enzyme to be determined and to establish whether the plant enzyme is a type I DHQase. Degenerate oligonucleotides could then be designed for use in the PCR to amplify the cDNA encoding DHQase. This cDNA could then be cloned and sequenced to further characterise the plant enzyme.

Pea DHQase has been previously purified by Mousdale *et al.* (1987), but nothing was known about its primary structure. This was due to the failure to obtain sequence data on the very small amounts of protein originally purified. Methods for obtaining protein sequence on microgram quantities of protein have improved substantially since 1987, in particular, the use of electrophoretic transfer (electroblotting) onto special filters has made it practical to obtain substantial amounts of sequence on a few micrograms of protein (Matsudaria, 1987).

This chapter describes the large-scale purification procedure and micro-sequencing studies on the DHQase-SDHase from *P. sativum*. It also describes some of the properties of plant DHQase.

5.2 Plant material

Pea seeds (*Pisum sativum* L. Var. "Onward") were planted in moistened vermiculite and grown at 28°C under 15h of illumination for 15 day prior to harvest. Plants were cut off at ground level and stored at -20°C.

5.3 Large-scale purification method

In the protocol, scaled-up and modified from that described previously (Mousdale *et al.*, 1987), a relatively large amount of 15 day old pea shoots were used.

Unless otherwise stated, all steps were carried out at 4°C. The buffers used in this procedure are listed in Table 5.1.

Table 5.1

Buffers used for DHQase-SDHase purification

Buffer A: 50mM-Tris/HCl, pH 7.5, 1mM-EDTA, 1mM-BE, 2mM-DTT, 1mM-PMSF

Buffer B: 20mM-Tris/HCl, pH 7.5, 2mM-DTT

Buffer C: 20mM-Tris/HCl, pH 7.5, 1mM-BE, 1mM-DTT, 1mM-PMSF

Buffer D: 20mM-Tris/HCl, pH 7.5, 1mM-BE, 1mM-DTT

Buffer E: 10mM-potassium phosphate, pH 6.8, 1mM-BE, 1mM-DTT

Buffer F: 20mM-Tris/HCl, pH 7.5, 1mM-BE, 0.5mM-DTT

Step 1: Extraction and clarification

1250g of frozen pea shoot tissues were homogenised in 1700 ml homogenisation Buffer-A, in a Waring blender (packing intermittently with a spatula to reduce the volume of leaves) for 2-3min in 30sec medium speed bursts.

The homogenate was filtered by squeezing through two layers of cheese-cloth and clarified by centrifugation for 25min at 5500 rpm in an *MSE* 6L rotor.

The supernatant (crude-extract =2100ml) was adjusted to a conductivity equivalent of 0.2M KCl (~15m mho) for further purification.

Step 2: Passage through DE-52 column

The crude-extract was then loaded onto a 1000ml DE-52 column (at ~600ml/h), equilibrated beforehand with 4 litres of 0.2M-KCl in Buffer-B. 150ml fractions were collected. The column was finally washed with Buffer-B containing 1mM-BE, 1mM-EDTA and 1mM-PMSF. Fractions containing SDHase activity were pooled (2300ml).

Step 3: Fractionation with $(NH_4)_2SO_4$

The combined fractions from the DE-52 column were brought to 35% saturation with $(NH_4)_2SO_4$ (197g/l). After 20 min of gentle stirring the precipitate was removed by centrifugation at 5500 rpm in an *MSE* 6L rotor for 25 min.

The supernatant (2420ml) was then adjusted to 55% saturation with $(NH_4)_2SO_4$

(120g/l) and stirred for 20 min. The precipitate was collected by centrifugation as above for 30 min, redissolved in 72ml of Buffer-C and dialysed for 48h against 3X2000ml of Buffer-C. The final dialysis was in Buffer-D and the volume recovered after dialysis was 100ml.

Step 4: DE-52 chromatography

Dialysed protein was loaded at 50ml/h onto a DE-52 column (350ml), pre-equilibrated with Buffer-D. The column was then washed with ~500ml of Buffer-D and eluted with a linear gradient (1000ml) of 0-250mM-KCl in Buffer-D. Fractions containing SDHase activity were pooled and dialysed against 2000ml of Buffer-D. The elution profile can be seen in Fig. 5.1.

Step 5: ADP-Sepharose chromatography

Dialysed protein was loaded at 60ml/h onto an ADP-Sepharose column (~40ml), pre-equilibrated with Buffer-D. The column was then washed with five column volumes of the same buffer; at this stage the A_{280} dropped to ~0.35. Finally, the column was eluted (at 30ml/h) with two column volumes of the same buffer containing 1mM-shikimic acid and 1mM-$NADP^+$. Fractions active in SDHase were pooled and dialysed overnight against 2000ml of Buffer-E. The elution profile can be seen in Fig. 5.2.

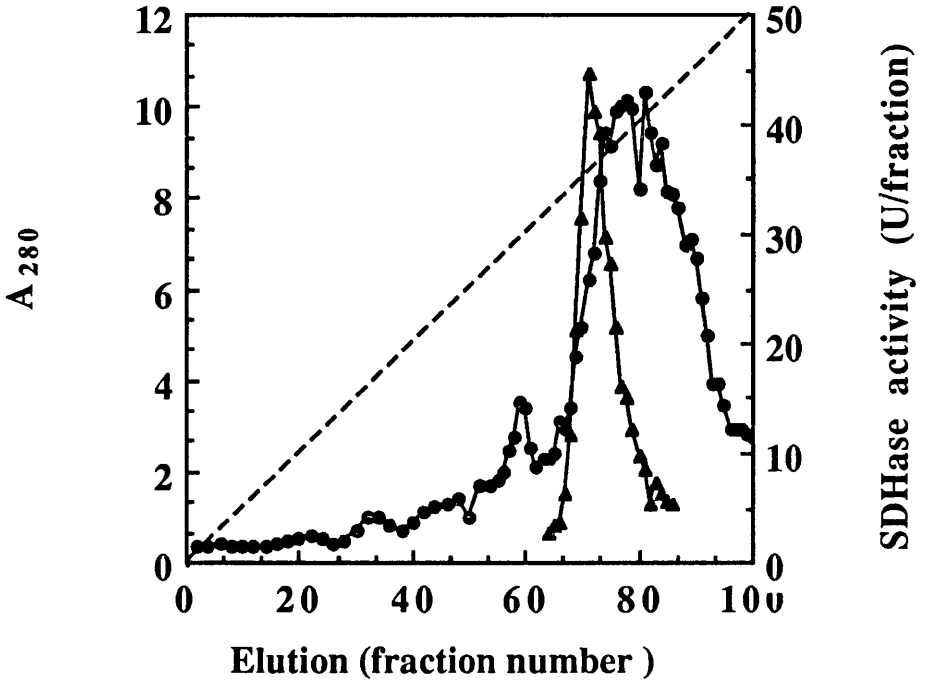


Fig. 5.1 DE-52 chromatography of DHQase-SDHase

The protein elution profile (*filled circles*) was measured by absorbance at 280nm (A_{280}). The salt gradient (0 to 250mM-KCl) is indicated by the *dashed line*. SDHase activity (*filled triangles*) was assayed as described in Section 2.4.2. Fraction numbers are indicated.

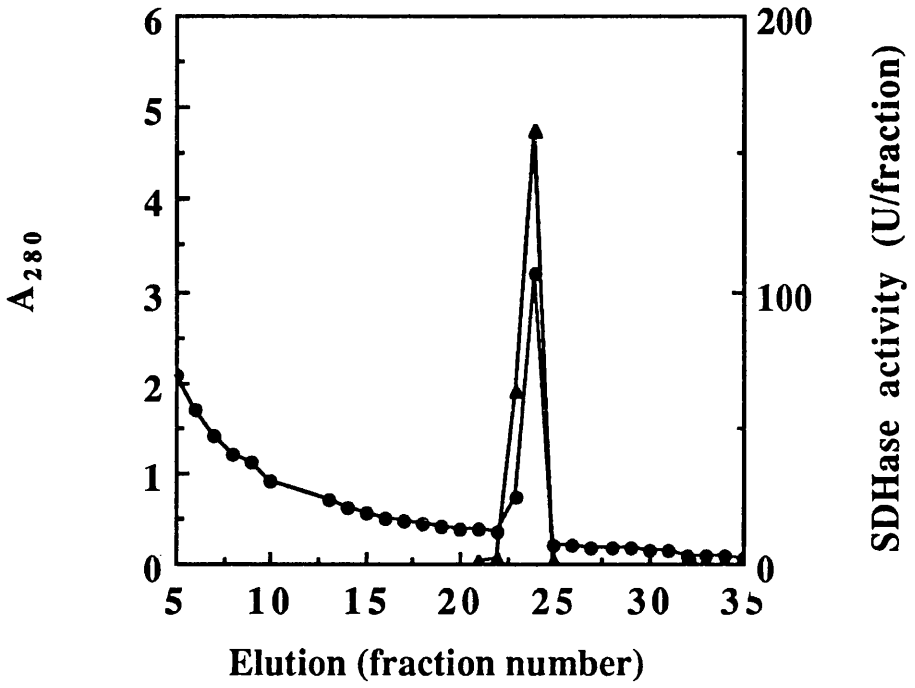


Fig. 5.2 ADP sepharose chromatography of DHQase-SDHase

The protein elution profile (*filled circles*) was determined by absorbance at 280nm (A_{280}). SDHase activity (*filled triangles*) was assayed as described in Section 2.4.2. Fraction numbers are indicated.

Step 6: Hydroxylapatite chromatography

Protein dialysed overnight was loaded (at 24ml/h) onto a hydroxylapatite column (~50ml) pre-equilibrated with Buffer-E. The column was then washed with two bed volumes of the same buffer (at 24ml/h) and, at that point the A_{280} dropped to 0.05. Finally, the column was eluted with a linear gradient (500ml) of an increasing phosphate concentration (10mM to 250mM-potassium phosphate, pH 6.8, containing 1mM-BE and 1mM-DTT). The elution profile can be seen in Fig. 5.3. Active fractions were dialysed against Buffer-F.

Step 7: Mono-Q anion-exchange chromatography

This step was carried out at room temperature using a *Pharmacia* FPLC system. The dialysate was applied to an HR5/5 Mono-Q anion-exchange column and was eluted with a 20ml linear gradient of 0-200mM-NaCl in Buffer-F (flow rate 1ml/min; 0.5ml fractions). The fractions with the highest activity were pooled and analysed by SDS PAGE (Fig.5.4). At this stage, two well-separated bands were visible. One corresponded to the bifunctional polypeptide (DHQase-SDHase) and the other, of low molecular weight was an unknown protein.

As shown in Fig. 5.5, the Mono Q chromatogram contained three activity peaks - one major peak comprising ~80% of the total activity, whereas the two minor peaks combined ~20% of the activity. These peaks were examined by SDS PAGE and in all cases a band corresponding to a bifunctional enzyme was observed (Fig. 5.6). The major peak with highest activity was collected and completely purified by SDS PAGE (see Section 5.6.2) for microsequence analysis.

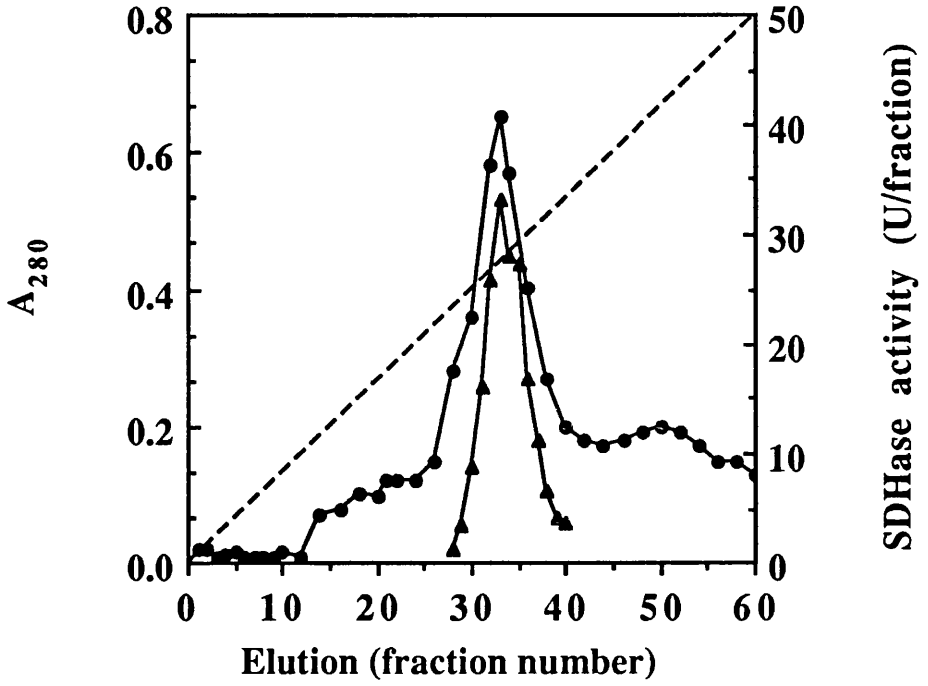


Fig. 5.3 Hydroxylapatite chromatography of DHQase-SDHase

The protein elution profile (*filled circles*) was determined by absorbance at 280nm (A_{280}). The phosphate gradient (10mM to 250mM) is indicated by the *dashed line*. SDHase activity (*filled triangles*) was assayed as described in Section 2.4.2. Fraction numbers are indicated.

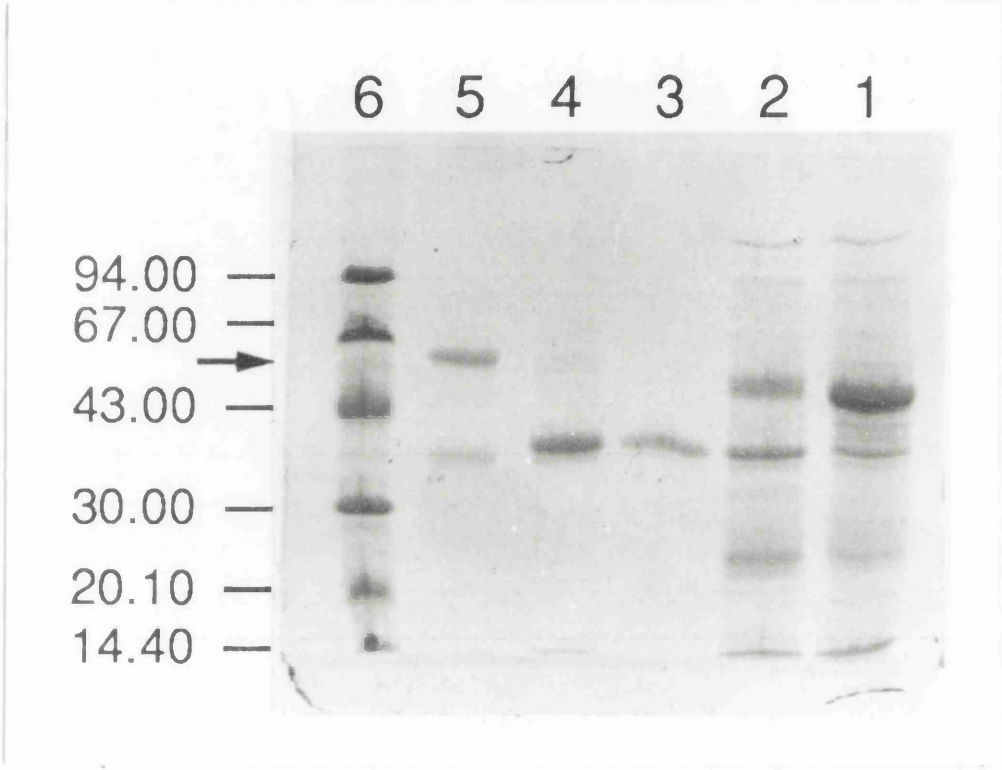


Fig. 5.4 SDS PAGE analysis of DHQase-SDHase

Samples from each step of the purification (indicated as 1, 2, 3, 4, and 5) were analysed on a 12.5% Phast Gel as described in Section 2.3. The DHQase-SDHase band is marked by an arrowhead (E2/E3). *Lanes* 1-5, $(NH_4)_2SO_4$ cut, DE-52 pool, ADP Sepharose pool, Hydroxylapatite pool and Mono Q pool respectively. *Lane* 6, molecular masses indicated in kilodaltons at the left of the panel.

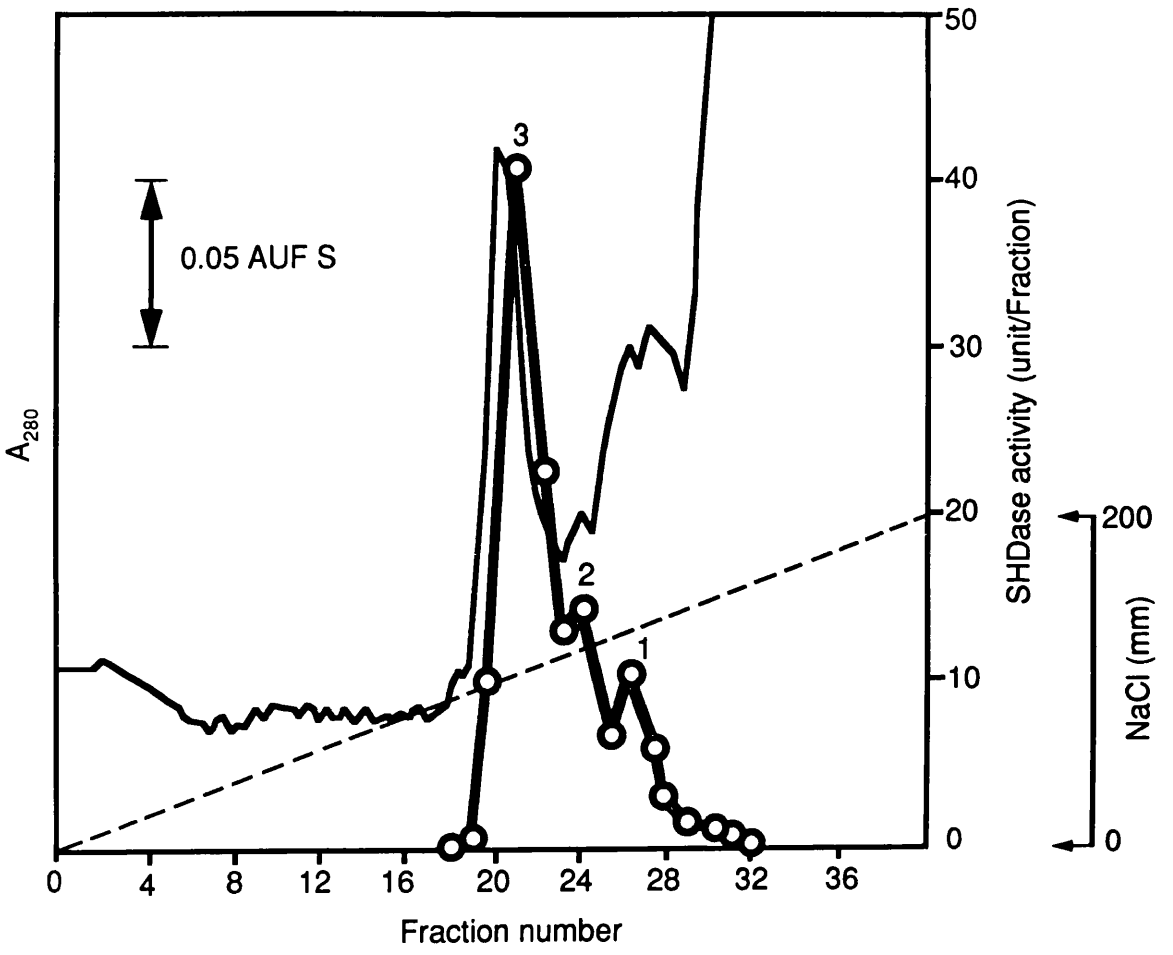


Fig. 5.5 Mono Q chromatography of DHQase-SDHase

The protein elution profile (*full line*) was monitored by absorbance at 280nm (A_{280}). The salt gradient is indicated by the *dashed line*. SDHase activity (*opened circles*) was assayed as described in Section 2.4.2. Three activity peaks are indicated as 1, 2 and 3. Fraction numbers are indicated. *AUFS*, absorbance units full scale.

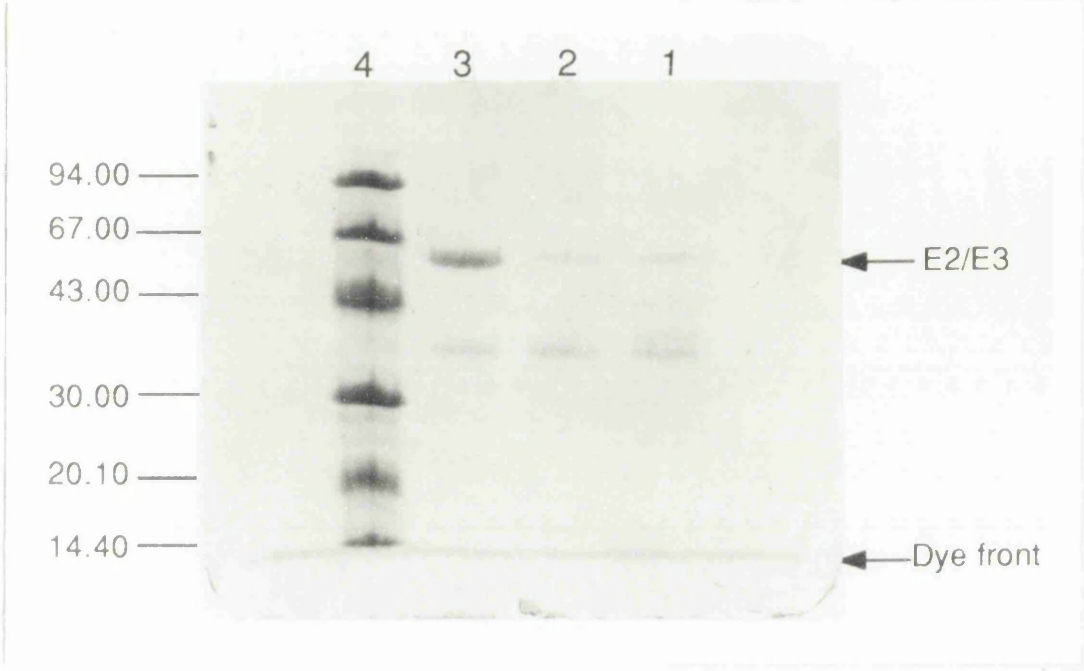


Fig. 5.6 SDS PAGE analyses of Mono Q purified activity peaks of DHQase-SDHase

Samples from each peak (Fig. 5.4) were analysed on 12.5% SDS Phast Gel as described in Section 2.3. *Lanes* 1, 2, and 3 corresponed to activity peaks 1, 2 and 3 respectively in Fig. 5.4. DHQase-SDHase (E2/E3) marked by an *arrowhead*. *Lane* 4, ~1μg of each protein used as a molecular weight marker (molecular masses indicated in kilodaltons at the left of the panel).

5.4 Molecular weight determination

DHQase-SDHase is a monomeric protein (Mousdale *et al.*, 1987). Its molecular weight is 57,500, as determined by SDS PAGE (12.5%) using *Pharmacia* low molecular weight markers (Fig.5.7). This value was slightly lower than the value reported by Mousdale *et al.* (1987). However, it is equivalent to the sum of the molecular weights of the polypeptide chain of the monofunctional *E. coli* DHQase and SDHase.

5.5 Mechanistic properties of the *P. sativum* DHQase

5.5.2 K_m value

The K_m value determined in 100mM-potassium phosphate, pH 7.0 at 25^{0}C was 37 μM, higher than the value for the type I DHQase from *E. coli* (16 μM). Both these values are much lower than that reported for the type II DHQase from *A. nidulans* (150 μM) (Kleanthous *et al.*,1992).

5.5.3 pH activity profile

The pH activity profile for the *P. sativum* DHQase was determined in 100mM-potassium phosphate, over the pH range 5.5 - 7.8 at 25^{0}C. The reaction showed an optimum at pH 7.0 (Fig. 5.8). This value was consistent with that of Mousdale *et al.* (1987), and very similar to the value reported for *E. coli* DHQase by Chaudhuri *et al.* (1986).

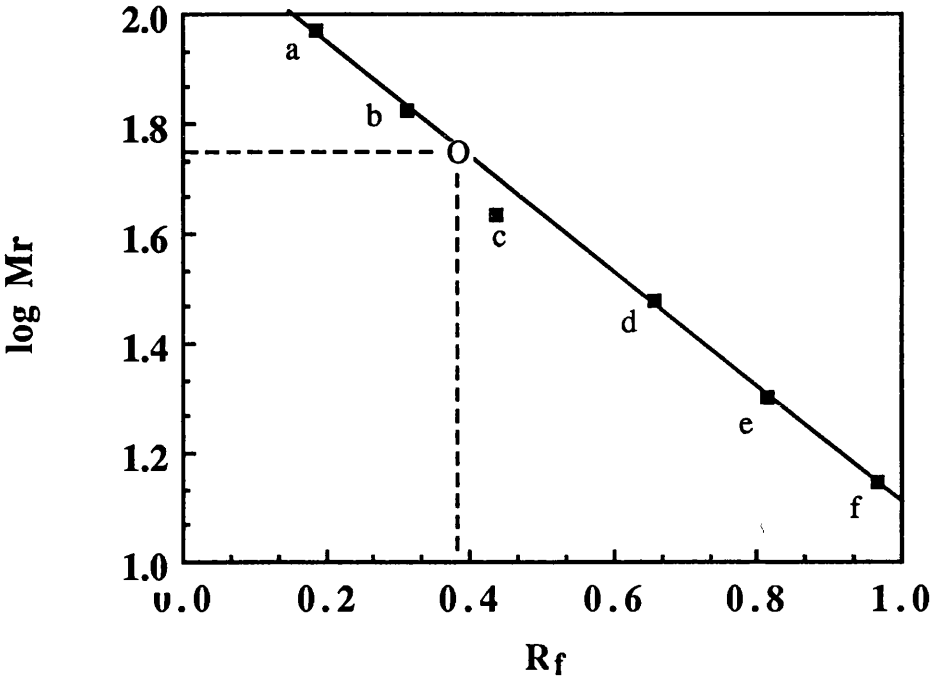


Fig. 5.7 Molecular masses of *P. sativum* DHQase-SDHase

Mono Q purified protein was subjected to SDS PAGE (12.5%) as described in Section 2.3. The R_f values of standard protein (Section 2.11.1) were calculated and plotted against the log Mr. The R_f of the purified DHQase-SDHase (O) is shown. Approximately 1 µg of each LMW as standard protein was applied (a=phosphorylase b, 94,000; b=albumin, 67,000; c=ovalbumin, 43,000; d=carbonic anhydrase, 30,000; e=trypsin inhibitor, 20,100; and f=α-Lactalbumin, 14,400).

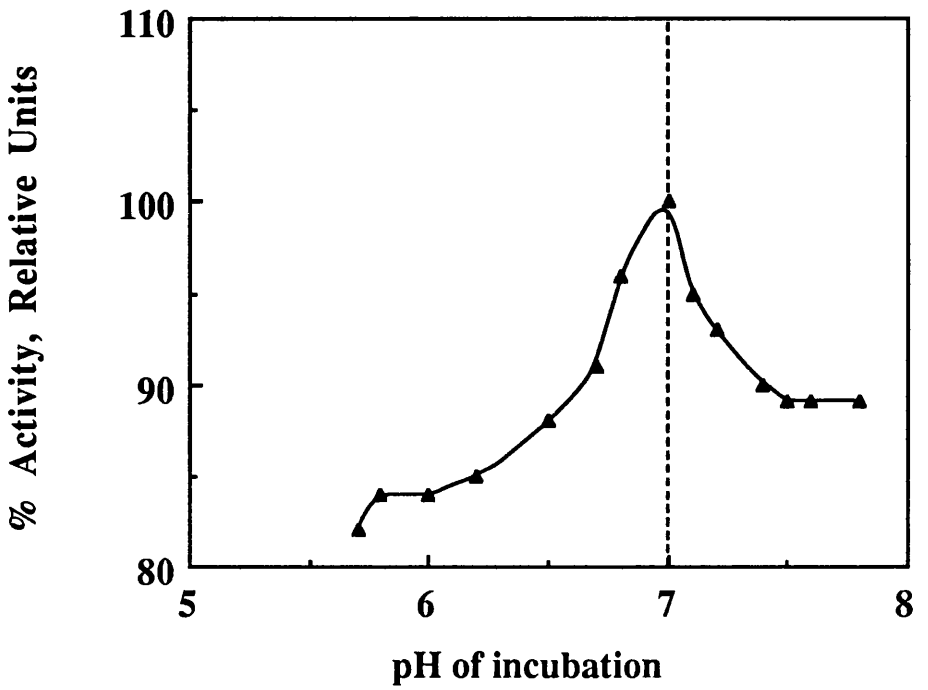


Fig. 5.8 pH optimum of the *P. sativum* DHQase

DHQase was assayed as described in Section 2.4 over a pH range 5.5 - 7.8. The values plotted are the maximal values at each pH. The curve was *theoretical line.*

5.5.4 DEPC inactivation

The DHQase and SDHase activities of the *P. sativum* bifunctional protein are inactivated by DEPC and in both cases inactivation follows pseudo first-order kinetics. Substrates protected both activities against DEPC (Fig.5.9). These results suggest the presence of a histidine residue at or near both active sites.

5.5.5 Borohydride reduction

The *P. sativum* DHQase activity, like that of *E. coli*, was >99% inactivated upon treatment with substrate and borohydride (Table 5.2 and Fig. 5.10), while control experiments showed no loss of activity. The SDHase activity was unaltered by this treatment, indicating that the activities are catalytically independent. The finding that borohydride inactivates in the presence of substrate indicates the formation of a Schiff base intermediate between the enzyme and its substrate.

Table 5.2

Inactivation of *P. sativum* DHQase-SDHase with ammonium dehydroquinate and sodium borohydride

Time of incubation with $NaBH_4$	% of initial activities remaining	
	DHQase	SDHase
02 min	1.50%	-
05 min	-	99.00%
10 min	1.00%	-
20 min	-	97.00%
20 h at 4°C	0.30%	87.00%

"-" represents not determined

Controls showed no loss of activity (data not shown)

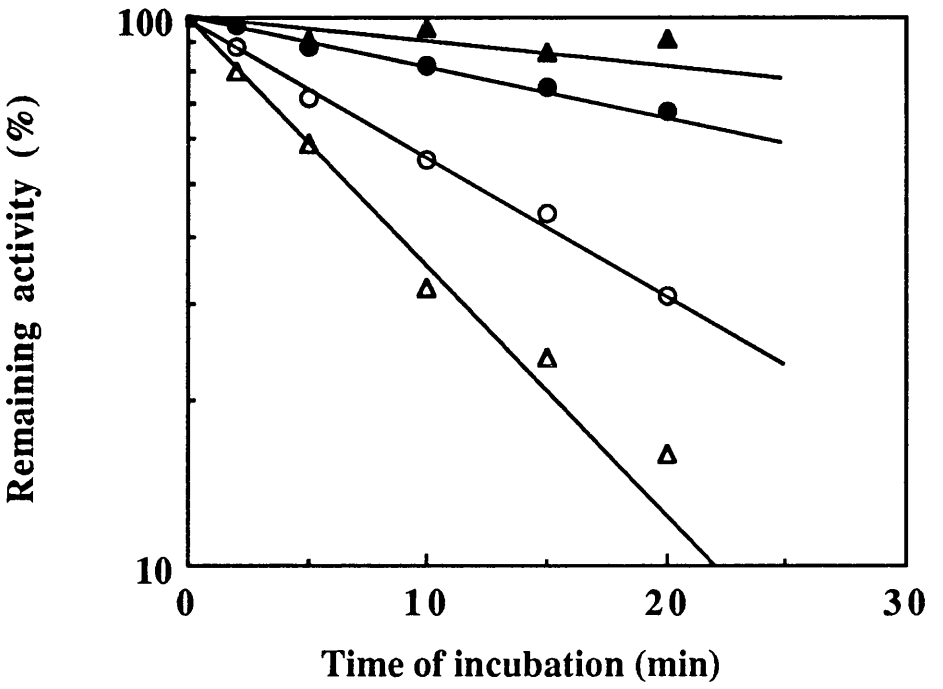


Fig. 5.9 Pseudo first-order plots for DEPC inactivation of DHQase-SDHase

The concentration of DEPC was 1mM (Δ-DHQase; O-SDHase). ▲ and ● show the results of substrate protection against 1mM DEPC inactivation of DHQase-SDHase by 500mM-ammonium dehydroquinate and 2mM-NADP/4mM-shikimic acid respectively.

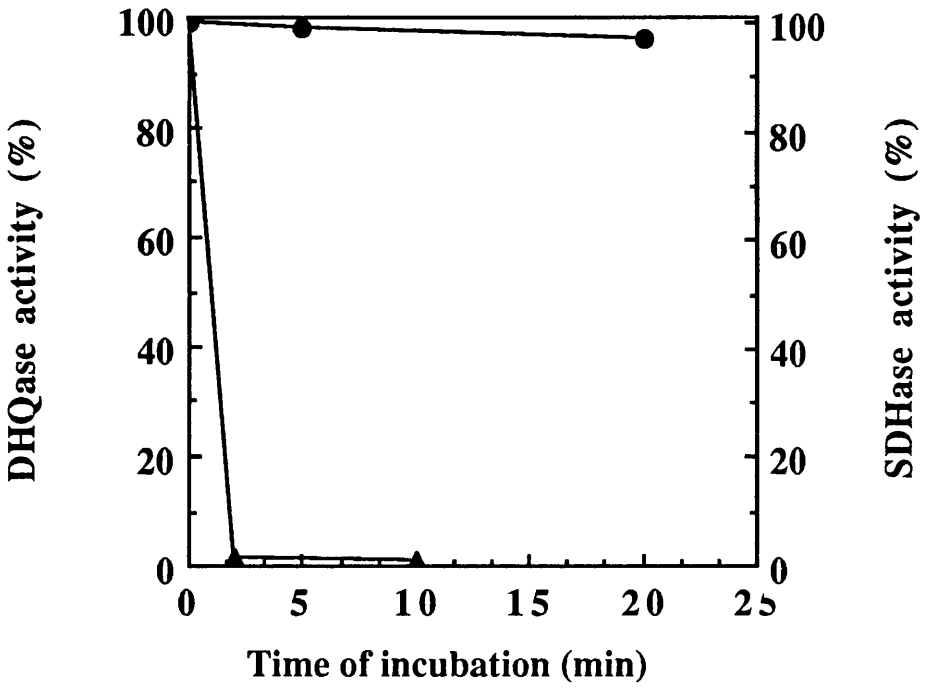


Fig. 5.10 Sodium borohydride inactivation experiment

After incubation of the enzyme with substrate for 5min at 20°C, sodium borohydride was added to the reaction mixture (see Section 2.11.3). DHQase activity (▲) rapidly decreased whereas SDHase activity was unaltered (●).

5.6 Microsequencing of DHQase-SDHase

Since its introduction in the late 1980s, the electroblotting technique for transferring proteins from SDS polyacrylamide gels onto PVDF membrane (Immobilon-P, Millipore) has become a routine procedure for microsequence analysis of polypeptides (Matsudaria, 1987). This technique can, of course, be carried out with partially purified protein provided that the desired protein band can be identified correctly. Moreover, gel purified protein can be subjected to *in situ* limited proteolysis and multiple peptides can be separated by SDS PAGE (Cleveland *et al.*, 1977) prior to blotting the bands onto a PVDF membrane for microsequence analysis. The following sections describe the sequence analysis of microgram quantities of the *P. sativum* bifunctional enzyme using this novel method. Primary sequence comparisons were then possible with the corresponding monofunctional *E. coli* enzyme.

5.6.1 N-terminal sequence of DHQase-SDHase

Partially purified DHQase-SDHase was separated on a 10% SDS polyacrylamide gel and electophoretically transferred onto PVDF membrane as described in "Materials and Methods" (Section 2.12). The band corresponding to the bifunctional polypeptide was cut out and the N-terminal amino acid sequence determined (Glasgow University) by automatic Edman degradation as described in "Materials and Methods" (Sections 2.12.4). The protein was not blocked at its N-terminus and methionine was found to be the N-terminal residue of the bifunctional polypeptide. The amount of sequence from each cycle of sequencing and the derived sequence are shown in Table 5.3. Although there was a sufficient amount of protein to sequence at least the first 30 residues, we were unable to manage more than 12 residues. This was due to the very poor yield at cycle 12. No assignment could be made for residue 8.

A second N-terminal sequencing run was carried out on a separate batch of protein at Aberdeen University. This produced a 26 amino acid sequence confirming the earlier sequence (Table 5.4). Residues 1 to 7 were identified unambiguously, residue 8 could not be identified initially but was later identified as a cysteine derivative. The sequence was then unambiguous up to residue 12. The later ambiguities (indicated in Table 5.4) were almost certainly due to the small amount of protein sequencing. A later sequencing run on a peptide obtained from a V8 protease digest of DHQase-SDHase unambiguously confirmed the first 12 amino acid residues from the N-terminus (see Section 5.6.3 and 5.6.5). This third run confirmed a further 9 residues and allowed correction of the 5 residues that were not convincingly identified on the second run.

Table 5.3

The N-terminal amino acid sequence of *P. sativum* DHQase-SDHase (First sequence run at Glasgow University)

Cycle number	Residue identification	Amount sequencing (pmol)
1	Met	69.27
2	Lys	26.14
3	Asn	22.93
4	Ala	83.86
5	Thr	13.55
6	Leu	74.79
7	Ile	67.19
8	*	-
9	Val	84.14
10	Pro	35.45
11	Ile	56.43
12	Met	9.11

* residue not identified

Table 5.4

The N-terminal amino acid sequence of *P. sativum* DHQase-SDHase (Second sequence run at Aberdeen University)

Cycle number	Residue identification	Amount sequencing (pmol)
1	Met	5.9
2	Lys	1.9
3	Asn	6.9
4	Ala	6.3
5	Thr	å
6	Leu	5.2
7	Ile	5.1
8	**	-
9	Val	6.3
10	Pro	3
11	Ile	2.3
12	Met	0.9
13	Val (?)	3.0
14	Glu	1.5
15	Phe/Pro (?)	1.6/0.7
16	Val	2.2
17	Glu	1.0
18	Gly (?)	0.8
19	Met	1
20	Val	2.1
21	*	-
22	Asp	1.3
23	Ile	1.4
24	Gln	2
25	*	-
26	Ala	1.1

* residue not identified

**later experiments in Aberdeen suggested that the peak on the h.p.l.c. traces was due to a cysteine derivative

å very low yield

? assignment uncertain

5.6.2 Generation of peptides from *in-situ* V8 proteolysis of intact protein

Protease V8 (also called endoproteinase Glu-C) is essentially specific for peptide bonds on the C-terminal side of glutamate residues, with some cleavage also occurring at aspartyl bonds (Drapeau, 1976). This proteinase is very frequently used for *in situ* proteolysis. Partially purified protein was completely purified in the presence of SDS on a 10% polyacrylamide gel (Laemmli, 1970). The desired protein bands were cut out and subjected to limited proteolysis as described in "Materials and Methods" (Section 2.12.3).

5.6.3 Sequencing of peptides

Peptides generated from *in-situ* proteolysis were transferred onto PVDF membrane and visualized by staining with Coomassie Blue-R250 (Fig. 5.11). Five peptides were excised and analysed directly on an Applied Biosystems 477A protein sequencer fitted with a reaction cartridge specially designed for PVDF bound samples (Aberdeen University). The derived sequences are shown in Table 5.5. I, II, III, IV and V.

5.6.4 Amino acid composition of the DHQase-SDHase

Amino acid analysis was carried out as described in "Materials and Methods" (Section 2.13). Table 5.6 shows the amino acid composition of DHQase-SDHase. No attempt was made to derivatise cysteine and methionine before hydrolysis and so these residues like tryptophan were destroyed by the hydrolysis conditions.

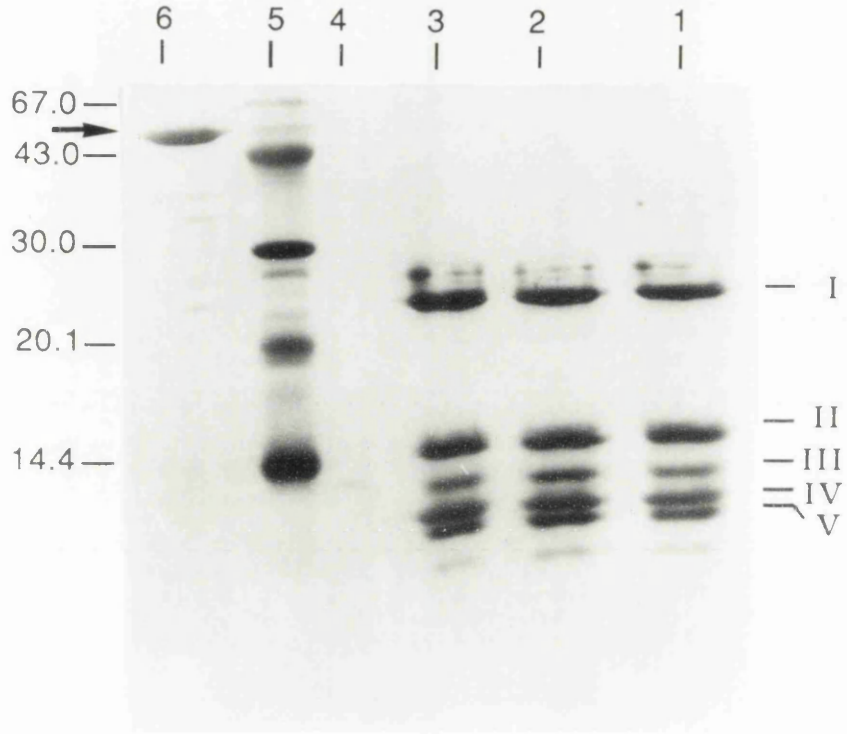


Fig. 5.11 *In situ* V8 protease digestion of DHQase-SDHase (Photograph of PVDF-bound protein/peptides)

Gel slices containing 25-30µg of DHQase-SDHase were digested with 2µg of V8-protease. Peptides were separated on 18% SDS PAG, transferred electrophoretically onto PVDF membrane and Coomassie stained on the blot as described in Section 2.12. *Lanes* 1-3, V8-digest; *Lane* 4, 2µg V8 as control; *Lane* 5, 10µg of each protein used as molecular weight markers (molecular masses indicated in kilodaltons at the left of the panel). *Lane* 6, approx. 10µg of intact protein (*arrowed*). The V8 fragments are referred as peptide-I, II, III, IV and V according to their mobility.

Table 5.5.I

The amino acid sequence of V8 peptide I

Cycle number	Residue identification	Amount sequencing (pmol)
1	Met	92
2	Lys	87
3	Asn	102
4	Ala	94
5	Thr	48
6	Leu	88
7	Ile	96
8	*	-
9	Val	76
10	Pro	48
11	Ile	49
12	Met	36
13	Gly	37
14	Glu	27
15	Thr	14
16	Val	35
17	Glu	19
18	Lys	25
19	Met	25
20	Val	26
21	Val	30
22	Asp	26
23	Ile	28
24	Gln	27
25	Lys	20
26	Ala	23
27	Lys	26
28	Leu	23
29	Asn	20
30	Gly	19
31	Ala	18
32	Asp	17
33	Leu	15
34	Val	13
35	Glu	11
36	Ile	15
37	Arg	4.7
38	Leu	10
39	Asp	7
40	Asp	8

* peak not identified
Peptide I later identified as the N-terminal peptide

Table 5.5.II

The amino acid sequence of V8 peptide II

Cycle number	Residue identification	Amount sequencing (pmol)
1	Gly	35
2	Gly	30
3	Lys	21
4	Tyr	23
5	Asp	26
6	Gly	20
7	Asp	22
8	Glu	16
9	Asn	15
10	Arg	15
11	Arg	15
12	Leu	16
13	Asp	17
14	Ala	14
15	Leu	13.5
16	Arg	9.5
17	Leu	15
18	Ala	13
19	Val	10.8
20	Glu	7.2
21	Leu	10.5
22	Gly	9.7
23	Ala	10.8
24	Asp	8.7
25	Tyr	6.8
26	Val	7.5
27	Asp	8.6
28	Ile	7.1
29	Glu	2.3
30	Leu	6.1
31	Lys	4.6
32	Val	5
33	Ala	5.7
34	His	2.3
35	Glu	2.6
36	Phe	4
37	Tyr	3.2
38	Asp	3.5
39	Ser	0.6
40	Ile	1.8

Table 5.5.III

The amino acid sequence of V8 peptide III

Cycle number	Residue identification	Amount sequencing (pmol)
1	Ser	10
2	Ala	25
3	Leu	25
4	Lys	20
5	**	-
6	**	-
7	Asp	22
8	Glu	15
9	Val	15
10	Asp	18
11	Pro	10
12	Val	10
14	Lys	7
15	Ser	2
16	Ile	4
17	Gly	4
18	Ala	4
19	Val	4
20	Asn	3.2
21	*	-
22	Ile	2.5
23	Val	2.7
24	Arg	1
25	*	-
26	Pro	1.4
27	Thr	0.8
28	Asp	1
29	Gly	1.3
30	Ala/Lys/Leu	0.3/0.4/0.7
31	Lys	1
32	Leu	1.2
33	*	-
34	Tyr	0.6

* residue assignment not possible

**later experiments in Aberdeen suggested that the peaks on the HPLC traces were due to cysteine derivatives

Where more than one amino acid was seen in the PTH profiles the first amino acid given is the one later identified by nucleotide sequencing.

Table 5.5.IV

The amino acid sequence of V8 peptide IV

Cycle number	Residue identification	Amount sequencing (pmol)
1	Ser	16
2	Gly	47
3	Val	50
4	Val	52
5	Ser	11
6	Ala	37
7	Pro	34
8	Gly	27
9	Gln	28
10	Pro	11
11	Thr/pro	4/14
12	Ile/Asn	8.3/2
13	Lys/Ile	4.8/3
14	Asp/Lys	7.2/8.3
15	Leu/Asp	7.4/11
16	Leu	15
17	His/Leu	6/14
18	Leu/His	15/3
19	Tyr	4
20	Asn	4.6
21	Phe	4.5
22	Arg	2.8
23	Gln	4.2

Where more than one amino acid was seen in the PTH profiles the first amino acid given is the one later identified by nucleotide sequencing.

Table 5.5.V

The amino acid sequence of V8 peptide V

Cycle number	Residue identification	Amount sequencing (pmol)
1	Asp	48
2	Gly	48
3	Met	29
4	Ile	40
5	Leu	35
6	Ala	39
7	Asn	34
8	Thr	22
9	Thr	22
10	Ser	9
11	Ile	23
12	Gly	19
13	Met	14
14	Gln	20
15	Pro	12
16	Lys	9
17	Val	9
18	Asp	8
19	Glu	5
20	Thr	3.5

Table 5.6

Amino acid analysis of DHQase-SHDase

Amino acid residue	pmol by height	%composition
Aspartic acid	125.51	12.29
Glutamic acid	90.17	8.83
Serine	74.24	7.27
Glycine	107.96	10.57
Histidine	24.45	2.40
Arginine	46.78	4.58
Threonine	60.08	5.88
Alanine	80.29	7.86
Proline	43.15	4.23
Tyrosine	34.03	3.33
Valine	78.74	7.71
Isoleucine	66.71	6.53
Leucine	83.15	8.14
Phenylalanine	41.19	4.03
Lysine	64.57	6.32

The conditions employed for acid hydrolysis destroy tryptophan; the sample was not pretreated with performic acid to allow identification of cysteine and methionine residues and the analyses of these amino acids are excluded.

% compositions were calculated assigning tryptophan, methionine and cysteine as 0%.

5.6.5 Alignment of amino acid sequences with the monofunctional *E. coli* DHQase and SDHase

A total of 157 residues were sequenced from five different peptides (Fig. 5.11) and compared with the corresponding monofunctional *E. coli* enzymes using the programme (*BESTFIT*) as described in "Materials and Methods" (Section 2.24.1). Three of these peptides were homologous to regions in DHQase and two of them to SDHase (Fig. 5.12). Peptide I contains the N-terminal sequence found for the intact protein and so is clearly the N-terminal region of the protein. It is homologous to the N-terminal sequence of *E. coli* DHQase which suggests that the DHQase domain of the *P. sativum* bifunctional enzyme is N-terminal (Fig. 5.13). Detailed discussion about the analysed sequences is presented in Chapter VI.

Peptide I

```
P. sativum    2 KNATLICVPIMGETVEKMVVDIQKAKLNGADLVEIRLDD 40
                ..|. |.|.:|:..:..:  :    : .: |::| |:|.
E. coli      13 TGAPKIIVSLMAKDIASVKSEALAYREADFDILEWRVDH 51
```

Peptide II

```
P. sativum    1 GGKYDGDENRRLDAL....RLAVELG.ADYVDIELKVAHE 35
                ::| :|::. . :|.    | |:: | .| :|:|| .:.:
E. coli      83 SAKEGGEQAISTEAYIALNRAAIDSGLVDMIDLELFTGDD 122
```

Peptide III

```
P. sativum    1 SALKCCDEVDPVAKSIGAVNXIVRXP........TDGAKL 32
                .|:  .||:.. |   |||| ::| .         |||. |
E. coli      67 EAFARADELTERAALAGAVNTLMRLEDGRLLGDNTDGVGL 106
```

Peptide IV

```
P. sativum    1 SGVVSAPGQPTNIDDLLHYNF 21
                .:  ||||| .  |    ..:
E. coli     230 KA..SAPGQISVNDLRTVLTI 248
```

Peptide V

```
P. sativum    1 DGMILA.NTTSIGMQPKVD 18
                ||.:|: ||.::|: ..::
E. coli      93 DGRLLGDNTDGVGLLSDLE 111
```

Fig. 5.12 Sequence alignments of the *P. sativum* DHQase-SDHase peptide sequences with the corresponding regions in monofunctional *E. coli* DHQase-SDHase

Identical amino acids are indicated by *vertical bars*, and *dotted lines* represent conserved changes. *Periods* indicate gaps for optimal alignment. Peptides I, II and IV were homologous to regions in DHQase and other two (III and V) to SDHase.

```
                  1                                                  * * *  60
E. coli                MKTVT VKDL.VIG.T GAPKIIVSLM AKDIASVKSE ALAYREADFD ILEWRVDH
S. typhi               MKTVT VKNL.IIG.E GMPKIIVSLM GRDINSVKAE ALAYREATFD ILEWRVDH
A. nidulans       .L..QVATGQ IDSLSIIK.E KEHSFFASLT LPDLRE.AGD ILEEVCVGSD AVELRVDL
S. cerevisiae  SFSKYIAT.. ITGVREIEIP SGKRSAFVCLT FDDLTE.QTE NLTPICYGCE AVEVRVDH
P. sativum             MKNAT LICVPIMG.E TVEKMVVDIQ KAKLNG.... ........AD LVEIRLDS
```

Fig. 5.13 Alignment of the *P. sativum* DHQase-SDHase N-terminal sequence with type I DHQases
Identical amino acids are indicated by *asterisks*. *Periods* indicate gaps for optimal alignment.

Table 5.7

Large scale purification of DHQase-SDHase from *P. sativum* shoot extract

Step	Volume (ml)	Protein (mg)	Total activities (units)		SDHase / DHQase	Recovery* (%)	Purification* (fold)
			DHQase	SDHase			
Crude extract	2100	6775	N.D.	406	N.D.	N.D.	N.D.
DE-52 (-)	2300	6095	N.D.	398	N.D.	100	1
$(NH_4)_2SO_4$	100	3925	35	385	11	97	1.5
DE-52 (+)	143	722	33	357	10.82	90	7.6
ADP-Sepharose	22	27.72	21	260	12.38	65	144
Hydroxylapatite	40	14.40	15	206	13.73	52	219
Mono-Q	4.3	0.45	7.74	73	9.44	18	2488

DHQase = Dehydroquinase; SDHase = Shikimate dehydrogenase; *Calculated for SDHase; N.D. = Not determined

5.7 Discussion

In an effort to scale-up the purification of the *P. sativum* bifunctional enzyme, we decided to start a large-scale purification using 1250g of pea shoots. We were able to purify ~400μg of nearly homogeneous protein containing both DHQase and SDHase activities with an activity ratio of about 1:10. An approximately 2500-fold purification was achieved (Table 5.7). Mousdale *et al.* (1987) used 250g of plant material to recover ~80μg of protein after FPLC on a Mono-Q column with 2500-fold purification. This indicates that the large scale purification essentially scaled-up the yield about 5-fold. This enabled us to determine some properties of the plant enzyme as well as to obtain primary sequence data to allow its cloning. It is interesting to note that there is evidence about the occurrence of isozymes of DHQase-SDHase in plants (Fiedler and Schultz, 1985; Mousdale *et al.*, 1987). In the final FPLC Mono-Q ion exchange step of the purification three activity peaks were seen in a ratio of about 8:1:1 (Fig. 5.5). SDS PAGE examination of these peaks showed that they all contained polypeptides with similar moleculr weights (Fig. 5.6). To avoid any ambiguity during primary structure determination we used only the major peak for microsequence analysis.

5.7.1 Properties of *P. sativum* DHQase-SDHase

Purified protein contained both DHQase and SDHase activities at a ratio about 1:10. This protein is quite stable at 4^0C in Buffer-F at least for a month and thereafter a drop in DHQase activity was observed.

It appears that this protein is very hydrophobic in nature. At one stage when we attempted to concentrate the protein by lyophilization, we faced some difficulty in redissolving it. Addition of 0.1% SDS slightly improved the solubility and it was possible to redissolve the protein completely in 50% acetonitrile.

5.7.2 *P. sativum* DHQase is a type I enzyme

Table 5.8 briefly compares some of the mechanistic properties of *P. sativum* DHQase with type I and type II DHQases. The most interesting property is that the plant DHQase activity can be inhibited by borohydride in the presence of substrate which is characteristic of type I DHQases. On this basis the *P. sativum* is a type I enzyme which works through an imine intermediate that can be trapped by reduction. Other properties like the pH optimum and K_m values are also very similar to those of the type I DHQase from *E. coli*.

Table 5.8

Properties of dehydroquinases

Parameters	*P. sativum* DHQase	*E. coli*[a] type I DHQase	*A. nidulans*[b] type II DHQase
K_m (μM)	35	16	150
pH optimum	7.0	6.8[c]	8.5-9.0
DEPC inactivation	+	+	+
Substrate protection	+	+	+
$NaBH_4$ inactivation	+	+[c]	-

+ = inactivation; - = no inactivation

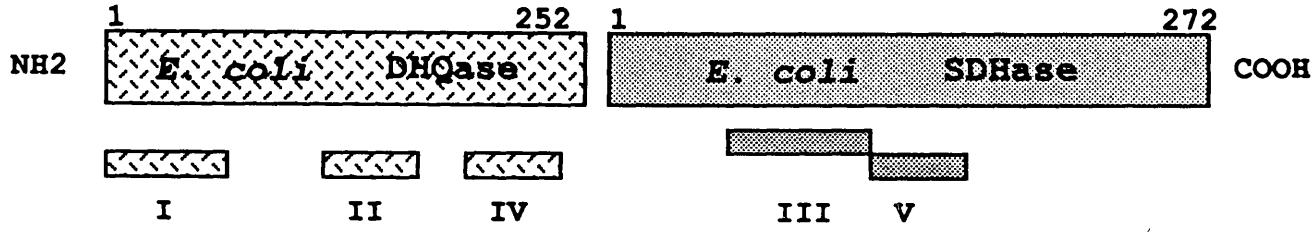
a = Deka *et al.*, 1992; b = Kleanthous *et al.*, 1991; c = Chaudhuri *et al.*, 1986

5.7.3 Domain order of bifunctional DHQase-SDHase

Nothing was known about the domain order of the bifunctional plant DHQase-SDHase. We carried out microsequencing of intact and proteolytically cleaved *P. sativum* protein in order to obtain enough sequence to design oligonucleotides and to undertake cDNA cloning with the aim of determining the complete nucleotide sequence and then deducing the amino acid sequence. The sequence data obtained from the micro-sequence analysis was considerably more than we expected and allowed regions homologous to *E. coli* enzymes to be identified. The peptides I, II and IV corresponded to the *E. coli* DHQase at positions 13 to 51, 83 to 122 and 230 to the C-terminus, respectively. The other two peptides III and V, were homologous to regions in the SDHase of *E. coli* (Fig. 5.14). The first V8-peptide (40 amino acids) which matched with the N-terminus of the intact protein was homologous to the N-terminus of other type I DHQases (Fig. 5.13), establishing that in the *P. sativum* bifunctional enzyme the DHQase domain is N-terminal (5.15).

Further characterisation of the *P. sativum* enzyme was limited by the availability of only small amounts of purified protein after purification. It was therefore decided to progress onto the cloning of the cDNA as no plant cDNA encoding a DHQase had yet been characterised.

Cloning and sequencing of the cDNA would allow the complete protein sequence to be deduced from the nucleotide sequence. The cloned cDNA would also be a valuable tool for later analysis of the gene encoding the plant bifunctional polypeptide and the gene transcripts, and would allow study of the origin of the reported isozymes.



DHQase regions

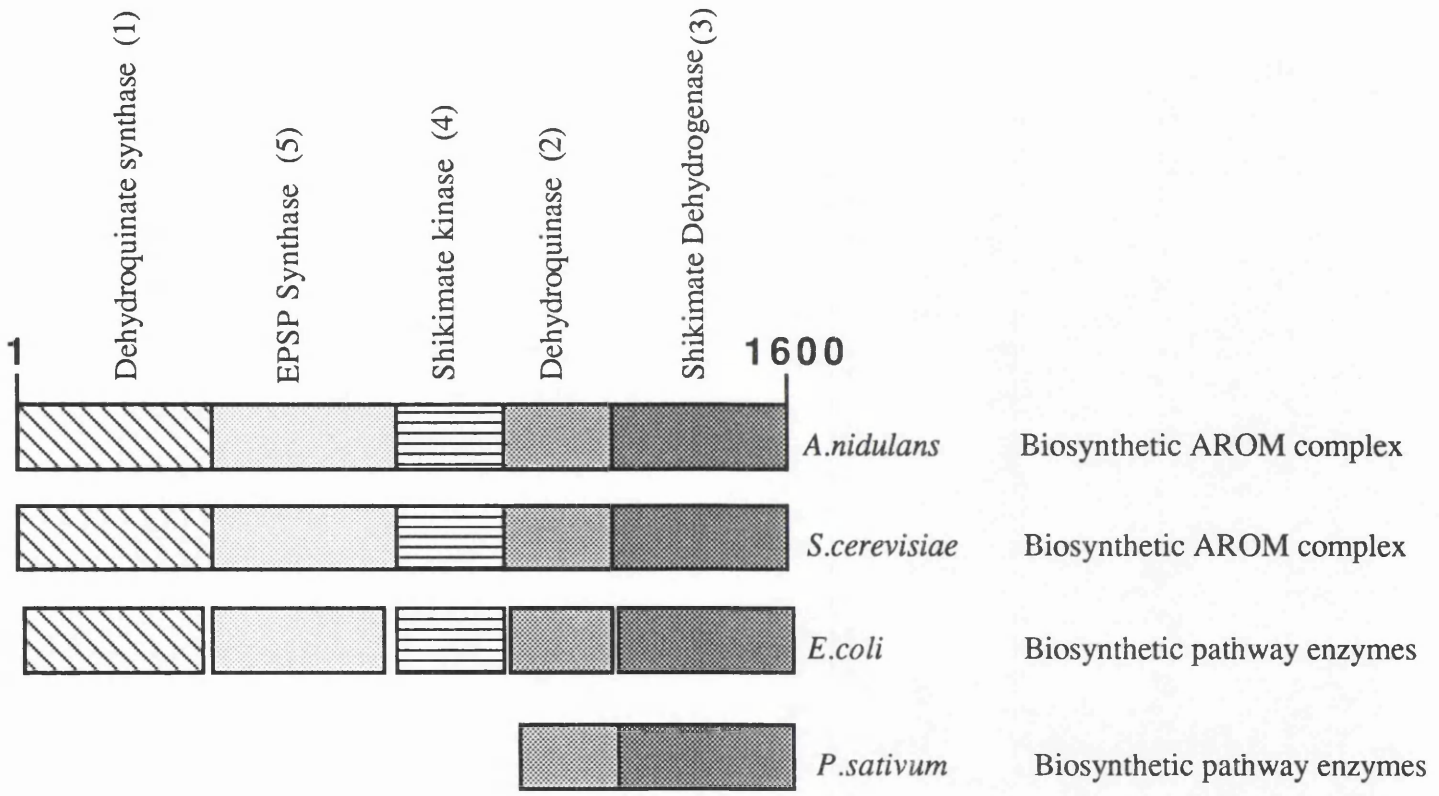
Peptide I = 13-51
peptide II = 83-122
Peptide IV = 230-248

SDHase regions

Peptide III = 67-106
Peptide V = 93-111

Fig. 5.14 Alignment of the *P. sativum* DHQase-SDHase V8-peptides against the monofunctional *E. coli* enzymes (also see Fig. 5.12)

Fig. 5.15 Domain arrangement of the P. sativum DHQase-SDHase with the corresponding shikimate pathway enzymes



CHAPTER VI

Cloning of a cDNA for the dehydroquinase region of *P. sativum* dehydroquinase-shikimate dehydrogenase

6.1 Introduction

One aim of this project was to isolate a cDNA encoding at least the DHQase region of dehydroquinase-shikimate dehydrogenase (DHQase-SDHase) from *P. sativum.* This was to enable us to obtain the amino acid sequence for comparison with other DHQases. Here I describe the first cloning of a cDNA for DHQase (and a part of SDHase) from a higher plant (*P. sativum*). This work used the polymerase chain reaction and confirmed the results from peptide sequencing with regard to the domain order and the identification of the DHQase activity as being of the type I family.

6.2 The polymerase chain reaction

The polymerase chain reaction (PCR) allows the *in vitro* amplification of nucleic acid sequences (Mullis and Faloona, 1987; for a recent review, see Bej *et al.*, 1991). This technique has been simplified and automated by making use of the thermostable *Taq* DNA polymerase, which can withstand the temperatures used in repeated cycles of DNA strand melting, primer annealing, and extension (Saiki *et al.*, 1988). The DNA region to be amplified is commonly defined by a pair of flanking oligonucleotide primers. The template DNA can be genomic DNA, double-stranded cDNA or simply first-strand cDNA. For amplification, "degenerate" primers can be designed from a given amino acid sequence. For example, the urate-oxidase cDNA (Lee *et al.*, 1988), the diabetes-associated peptide gene (Girgis *et al.*, 1988), the mammalian and avian hepadnaviruses (Mack and Sninsky 1988), several genes for ribosomal proteins from *Bacillus stearothermophilus* (Ramakrishnan and Gerchman, 1991), the genes for plant abrin A-chains (Evensen *et al.*, 1991) and the tomato shikimate kinase cDNA (Schmid *et al.*, 1992) have all been cloned using degenerate primers.

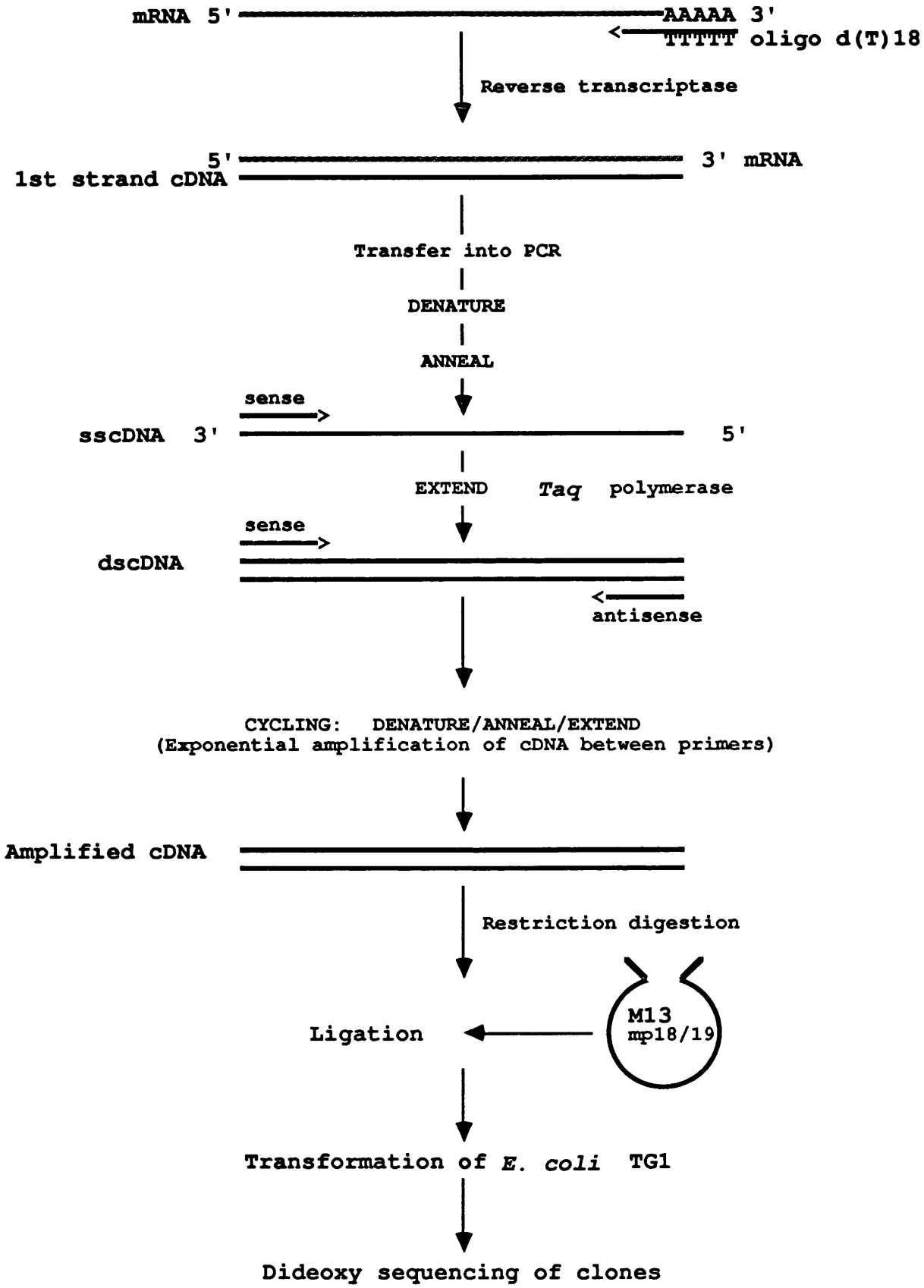


Fig.6.1 The strategy for cDNA amplification, cloning and sequencing

6.3 Design of degenerate PCR oligonucleotide primers

The exact nucleic acid sequence of a segment of DNA cannot usually be deduced from the amino acid sequence, since the genetic code is degenerate, so it is necessary to design primers representing all possible DNA coding sequences for a given peptide sequence. Mixed primers representing all combinations of codon choices have been used for PCR amplification of the target DNA (Patil and Dekker, 1990). In designing a primer with minimal degeneracy one should avoid amino acid residues such as serine, leucine or arginine since these amino acids can be encoded by any of six synonymous codons. Conversely, regions containing methionine and tryptophan are highly advantageous as the amino acids are encoded by a single codon. Deoxyinosine, which forms stable base pairs with all four normal bases, can be used in primers at positions of 3 and 4-fold degeneracy (Knoth *et al.*, 1988; Patil and Dekker, 1990).

Each primer contained two regions: a 3' region complementary to the template sequence, which provides specificity in hybridization of the primer to template, and a 5' tail containing a restriction enzyme site to facilitate subsequent cloning of the PCR products.

6.4 PCR amplification of *P. sativum* DHQase-SDHase cDNA using degenerate primers

6.4.1 The strategy for cDNA amplification cloning

To isolate the DHQase-SDHase cDNAs the polymerase chain reaction was performed using pea cDNA as template and degenerate PCR primers. Fig. 6.1 illustrates schematically the steps involved in the synthesis of cDNA and the subsequent generation, cloning and sequencing of the PCR products. Total RNA [rather than poly $(A)^+$ enriched

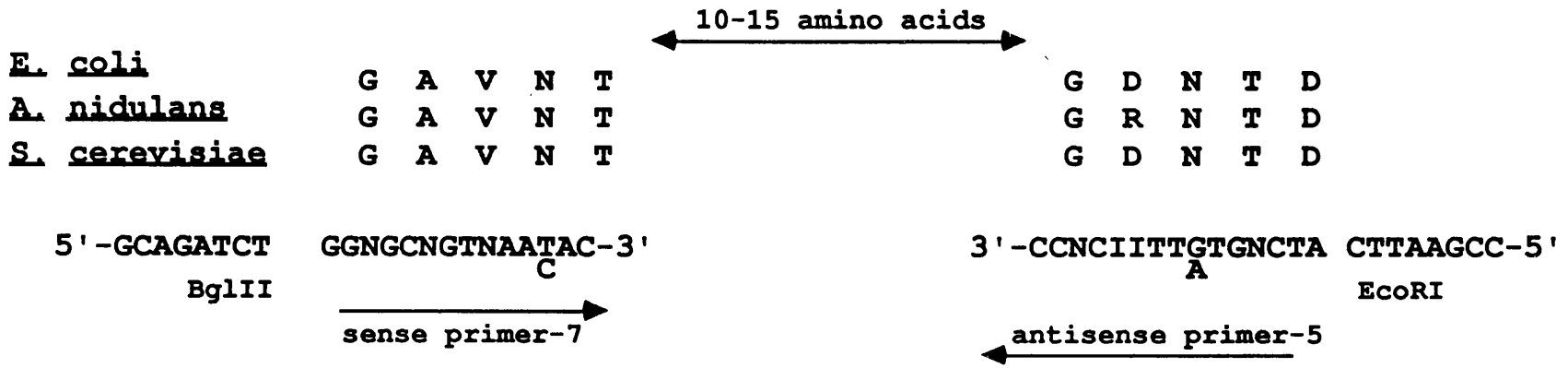


Fig. 6.2 Design of degenerate PCR primers for amplification of a segment of SHDase cDNA -the "first approach"
The amino acid sequences of two short, highly conserved regions of the three SHDases are shown. These regions are 10 residues apart in E. coli and S. cerevisiae but 15 residues apart in A. nidulans. The degenerate PCR primers are shown at the bottom of the figure (see also Table 6.1). N indicates that a mixture of all four bases was used.

RNA] was used for the synthesis of first strand cDNA from the oligo d(T) primer complementary to the poly-$(A)^+$ tails of the mRNAs (Section 2.18.1). Sometimes cDNA primed with random hexanucleotides, $pd(N)_6$, was used.

6.4.2 First approach: primers based on consensus sequences

This approach was tried prior to extensive primary sequence data being available from pea DHQase-SDHase. Genes for DHQase and SDHase have been cloned from several species (see Section 1.3.3 and 1.3.4). This work revealed considerable overall homology between different SDHases and some short highly conserved regions. The initial cloning strategy was based on the hope that these regions were also conserved in higher plants.

Fig. 6.2 shows the two highly conserved regions of SDHase that were used to design a pair of primers (primers-5 and -7 in Table 6.1).

As a positive controls both an *E. coli aro*E clone (encoding SDHase, Anton and Coggins 1988) and *E. coli* genomic DNA were used as templates (The *E. coli* genomic DNA was the generous gift of Dr. G. A. Nimmo) and from these we were able to amplify fragments of the expected size using primers-5 and -7. However, when the same primers were used with *P. sativum*, *A. thaliana* and *Brassica napus* cDNAs, negative results were obtained. These results led me to suspect that the regions which were used to design the primers (Fig. 6.2) might not be well conserved in plants. This turned out to be the case.

Table 6.1

Primers for amplification of pea cDNA in the initial unsuccessful experiments ("first" and "second" approaches)

		BamHI	M K N A T
Primer-1	5'-CGGGATCC	ATGAA(AG)AA(TC)GCNAC-3'	
(Sense/22mer/16fold)			
		ClaI	T N V A G
Primer-2	5'-GCATCGAT	GT(AG)TTNACNGCNCC-3'	
(Antisense/22mer/128fold)			
		ClaI	T N V A G
Primer-3	5'-GCATCGAT	GT(AG)TT(AG)ACNGCNCC-3'	
(Antisense/22mer/64fold)			
		XbaI	V P I M
Primer-4	5'-GCTCTAGA	CAGTNCCNAT(CTA)ATG-3'	
(Sense/22mer/48fold)			
		EcoRI	D T N X G
Primer-5	5'-CCGAATTC	ATCNGT(AG)TTIICNCC-3'	
(Antisense/23mer/32fold)			
		ClaI	
Primer-6	5'-CGATCGAT	AACTGGAAGAATTC-3'	
(antisense/22mer)			
		BglII	G A V N T
Primer-7	5'-GCAGATCT	GGNGCNGTNAA(TC)AC-3'	
(Sense/22mer/128fold)			

N indicates that a mixture of all four bases was used; brackets indicate where a simpler mixture of bases was used. Restriction sites and two extra bases were present at the 5' end of each primer. Primers are shown either in sense or antisense orientation with their length and degenerecies.

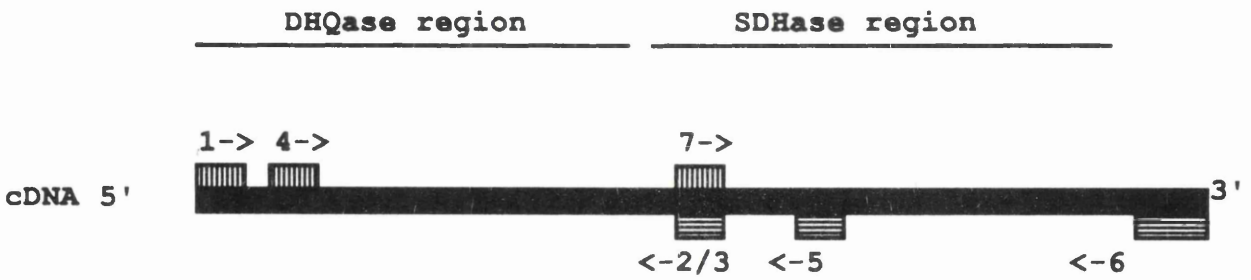


Fig.6.3 Schematic representation of primer positions for amplification

Vertical lines represent sense primers and horizontal lines represent antisense primers. The combinations of primers used in the "first" and "second" approaches were as follows: Primers-1 and 2 or 3; 4 and 2 or 3; 1 and 5; 4 and 5; 1 and 6; 4 and 6; 7 and 5; 7 and 6.

6.4.3 Second approach: primers based on the N-terminal sequence of DHQase-SDHase and the first-strand "universal" primer

This approach utilizes a combination of the universal primer which associates with polyadenylated mRNA molecules and a specific degenerate primer that corresponds to the N-terminal protein sequence (see Chapter V).

The N-terminal sequence of DHQase-SDHase was used to design oligonucleotides for use as sense primers (primers-1 and -4 in Table 6.1 and in Fig. 6.3) in PCR amplification.

The first-strand "universal" primer was redesigned for use in PCR amplification as an antisense primer (primer-6 in Table 6.1 and Fig. 6.3). This was based on part of the "anchor" portion of the Pharmacia *Not* I $d(T)_{18}$ bifunctional primer (Fig. 6.4).

PCR amplification did not give bands of the expected sizes. Pairing the primers based on the N-terminal sequence of the protein with the antisense primer designed from the consensus regions was also tried, again without success. All the various combination of primers tried are shown in Fig. 6.3.

At this stage, substantially more amino acid sequence data from peptides of DHQase-SDHase was obtained (see Chapter V). Analysis of these sequences revealed that the GXNTD box which was used in designing primer-5 is not conserved in the pea SDHase region (peptide V in Fig. 5.12 of Chapter V). This explained why the primer pair-7 and-5 gave no specific products with any of the plant cDNAs tried (*P. sativum*, *B. napus* and *A. thaliana*), while the same set of primers amplified a fragment of the correct size when used either with *E. coli* genomic DNA or M13 DNA containing the *aro*E gene (Anton and Coggins 1988) as template.

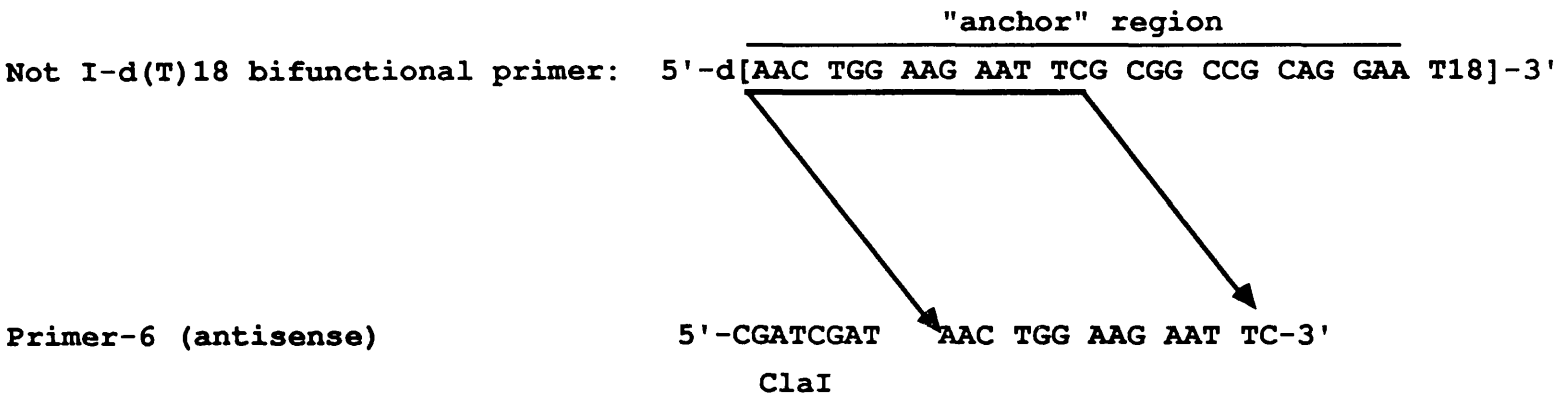


Fig. 6.4 The oligo d(T)18 primer used for the design of an antisense primer for PCR use

The antisense orientation of the bifunctional primer is shown. Underneath is shown the antisense primer-6 that was based on the "anchor" region of the bifunctional primer with the Cla I restriction site.

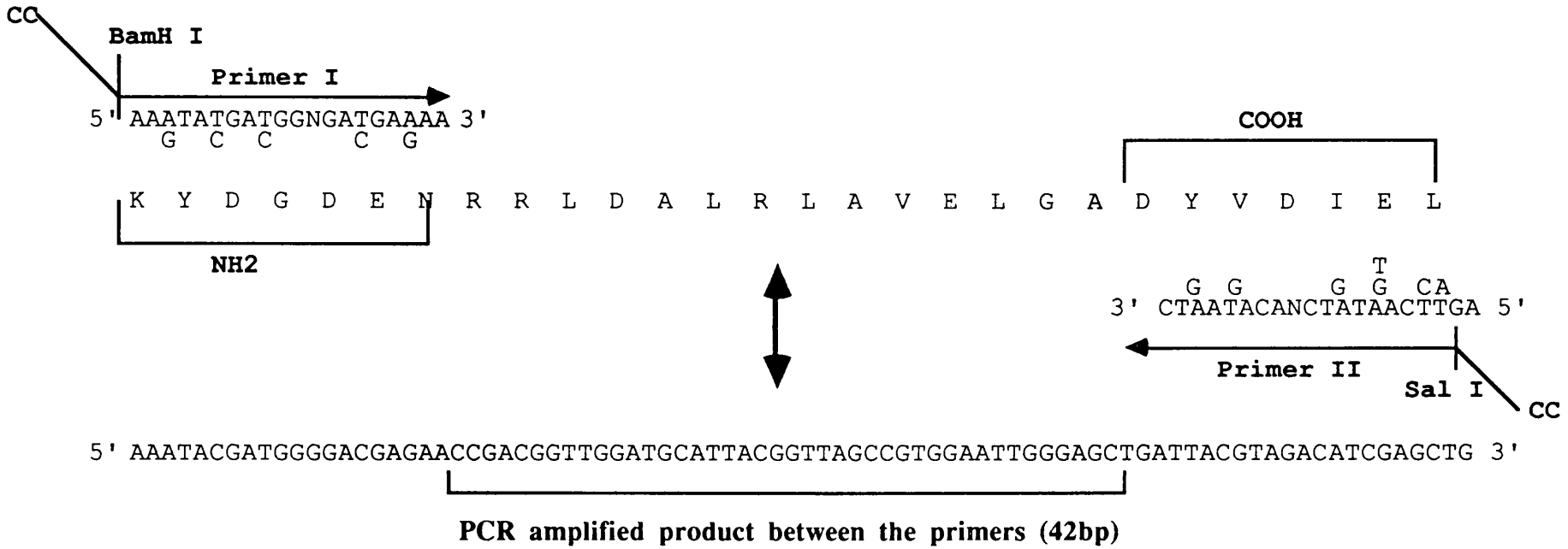


Fig. 6.5 Design of degenerate PCR primers from peptide sequence data

The sequence of peptide II from the pea DHQase region shown in single letter code. "Back translation" of this peptide into all possible corresponding DNA codons was used to design degenerate primers-I and II. N indicates that a mixture of all four bases was used. The primers are indicated by arrows pointing 5' to 3'. Restriction sites and two dCs were present at the 5' end of each primer. Underneath is shown the product of the PCR-reaction cloned and sequenced between the primers (underlined).

6.4.4 Third approach: successful use of internal pea DHQase specific primers for PCR amplification

Following the failure of the early PCR work it was decided to exploit the new internal amino acid sequence data obtained for pea DHQase-SDHase (see Chapter V). In particular, one excellent stretch of 40 amino acids (peptide II in Section 5.6.3) within the DHQase region seemed favourable for designing a pair of relatively long and closely spaced PCR primers (Fig. 6.5; primers I and II in Table 6.2). Various workers have obtained good results using such closely spaced pairs of primers (eg. Lingner *et al.*, 1991). This approach has the advantage that the size of the desired PCR product is known exactly in advance, which helps spurious products to be ignored.

PCR was performed as described in "Materials and Methods", Section 2.18.4. After initial denaturation at 95°C for 3min the PCR reaction was repeated for 40 cycles using the following cycle parameters: denaturation 94°C for 1min, annealing 60°C for 1min, extension 72°C for 2min. After the last cycle, the extension stage was continued for another 7min.

A band of the anticipated size was obtained (Fig. 6.6) and was extracted from a low melting point agarose/ethidium bromide gel (Section 2.18.4), cloned in M13 (Section 2.20.1) and sequenced (Section 2.23). The sequence, when translated, gave an exact match to the known peptide sequence between the two primers used (Fig. 6.5). Thus, the first piece of pea DHQase-SDHase cDNA had been obtained.

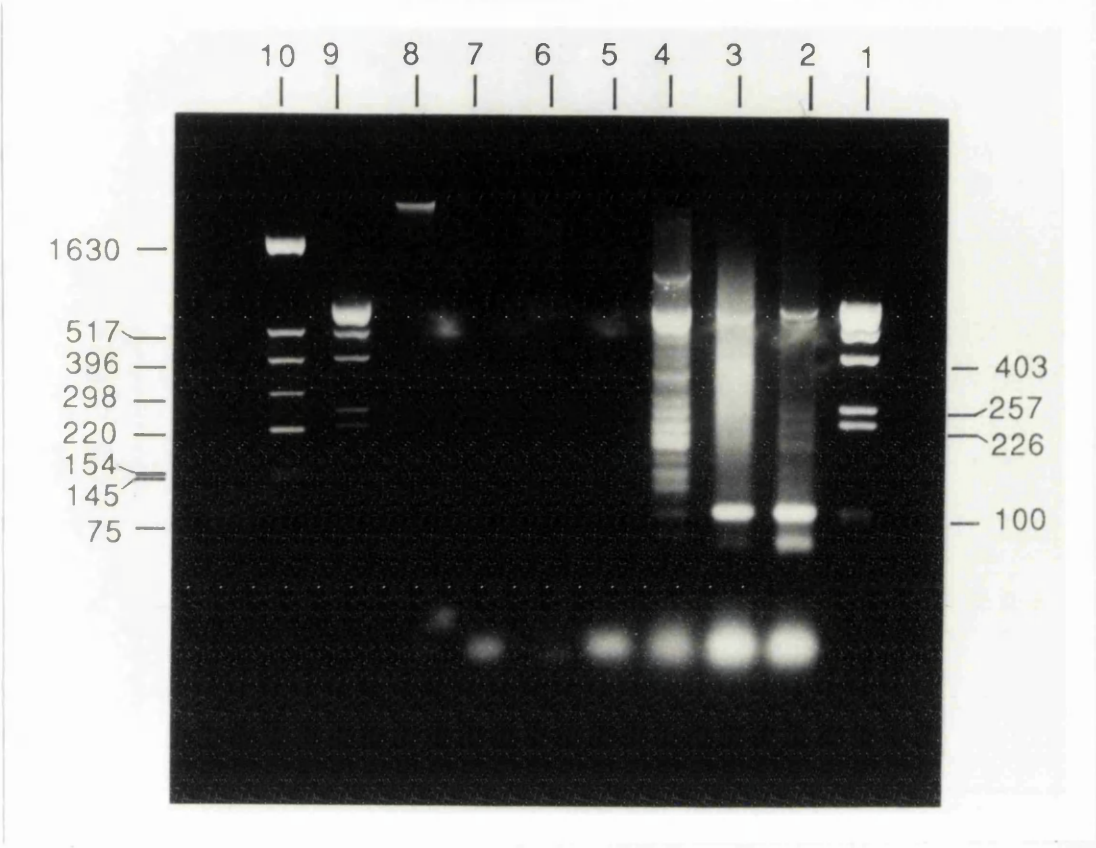


Fig. 6.6 Amplification of first-strand pea cDNA with DHQase specific primers (of pea DHQase-SDHase) by PCR

The PCR reaction mixture was run on a 3% agarose/ethidium bromide gel and photographed by UV illumination. *Lanes* 1 and 9-samples of pAT153 cut with *Alu* I and lane 10, pAT153 cut with *Hinf* I used as size markers; fragment sizes in bp are shown at both sides of the panel. *Lanes* 2 - 7, are the results of amplification of the cDNA and *lane* 8, the pea genomic DNA as template. The expected product (*arrowed*) is a 100bp fragment (see Fig. 6.5) in *lanes* 2 [cDNA primed with pd(N)6] and 3 [cDNA primed with Not I d(T)18] which is visible only when both primers (I and II) are present together. Unincorporated primers can be seen at the bottom of the gel. A number of other bands are also visible in the controls. *Lanes* 4 and 5, are single primer controls identical to *lanes* 2 and 3, but containing respectively either primer I or II only. *Lane* 6, contains no primer and *lane* 7, contains no template. 10% of PCR products were applied to each lane.

Table 6.2

Primers for amplification and sequencing of pea cDNA

	BamHI K Y D G D E N	
Primer-I	5'-CCGGATCC AA(AG)TA(TC)GA(TC)GGNGA(TC)GA(AG)AA-3'	
(Sense/28mer/256fold)		
	SalI L E I D V Y D	
Primer-II	5'-CCGTCGAC A(GA)(TC)TC(AGT)AT(AG)TCNAC(AG)TA(AG)TC-3'	
(Antisense/28mer/384fold)		
	EcoRI G E T V E K M V	
Primer-III	5'-CCGAATTC GGNGA(AG)ACNGTNGA(AG)AA(AG)ATGGT-3'	
(Sense/31mer/512fold)		
	EcoRI A V P D V E D	
Primer-IV	5'-CCGAATTC GCNACNGG(AG)TCNAC(TC)TC(AG)TC-3'	
(Antisense/28mer/512fold)		
	ClaI	
Primer-6	5'-CGATCGAT AACTGGAAGAATTC-3'	
(antisense/22mer)		
Primer-SP1	5'-TCCATTCATTGGACTTG-3'	
Primer-SP2	5'-CAAGTCCAATGAATGGA-3'	

N indicates that a mixture of all four bases was used; brackets indicate where a simpler mixture of bases was used. Restriction sites and two extra bases were present at the 5' end of each primer. Primers are shown either in sense or antisense orientation with their length and degenerecies. SP1 and SP2 were the specealized sequencing primers.

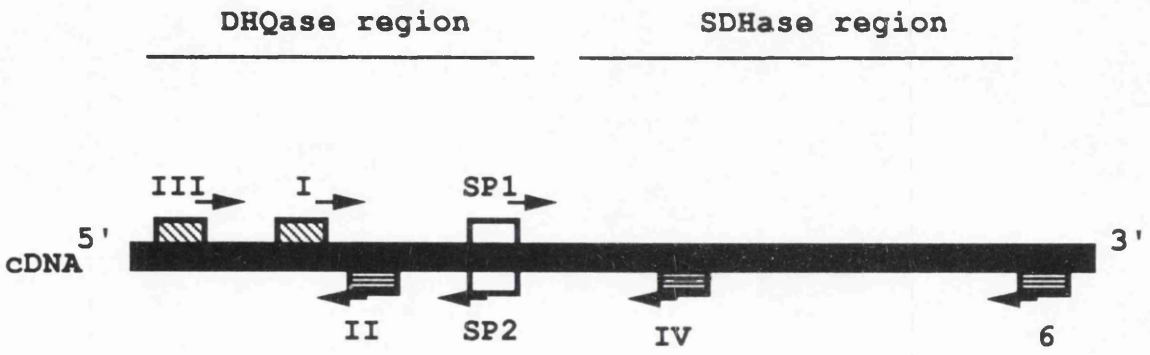


Fig. 6.7 Location of primers used for amplification and sequencing of the DHQase-SDHase cDNA fragments

Primers-I and III are sense and II, IV and 6 are antisense primers. SP1 and SP2 are specialised sequencing primers (see Table 6.2).

6.4.5 Building on the success of the third approach: strategies for amplification of DHQase cDNA fragments

The PCR-based approach to cloning cDNA has been used successfully during the past few years using degenerate primers made from protein sequence data. The success in amplifying a segment of DHQase-SDHase cDNA using specific primers (described in Section 6.4.4) encouraged us to design more primers to clone and sequence cDNAs encoding at least the DHQase region of the bifunctional plant protein.

The amplification strategy for isolation of overlapping DHQase cDNA fragments is shown in Fig. 6.7 (and see also Table 6.2). In outline, the primers I and II (used in above Section 6.4.4) were used separately in combination with new primers. Primer II (antisense) was used with a new primer (III) based on the extended N-terminal amino acid sequence. Primer III was longer than the previously tried N-terminal primers. Primer I (sense) was used with a new primer (IV) based on an internal peptide from the SDHase region. Thus it was hoped that the whole pea DHQase region would be obtained as a pair of overlapping cDNA segments. The likely PCR product sizes for primers III and II (~320bp), and primers I and IV (~720bp), were calculated from the polypeptide lengths of the *E. coli* enzymes. Amplification of single-stranded cDNA with primers I/IV produced a fragment of the expected size(Fig. 6.8). However, there was a problem with primers III and II: they gave two bands on the gels and neither of these exactly matched the expected size (Fig. 6.9). The band which was more pronounced appeared smaller than the expected size whereas the other band appeared longer. Both these bands were cloned for sequencing and their relationship will be discussed in Section 6.5.

To improve the yield of the ~720 bp fragment obtained with primers I and IV "nested" PCR was performed: here PCR was performed for 30 cycles with primers III and 6 (see Fig. 6.7) as described above (Section 6.4.4) but with a 50°C annealing temperature using first strand cDNA made from pea total RNA. Then 1% of this material was used for an additional 30 cycles using primers I and IV which lie inside the internally amplified DNA. During nested PCR (*i.e.* in the second stage) a higher annealing temperature (58°C for 2min) was used.

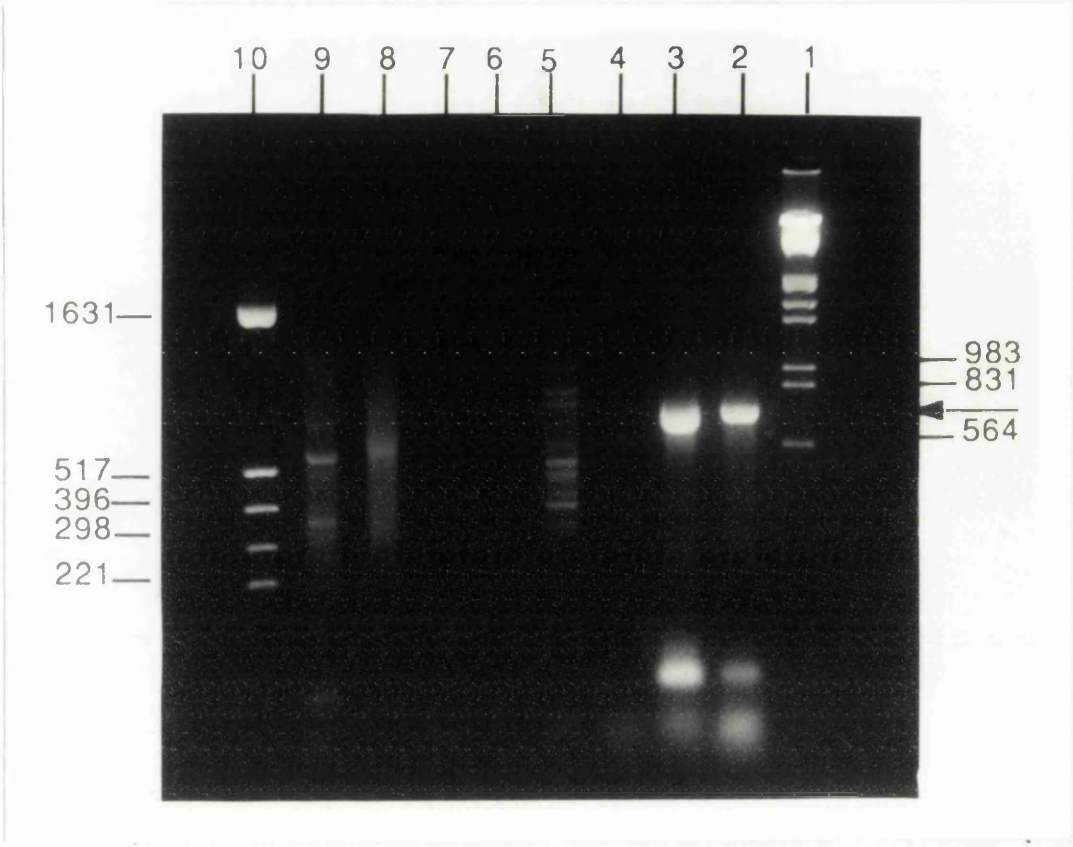


Fig. 6.8 Amplification of first-strand pea cDNA with DHQase-SDHase specific primers followed by "Nested" PCR

The PCR reaction mixture was run on a 2% agarose/ethidium bromide gel and photographed by UV illumination. *Lane* 1, samples of λ-DNA cut with *EcoR* I and *Hind* III and *lane* 10, pAT153 cut with *Hinf* I used as size markers; fragment sizes in bp are indicated at both sides of the panel. *Lanes* 8 [cDNA primed with Not I d(T)18] and 9 [cDNA primed with pd(N)6], primary amplified products generated as described in Section 6.4.4, using primers III and 6 (see Fig. 6.7). *Lanes* 2 - 7, nested amplification products of produced from the materials shown in *lane* 8. 1% samples of the reaction product shown in *lane* 8 were amplified for 30 cycles using primers-I and IV (see Fig. 6.7 and Section 6.4.5). The expected product (*arrowed*) is a 720bp fragment seen in *lanes* 2 and 3, of the nested PCR experiment. Unincorporated primers can be seen at the bottom of the gel. A number of other bands are also visible in controls. *Lanes* 4 and 5, are single primer controls identical to *lanes* 2 and 3, but containing respectively either primer I or IV only. *Lane* 6, contains no primer and *lane* 7, contains no template. 10% of the PCR products were applied to each lane.

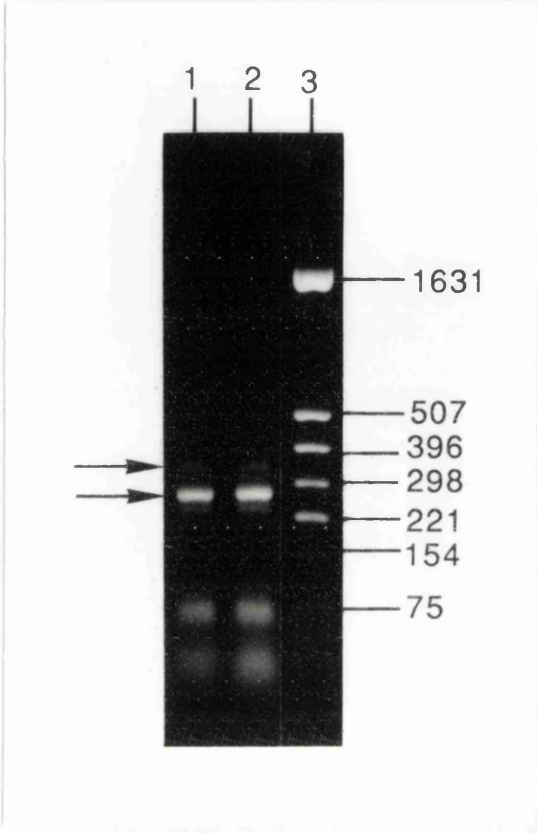


Fig. 6.9 Amplification of first-strand pea cDNA with DHQase specific primers (of pea DHQase-SDHase) by PCR

The PCR reaction mixture was run on a 2% agarose/ethidium bromide gel and photographed by UV illumination. *Lane* 3, sample of pAT153 cut with *Hinf* I used as size markers; fragment sizes are indicated in bp at the right side of the panel. *Lanes* 1 and 2, are the results of amplification of first-strand cDNA primed with pd(N)6. The expected product (*arrowed*) is in between 266bp and 346bp fragments and is visible when both primers (III and II) (see Section 6.4.5 and Fig. 6.7) are present together. Unincorporated primers can be seen at the bottom of the gel. A number of other bands are also visible. 10% of the PCR products were applied to each lane.

6.5 Cloning and sequencing of amplified cDNA fragments

PCR amplified cDNA segments were purified from low melting point agarose gels, cloned into M13 mp18 and 19, and sequenced. Clones containing inserts were sequenced on both strands as described in "Materials and Methods" (Section 2.23). Specialized sequencing primers (SP1 and SP2 in Table 6.2) were used as well as the universal M13 sequencing primer. dITP was used to eliminate the compressions in GC rich regions. Precautions were taken against the possibility of sequencing a mutant cDNA fragment arising from errors either by the reverse transcriptase or the *Taq* polymerase. The sequence of the 714 bp region was determined initially on both strands. It was checked by further sequencing of an independent clone (in M13) derived from a separate PCR reaction which used cDNA from a different synthesis to that used in generating the initial PCR template. It is highly improbable that two independent experiments would generate an identical mutation.

As mentioned in Section 6.4.5, with PCR primers III and II two products were obtained which differred slightly in size. Both were cloned and sequenced. One produced a sequence of 266bp and the other 346bp. The 266bp insert was a part of DHQase. The 346bp insert exactly matched the sequence of the 266bp insert except for an 80bp insertion which is almost certainly an intron (see Fig. 6.10). Although this second clone came from the same PCR reaction as the first, clearly the presence of the intron fortuitously marks this clone as having an independent origin. The second clone originated either from an independent cDNA derived from an aberrantly spliced mRNA or conceivably from contaminating genomic DNA in the total RNA preparation (Bürglin and Barnes, 1992).

```
1    GGAGAGACCGTGGAGAAAATGGTAGTTGATATACAAAAAGCTAAACTCAATGGAGCAGAC
       G   E   T   V   E   K   M   V   V   D   I   Q   K   A   K   L   N   G   A   D
61   CTTGTCGAAATTCGATTAGATTCTCTCAGCACTTTCAATCCTCATCAAGATCTTAACACT
       L   V   E   I   R   L   D   S   L   S   T   F   N   P   H   Q   D   L   N   T
121  TTCATTCAGCAACACCATTCCTTGCCCTTCTTGTTCACTTACAGGTTCCTTTTCGTTACT
       F   I   Q   Q   H   H   S   L   P   F   L   F   T   Y   _____________________
181  ATGTTCATTTTTGATTAATTGTTGAATTGTATTGAAGATGATAAAACATTATACCTTTTC
     _____________*_____________________*___*___*_____________________
241  TTAGACCGATATGGGAAGGTGGTAAGTATGATGGTGATGAAAACCGACGGTTGGATGCAT
     ___ R   P   I   W   E   G   G   K   Y   D   G   D   E   N   R   R   L   D   A
301  TACGGTTAGCCGTGGAATTGGGAGCTGATTACGTTGATATTGAACTT
       L   R   L   A   V   E   L   G   A   D   Y   V   D   I   E   L
```

Fig. 6.10 Location and characteristics of an intron (80bp) in the 346bp insert

The position of intron between codons is indicated by *underlining*. As in other eukaryotes the intron begins with GT and ends with AG (*bold face*), it is AT rich (75%) and it contains lots of stop codons (*shown by a asterisks*).

6.6 Sequence alignments

The 266 and 714 bp fragments overlapped by 40 bp (excluding primers, see Fig. 6.11). The 897 base pairs sequenced encode 299 amino acid residues (see Fig. 6.11). The N-terminal amino acid sequence (peptide-I in Table 5.5.I of Chapter V) of pea DHQase-SDHase overlaps with the start of the amino acid sequence deduced from the cDNA sequences. Only the amino acid sequence deduced from the cDNA sequences is shown in Fig. 6.11. The sequence right from the N-terminus of the protein is shown in the multiple alignment figure (Fig. 6.12). Multiple alignment showed that the bulk of the polypeptide, from its N-terminus, was homologous to all other known type I DHQases. The C-terminal region was homologous to SDHases. This confirmed our earlier tentative conclusion that the DHQase domain of the plant bifunctional protein is N-terminal. The plant DHQase region is shorter than *E. coli* DHQase by 20 amino acid residues and the majority of these were towards the N-terminus of the polypeptide (see Fig. 6.12). This explained why primers III and II gave a slightly smaller PCR product than that calculated on the basis of the *E. coli* enzyme.

Table 6.3 summarises the percentage identities from pair-wise comparisons of *P. sativum* DHQase with all the other type I DHQases available. These comparisons showed that at the amino acid level there was from 23 to 28% identity between the *P. sativum* dehydroquinase domain and the other type I dehydroquinase domains.

```
1    GGAGAGACCGTGGAGAAAATGGTAGTTGATATACAAAAAGCTAAACTCAATGGAGCAGAC
       G  E  T  V  E  K  M  V  V  D  I  Q  K  A  K  L  N  G  A  D
61   CTTGTCGAAATTCGATTAGATTCTCTCAGCACTTTCAATCCTCATCAAGATCTTAACACT
       L  V  E  I  R  L  D  S  L  S  T  F  N  P  H  Q  D  L  N  T
121  TTCATTCAGCAACACCATTCCTTGCCCTTCTTGTTCACTTACAGACCGATATGGGAAGGT
       F  I  Q  Q  H  H  S  L  P  F  L  F  T  Y  R  P  I  W  E  G
181  GGTAAGTATGATGGTGATGAAAACCGACGGTTGGATGCATTACGGTTAGCCGTGGAATTG
       G  K  Y  D  G  D  E  N  R  R  L  D  A  L  R  L  A  V  E  L
241  GGAGCTGATTACGTTGATATTGAACTTAAGGTTGCTCACGAGTTCTATGATTCTATACGT
       G  A  D  Y  V  D  I  E  L  K  V  A  H  E  F  Y  D  S  I  R
301  GGGAAGATGTTCAACAAAACCAAAGTCATTGTTTCATCTCACAATTATCAGTATACTCCG
       G  K  M  F  N  K  T  K  V  I  V  S  S  H  N  Y  Q  Y  T  P
361  TCAGTTGAGGATCTTGGCGACCTTGTGGCAAGAATACAAGCAACTGGGGCAGACATAGTA
       S  V  E  D  L  G  D  L  V  A  R  I  Q  A  T  G  A  D  I  V
421  AAGATTGCAACGACTGCCGTGGAGATCACTGACGTGGCACGTATGTTTCAAATATTGGTG
       K  I  A  T  T  A  V  E  I  T  D  V  A  R  M  F  Q  I  L  V
481  CATTCCCAAGTTCCATTCATTGGACTTGTTATGGGTGACAGGGGATTAGTTTCACGCGTG
       H  S  Q  V  P  F  I  G  L  V  M  G  D  R  G  L  V  S  R  V
541  CTTTGTGCAAAATTTGGTGGGTATCTCACTTTTGGTACACTCGAGTCAGGGGTAGTTTCA
       L  C  A  K  F  G  G  Y  L  T  F  G  T  L  E  S  G  V  V  S
601  GCTCCTGGTCAACCTACAATTAAGGATCTGTTGCATCTATACAATTTCAGACAACTGGGA
       A  P  G  Q  P  T  I  K  D  L  L  H  L  Y  N  F  R  Q  L  G
661  CCTGAGACAAAAGTATATGGGATTATTGGGAAGCCTGTCAGTCACAGTAAATCACCCATA
       P  E  T  K  V  Y  G  I  I  G  K  P  V  S  H  S  K  S  P  I
721  TTGTTTAATGAAGCCTTCAAGACAGTTGGTTTCAATGGAGTTTTTGTATTTTTATTGGTG
       L  F  N  E  A  F  K  T  V  G  F  N  G  V  F  V  F  L  L  V
781  GATGACCTTGCCAATTTTCTCAGGACTTACTCATCAACAGATTTTGTGGGATTCAGTGTT
       D  D  L  A  N  F  L  R  T  Y  S  S  T  D  F  V  G  F  S  V
841  ACTATTCCTCACAAGGAGTCTGCCCTTAAGTGCTGTGACGAGGTAGACCCAGTCGCA
       T  I  P  H  K  E  S  A  L  K  C  C  D  E  V  D  P  V  A
```

Fig.6.11 Nucleotide and predicted amino acid sequence of cDNA encoding part of pea DHQase-SDHase

The nucleotide sequence represents two overlapping (*shown double underlined*) cDNA clones of 266bp and 714bp. The predicted amino acid sequence (*single letter code*) is shown. The regions of sequence shown in *bold face* represent the sequences of the primers used for PCR amplification of cDNA. The amino acid sequences corresponding to those determined for V8 peptides are *underlined.*

```
                1                                                          60
E.coli          .....MKTVT VKDL.VIG.T GAPKIIVSLM AKDIASVKSE ALAYREADFD ILEWRVDHYA
S.typhi         .....MKTVT VKNL.IIG.E GMPKIIVSLM GRDINSVKAE ALAYREATFD ILEWRVDHFM
A.nidulans      .L..QVATGQ IDSLSIIK.E KEHSFFASLT LPDLRE.AGD ILEEVCVGSD AVELRVDLLK
S.cerevisiae    SFSKYIAT.. ITGVREIEIP SGRSAFVCLT FDDLTE.QTE NLTPICYGCE AVEVRVDHLA
P.sativum       .....MKNAT LICVPIMG.E TVEKMVVDIQ KAKLNG.... .........AD LVEIRLDSLS

               61                                                         120
E.coli          ......DLSN VESVMAAAKI LRETMPEKPL LFTFRSAKEG GE.QAISTEA YIALNRAAID
S.typhi         ......DIAS TQSVLTAARV IRDAMPDIPL LFTFRSAKEG GE.QTITTQH YLTLNRAAID
A.nidulans      DPASNNDIPS VDYVVEQLSF LRSRV.TLPI IFTIRTQSQG GRFPDNAHDA ALELYRLAFR
S.cerevisiae    NY.......S ADFVSKQLSI LRKATDSIPI IFTVRTMKQG GNFPDEEFKT LRELYDIALK
P.sativum       TFNPHQD... .......LNT FIQQHHSLPF LFTYRPIWEG GK.YDGDENR RLDALRLAVE

               121                                 *                      180
E.coli          SGLVDMIDLE LFTGDDQVKE TVAYAHAHDV KVVMSNHDFH KTPEAEEI.. IARLRKMQSF
S.typhi         SGLVDMIDLE LFTGDADVKA TVDYAHAHNV YVVMSNHDFH QTPSAEEM.. VSRLRKMQAL
A.nidulans      SG.CEFVDLD I.AFPEDMLR AVTEMKG.FS KIIASHHDPK GELSWANMSW IKFYNKALEY
S.cerevisiae    NG.VEFLDLE L.TLPTDIQY EVINKRG.NT KIIGSHHDFQ GLYSWDDAEW ENRFNQALTL
P.sativum       LG.ADYVDIE L.KVAHEFYD SIRGKMFNKT KVIVSSHNYQ YTPSVEDLG. .DLVARIQAT

               181  *                                                      240
E.coli          DADIPKIALM PQSTSDVLTL LAATLEMQEQ YADRPIITMS MAKTGVISRL AGEVFGSAAT
S.typhi         GADIPKIAVM PQSKHDVLTL LTATLEMQQH YADRPVIMTS MAKEGVISRL AGEVFGSAAT
A.nidulans      G.DIIKLVGV ARNIDDN.TA LRKFKNWAAE AHDVPLIAIN MGDQGQLSRI LNGFMTPVSH
S.cerevisiae    DVDVVKFVGT AVNFEDN... LR.LEHFRDT HKNKPLIAVN MTSKGSISRV LNNVLTPVTS
P.sativum       GADIVKIATT AVEITDVARM FQ.....ILV HSQVPFIGLV MGDRGLVSRV LCAKFGGYLT

               241                                277                     300
E.coli          FGAVKKA..S APGQISVNDL RTVLTILHQA .......M.. ....ETYAVF GNPIAHSKSP
S.typhi         FGAVKQA..S APGQIAVNDL RSVLMILHNA ..........
A.nidulans      PSLPFKA... APGQLSATEI RKGLSLMGEI KPKKFAI... ...PKKFAIF GSPISQSRSP
S.cerevisiae    DLLPNSA... APGQLTVAQI NKMYTSMGGI EPKELF.SMG GIEPKELFVV GKPIGHSRSP
P.sativum       FGTLESGVVS APGQPTIKDL LHLYNFR... .......QLG P.ETKVYGII GKPVSHSKSP

               301                                                        360
E.coli          FIHQQFAQQL NIEHPYGRVL APINDFINTL NAFFSAGGKG ANVTVPFKEE AFARADELTE
A.nidulans      ALHNTLFAQV GLPHNYTRLE TTNAQDVQEF I..RSPDFGG AFRNNSLKLD IMPLLDEVAA
S.cerevisiae    ILHNTGYEIL GLPHKFDKFE TESAQLVKEK LLDGNKNFGG AAVTIPLKLD IMQYMDELTD
P.sativum       ILFNEAFKTV GFNGVFVFLL VDDLANFLR. .TYSSTDFVG FSVTIPHKES ALKCCDEVDP

               361                                                        420
E.coli          RAALAGAVNT LMR..LE... DGRLLGDNTD GVGLLSDLER
A.nidulans      EAEIIGAVNT IIPVSTGKNT PSRLVGRNTD WQGMILSLRK
S.cerevisiae    AAKVIGAVNT VIP..LG... NKKFKGDNTD WLGIRNALIN
P.sativum       VAKSIGAVNx IVRxPT
```

Fig. 6.12 Multiple alignment of the deduced amino acid sequence of the *P. sativum* DHQase domain, and part of the SDHase domain, with other type I DHQases and SDHases

The first 12 and last 15 residues (*shown double underlined*) of the *P. sativum* sequence were from peptide sequences. Amino acids 1 to 277 from the DHQase region, and 277 from onwards are part of the SDHase region. Aligned positions of identical amino acid residues with the *P. sativum* are shown with *bold face*. Gaps indicated by *dots* were introduced for optimal alignment. * mark the active site residues of *E. coli* DHQase (His-143 and Lys-170). x refers to the residue unidentified during peptide sequencing.

Table 6.3

Sequence identities between of *P. sativum* DHQase and the type I DHQases of *S. cerevisiae*, *A. nidulans*, *S. typhi* and *E. coli*

P. sativum DHQase length (amino acids)	Conserved residues between *P. sativum* and *S. cerevisiae*	*A. nidulans*	*S. typhi*	*E. coli*
232	55 (24%)	53 (23%)	66 (28%)	60 (26%)

6.7 Discussion

6.7.1 Is DHQase cDNA a chimera?

As discussed in Chapter V, there is evidence for isozymic forms of DHQase-SDHase in diverse plants including peas (Fiedler and Schultz, 1985; Mousdale *et al.*, 1987). We found no size differences amongst the three activity peaks that were observed after an FPLC Mono Q column (Section 5.3). If there are isozymes, then these may result from differential splicing that involves small alterations in the coding potential of the mature mRNA resulting in multiple forms of DHQase-SDHase transcripts rather than as a result of differences in transcriptional read-through (Mattox *et al.*, 1992 and references therein). Nothern blot analysis may be helpful in detecting the existance of three messenger RNAs. There is a chance that the DHQase cDNA sequence presented here is chimeric with respect to isozymic sequences. However, we think this is unlikely for three reasons. Firstly, out of the three activity peaks, only the major peak was taken for protein sequencing and those data were used for primer design. Secondly, there is an exact match between the actual peptide sequences and those deduced from the DNA. Thirdly, in the 40 bp overlap region (excluding primers; see Fig. 6.11) both PCR products have identical sequences (including the third base positions). In both parts of the sequence there is unequivocal homology to other type I DHQases.

6.7.2 Active site histidine and lysine residues in a pea DHQase

As discussed in Section 1.5.3, several amino acid residues have been identified at or near the active site of type I DHQases. The residues which have been identified in the *E. coli* enzymes are: Lys-170 (Chaudhuri *et al.*, 1991) and His-143 (Deka *et al.*, 1992;

```
          137    * *                          *  *          *    184
E.coli          KVVMSNHDFH KTPEAEEI.. IARLRKMQSF DADIPKIALM PQSTSDVLTL
S.typhi         YVVMSNHDFH QTPSAEEM.. VSRLRKMQAL GADIPKIAVM PQSKHDVLTL
A.nidulans      KIIASHHDPK GELSWANMSW IKFYNKALEY G.DIIKLVGV ARNIDDN.TA
S.cerevisiae    KIIGSHHDFQ GLYSWDDAEW ENRFNQALTL DVDVVKFVGT AVNFEDN...
P.sativum       KVIVSSHNYQ YTPSVEDLG. .DLVARIQAT GADIVKIATT AVEITDVARM
```

Fig. 6.13 Comparison of amino acid sequences showing the proposed active site His-143 and Lys-170 regions in the type I enzymes

Identical amino acid residues are indicated by *asterisks*. Periods indicate gaps for optimal alignment. *Bold faces* indicate the active site residues identified in the *E. coli* enzyme, His-143 (Deka *et al.*, 1992) and Lys-170 (Chaudhuri *et al.*, 1991) are conserved in all type I enzymes including the *P. sativum* dehydroquinase. Underlined is the change of Asp to Asn in plant enzyme. The numbering is according to the *E. coli* enzyme.

Chapter III) which are both involved in the mechanism of type I DHQase catalysed reaction. These two conserved residue containing sequences, termed His-143 and Lys-170 regions, were found in all type I enzymes including pea DHQase (Fig. 6.13). This together with our chemical modification data presented in Chapter V (Section 5.5) establish the *P. sativum* DHQase as a type I enzyme.

In summary, we have cloned a part of DHQase-SDHase cDNA from pea using degenerate pea-specific PCR primers and this, in combination with N-terminal sequence data, allowed us to determine the complete amino acid sequence of the DHQase domain. In the future, it should be possible to obtain a full-length cDNA from *P. sativum* for the DHQase-SDHase.

CHAPTER VII

General discussion and future prospects

7.1 The achievements of this project

The aim of the work presented in this thesis was to study and compare the enzyme 3-dehydroquinase from *E. coli*, *A. nidulans* and *P. sativum*. The *E. coli* enzyme is an authentic type I dehydroquinase (White *et al.*, 1990; Kleanthous *et al.*, 1992), while the catabolic *A. nidulans* enzyme is an authentic type II dehydroquinase (Kleanthous *et al.*, 1992). The key objectives were to ientify the general base involved in the mechanism of the type I enzyme, to define some of the kinetic and mechanistic differences between the type I and type II enzymes, to establish whether the *P. sativum* enzyme was a type I or a type II enzyme and to obtain and sequence a cDNA encoding the *P. sativum* dehydroquinase domain. The first three objectives have been achieved through a combination of kinetic analyses and chemical modification studies on all three enzymes. The final objective has also been achieved by partial protein sequencing of the *P. sativum* bifunctional dehydroquinase-shikimate dehydrogenase together with PCR assisted cDNA cloning and sequencing of the *P. sativum* dehydroquinase coding region.

7.2 The mechanism and the active site of the type I dehydroquinases

The proposed mechanism of *E. coli* dehydroquinase presented in Chapter III (Section 3.1, page 65) requires a general base for the abstraction of the pro-R hydrogen which follows initial imine formation. The results presented in chapter III (which have now been published: Deka *et al.*, 1992) indicate that the most probable side chain involved is the imidazole group of His-143. This residue is conserved in all type I dehydroquinases including the *P. sativum* enzyme (see.Chapter VI, Fig. 6.13, page 199) and is the only conserved histidine residue. The pH-activity profile (Chaudhuri *et al.*, 1986) and the chemical modification data (especially the pH-dependence of the rate of inactivation, see

Chapter III) strongly implicate the histidine residue. However, it is possible that another side chain may also be involved, with the histidine, in the proton abstraction process. Asp-144 is conserved in all the bacterial and fungal dehydroquinases and the participation of this residue in a catalytic diad had been suggested (Bugg, 1989). It is very interesting to note that Asp-144 is not conserved in the *P. sativum* enzyme; it is replaced by an asparagine residue. This argues against the catalytic diad idea. Changing His-143 by site-directed mutagenesis would be expected to produce an enzyme which although still capable of forming an imine with the substrate would give a much diminished rate of product formation because of the absence of a correctly placed catalytic group.

The lysine residue involved in imine formation during the catalytic cycle and previously identified as Lys-170 in the *E. coli* sequence (Chaudhuri *et al.*, 1991) is conserved in all the known sequences. It is the only conserved lysine residue (Fig. 6.13, page 199). Changing this residue by site-directed mutagenesis would be expected to give rise to an inactive enzyme.

Chemical modification studies by Kleanthous *et al.* (1990) identified two methionine residues in the active site of *E. coli* dehydroquinase. One of these Met-23 is not generally conserved in the type I dehydroquinases but the other residue Met-205 is conserved in all species including *P. sativum* (Fig. 6.12, page 196).

The carboxylate-binding site of the type I enzymes has not yet been identified. There are only 5 completely conserved basic residues in the type I dehydroquinases. Two of these, His-143 and Lys-170, have been clearly identified as the general base and the imine forming residue. The remaining conserved basic residues are all arginines (Arg-48, -82 and -213, see Fig. 6.12). It seems very likely that one of these residues is the carboxylate binding site. It would therefore be of interest to use either chemical modification (probably with phenylglyoxal) or site-directed mutagenesis to establish which is the carboxylate-binding site.

7.3 The mechanism and the active site of the type II dehydroquinases

It is now clear that the type II enzymes must work by a completely different mechanism from the type I enzymes. The first convincing evidence for this was the observation that the type II enzymes cannot be inhibited by treatment with substrate and sodium borohydride (Chapter IV and Kleanthous *et al.*, 1992). The lack of sequence similarity and particularly the lack of a conserved lysine residue re-inforced the idea that the type II enzymes could not use a lysine-bound imine intermediate. The recent finding by Shneier *et al.* (1992) (mentioned in Chapter IV, page 136) that the type II enzymes catalyse an *anti* (*trans*) elimination rather than the *syn* (*cis*) elimination catalysed by the type I enzymes has provided the final proof that the mechanisms are distinct.

Although the mechanism of the type II dehydroquinase does not involve imine formation it is still very likely that the transition state of the reaction will involve a build up of negative charge on the keto oxygen at C3. One possible way of neutralising this charge would be for the enzyme to use a metal cation, for example a zinc ion, as a catalytic group. Metal analyses on the *A. nidulans* enzyme have provide no supporting evidence for this idea (C. Kleanthous, personal communication); also there is no evidence that the type II enzymes are inhibited by metal chelating agents such as EDTA (R.K.Deka, C. Kleanthous and J.R. Coggins, unpublished results). An alternative mechanism for neutralising the build up of charge would involve hydrogen bonding for example with backbone or side chain NH groups. Interactions of this kind are known to contribute to the stabilisation of the tetrahedral transition states of the chymotrypsin and subtilisin reactions (Henderson, 1970; Robertus *et al.*, 1972; Wells *et al.*, 1986)

It is also likely that there will be a role for a basic residue in proton abstaction in the type

II dehydroquinases. The chemical modification experiments with DEPC described in Chapter IV suggest that a histidine residue may be involved in the mechanism but the situation is clearly more complicated than for the type I enzymes since the pH-activity profiles are complex.

The experiments with phenylglyoxal suggest that there may be an arginine residue (or residues) involved in carboxylate binding.

Clearly much remains to be understood about the mechanism of the type II dehydroquinases. So far chemical modification experiments and studies on the role of metal ions have not provided an answer. Recently small crystals of the type II dehydroquinase of *Mycobacterium tuberculosis* have been obtained (D. Gourley, J. Moore, A. Hawkins, N.W. Isaacs and J.R. Coggins, unpublished results). If these crystals diffract determination of the three dimensional structure of the enzyme is the most promising route to understanding the mechanism.

7.4 The evolutionary implications of the existence of two classes of dehydroquinase

The mechanistic differences and the complete lack of sequence homology suggest that the two classes of dehydroquinase are not related. It appears that two different mechanisms for the dehydration of dehdroquinic acid have emerged independently during evolution. It is interesting that both types of enzyme are used for biosynthesis (White *et al.*, 1990; Garbe *et al.*, 1991) while only the type II enzymes appear to be used for catabolism (Giles *et al.*, 1985; da Silva *et al.*, 1986). In at least one species, *Amycolatopsis methanolica*, a type II enzyme is used both for biosynthesis and catabolism (Euverink *et al.*, 1992).

7.5 *Pisum sativum* dehydroquinase

The experiments described in Chapter V established that the *Pisum sativum* dehydroquinase is a type I enzyme. The sequence studies on the intact enzyme also established that the dehydroquinase domain is at the N-terminus of the bifunctional dehydroquinase-shikimate dehydrogenase protein. This finding was confirmed by the sequencing studies at the DNA level on the cDNA clone which contained the entire coding sequence of the dehydroquinase domain and the coding sequence of the first third of the shikimate dehydrogenase domain (see Chapter VI). This domain order is the same as the domain order of these two enzymes in the pentafunctional fungal *arom* proteins (Charles *et al.*, 1985; Coggins and Boocock, 1986; Duncan *et al.*, 1987). In the *arom* proteins the dehydroquinase and shikimate dehydrogenase domains occur next to each other at the C-terminal end of the pentafunctional polypeptide chains.

7.6 The *P. sativum* dehydroquinase-shikimate dehydrogenase cDNA clone

The PCR experiments described in Chapter VI resulted in the isolation of first plant cDNA clone containing the coding sequence of dehydroquinase. The cDNA clone was incomplete. Dehydroquinase occurs on a bifunctional polypeptide with shikimate dehydrogenase and although the clone appeared to carry the complete coding sequence for the dehydroquinase domain it carried only the first third of the coding sequence of the shikimate dehydrogenase domain. In future work it will be possible to use this clone to isolate a full length cDNA for the complete dehydroquinase-shikimate dehydrogenase bifunctional protein. Such a clone could then be used for the overexpression of plant dehydroquinase-shikimate dehydrogenase. The successful overexpression and purification of the plant bifunctional protein in, for example, a simple *E. coli* expression system would facilitate both further enzymological studies and structural studies on the protein.

7.7 The transit peptide

As discussed in the Introduction (pages 15-16) all the shikimate pathway enzymes are found predominantly in the plastids. Thus dehydroquinase-shikimate dehydrogenase is found mainly, but not exclusively, in the chloroplasts of peas (Mousdale and Coggins, 1985; Mousdale and Coggins, 1992). Proper comparmentalisation of nuclear-encoded chloroplast proteins relies on transit peptides (Karlin-Neumann and Tobin, 1986). To date full length cDNA's encoding four of the shikimate pathway enzymes have been obtained from higher plants (DAHP synthase, Dyer *et al.*, 1990; shikimate kinase, Schmid *et al.*, 1992; EPSP synthase, Shah *et al.*, 1986, Gasser *et al.*, 1988; chorismate synthase, Schaller *et al.*, 1991). All of these full length cDNA clones include the coding region for a transit peptide. The present clone was obtained using an amplification procedure which used degenerate primers based on sequence data from the mature protein which lacks the transit peptide. cDNA species containing the coding region for the transit peptide would not have been amplified. To characterise the transit peptide it will be necessary to obtain a full length cDNA clone. Alternatively information about the transit peptide could be deduced from a genomic clone for *P. sativum* dehydroquinase-shikimate dehydrogenase.

7.8 Isoenzymes

There is good evidence at the enzyme level for the occurrence of isoenzymes of *P. sativum* dehydroquinase-shikimate dehydrogenase (Mousdale *et al.*, 1987; Mousdale and Coggins, 1992). It is not known whether these isoenzymes are due to the occurrence of a gene family or whether they arise through alternative splicing at the mRNA level. Similarly the occurrence of the enzyme in different sub-cellular locations may also be due to multiple genes or to alternative splicing. The key question to answer first is whether there are multiple copies of the gene. This will require a careful southern blot analysis of pea genomic DNA

digested with restriction enzymes and probed with the PCR amplified cDNA. Hybridisation of gene-specific cDNA to different restriction fragments in genomic DNA would suggest that there was more than one copy of the dehydroquinase-shikimate dehydrogenase gene per haploid genome. In view of the interest in isozymes of the shikimate pathway enzymes it would be interesting to investigate such bands by attempting to clone them.

Only when the number of genes have been identified would it be sensible to address the possibility of alternative splicing as an explanation for the multiple forms and locations of the bifunctional enzyme. At this stage it would also be possible to study expression of the dehydroquinase-shikimate dehydrogenase gene. RNA could be isolated from various parts of pea plants during growth, blotted onto nitrocellulose and probed with a suitable cDNA fragment. This would indicate the regions of the pea plant where the dehydroquinase-shikimate dehydrogenase gene was being expressed most highly.

References

Ahmed, S.I. & Giles, N.H. (1969) *J. Bacteriol.* **99**, 231-237

Alberg, D.G. & Bartlett, P.A. (1989) *J. Am. Chem. Soc.* **108**, 2337-2340

Amrhein, N. (1986) In *The Shikimic Acid Pathway: Recent Advances in Phytochemistry* (E.E. Conn, ed.) Vol.20, Plenum Press.

Amrhein, N., Deus, B., Gehrke, P., Hollander, H., Schab, J., Schulz, A. & Steinrucken, H.C. (1981) *Proc. Plant Growth Regul. Soc.* America **8**, 99-106

Anderson, K.S., Sikorski, J.A. & Johnson, K.A. (1988a) *Biochemistry* **27**, 7395-7400

Anderson, K.S., Sikorski, J.A., Benesi, A.J. & Johnson, K.A. (1988b) *J. Am. Chem. Soc.* **110**, 6577-6580

Anton, D.L., Hedatorm, L., Fish, E.M. & Abeles, R.H. (1983) *Biochemistry* **22**, 5903-5908

Anton, I.A. & Coggins, J.R. (1988) *Biochem.J.* **249**, 319-326

Badet-Denisot, M-A. & Badet, B. (1992) *Arch. Biochem. Biophys.* **292**, 475-478

Baillie, A.C., Corbett, J.R., Dowsett, J.R. & McCloskey, P. (1972) *Pestic. Sci.* **3**, 113-120

Balasubramanian, S., Davies, G.M., Coggins, J.R. & Abell, C. (1991) *J. Am. Chem. Soc.* **113**, 8945-8946

Bartlett, P.A. & Satake, K.J. (1988) *J. Am. Chem. Soc.* **110**, 1628-1630

Bej, A.K., Mahbubani, M.H. & Atlas, R.M. (1991) *CRC Crit. Rev. in Biochem. Mol. Biol.* **26**, 301-334

Bell, L.H., Coggins, J.R. & Milner-White, E.J.(1993) Submitted to *Protein Engineering.*

Bentley, R. (1990) *CRC Crit. Rev. Biochem. Mol. Biol.* **25**, 306-384

Beri, R.K., Grant, S., Roberts, C.F., Smith, M. & Hawkins, A.R. (1990) *Biochem. J.* **265**, 337-342

Berlyn, M.B. & Giles, N.H. (1969) *J. Bacteriol.* **99**, 222-230

Berlyn, M.B., Last, R.L. & Fink, G.R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4604-4608

Bickel, H., Palme, L. & Schultz, G. (1978) *Phytochemistry* **17**, 119-124

Biscoglio de Jimenez Bonino, M.J., Fukushima, J.G. & Cascone, O. (1986) *Anal. Biochem.* **157**, 8-11

Blanke, S.R., & Hager, L.P. (1989) *J. Biol. Chem.* **265**, 12454-12461

Bondinell, W.E., Vnek, J., Knowles, P.F., Sprecher, M. & Sprinson, D.B. (1971) *J.Biol.Chem.* **246**, 6191-6196

Boocock, M.R. & Coggins, J.R. (1983) *FEBS Lett.* **154**, 127-133

Boudet, A.M., Graziana, A. & Ranjeva, R. (1985) In *Annual Proceedings of the Phytochemical Society of Europe* (C.F. Van Sumere and P.J. Lea, eds.), Vol. 25, pp. 135-159. Clareendon Press, Oxford.

Boys, C.W.G., Fawcett, S.M., Sawyer, L., Moore, J.D., Charles, I.G., Hawkins, A.R., Deka, R., Kleanthous, C. & Coggins, J.R. (1992) *J. Mol. Biol.* **227**, 352-355

Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248-254

Buchholz, B. & Schuktz, G. (1980) *Z. Pflanzenphysiol* **100**, 209-215

Buchholz, B., Reupke, B., Bickel, H. & Schultz, G. (1979) *Phytochemistry* **18**, 1109-1111

Bugg, T.D.H. (1989) In *Active site studies on enzymes of the shikimate pathway*. Ph.D. Thesis, University of Cambridge, England.

REFERENCES

Bugg, T.D.H., Abell, C. & Coggins, J.R. (1988) *Tetrahedron Lett.* **29**, 6783-6786

Bürglin, T.R. & Barnes, T.M. (1992) *Nature* **357**, 367

Burstein, Y., Waish, K.A., & Neurath, H. (1974) *Biochemistry* **13**, 205-210

Butler, J.R., Alworth, W.L.& Nugent, M.J. (1974) *J. Am. Chem. Soc.* **96**, 175-184

Cain, R.B. (1981) In *Actinomycetes: Proceedings of the Fourth International Symposium on Actinomycete Biology* (K.P. Schaal & G. Pulverer, Eds.), pp. 335-354. New York: Gustav Fischer.

Carter, P., & Wells, J. A. (1988) *Nature* **332**, 564-568

Castellino, S., Leo, G.C., Sammons, R.D. & Sikorski, J.A. (1989) *Biochemistry* **28**, 3856-3862

Charles, I.G., Keyte, J.W., Brammer, W.J. & Hawkins, A.R. (1985) *Nucleic Acid Res.* **13**, 8119-8128

Charles, I.G., Keyte, J.W., Brammer, W.J., Smith, M. & Hawkins, A.R. (1986) *Nucleic Acid Res.* **14**, 2201-2213

Chaudhuri, S. & Coggins, J.R. (1985) *Biochem. J.* **226**, 217-223

Chaudhuri, S., Lambert, J.M., McColl, L.A. & Coggins, J.R. (1986) *Biochem. J.* **239**, 699-704

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

Cheung, S.-T. & Fonda, M.L. (1979) *Biochem. Biophys. Res. Commun.* **90**, 940-947

Church, F.C., Lundblad, R.L. & Noyes, C.M. (1985) *J. Biol. Chem.* **260**, 4936-4940

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. & Laemmli, U.K. (1977) *J. Biol. Chem.***252**, 1102-1106

REFERENCES

Coggins, J.R. (1986) In *Biotechnology and Crop Improvement and Protection Report* (P.R. Day, ed.), No.34, pp 101-110

Coggins, J.R. & Boocock, M.R.(1986) In *Multidomain proteins-structure and evolution* (D.G. Hardie and J.R. Coggins, eds.) Elsevier, Oxford.

Coggins, J.R., Boocock, M.R., Chaudhuri, S., Lambert, J.M., Lumsden, J., Nimmo, G.A. & Smith, D.D.S (1987) In *Methods Enzymol.* (S. Kaufman, ed.) Vol. 142, pp. 325-341. Academic Press, New York and London.

Coggins, J.R. (1989) In *Herbicides and Plant Metabolism* (A.D. Dodge, ed.), pp 97-112, Cambridge University Press, Cambridge.

Cohen, L.A.(1970) In *The Enzyme* (P.D. Boyer, ed.), 3rd ed., Vol. 1, pp.148-211, Academic Press, New York.

Cohen, S.N., Chang, A.C.Y. & Hsu, L. (1972) *Proc. Nat. Acad.Sci. USA* **69**, 2110-2116

Dan, L.G. & Britton, H.G. (1974) *Biochem. J.* **137**, 405-407

Dansette, P. & Azerad, R. (1974) *Biochimie* **56**, 751-755

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

Davis, B.D. (1951) *J. Biol. Chem.* **191**, 315-325

Davis, B.D. (1955) *Adv. Enzymol.* **16**, 247-312

DeFeyter, R.C. & Pittard, J. (1886) *J. Bacteriol.* **165**, 331-333

Deka, R.K., Kleanthous, C. & Coggins, J.R. (1992) *J. Biol. Chem.* **267**, 22237-22242

DeLeo, A.B., Dayan, J. and Sprinson, D.B. (1973) *J. Biol. Chem.* **248**, 2344-2353

Devereux, J. Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395

DiPersio, L.P., Fontaine, R.N. & Hui, D.Y.(1991) *J. Biol. Chem.* **266**, 4033-4036

Drapeau, G.R. (1976) In *Methods in Enzymology.* (L. Lorland, ed.) Vol.45, pp. 469, Academic Press, New York.

Duncan, K., Lewendon, A. & Coggins, J.R. (1984) *FEBS Lett.* **165**, 121-127

Duncan, K., Lewendon, A. & Coggins, J.R. (1985) *FEBS Lett.* **170**, 59-63

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

Duncan, K., Edwards, R.M. & Coggins, J.R. (1987) *Biochem.J.* **246**, 375-386

Dyer, W.E., Weller, S.C., Bressan, R.A. & Herrmann, K.M. (1988) *Plant Physiol.* **88**, 661-666

Dyer, W.E., Henstrand, J.H., Handa, A.K. & Herrmann, K.M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7370-7373

Dyer, W.E., Weaver, L.M., Zhao, J., Kuhn, D.N., Weller, S.C. & Herrmann, K.M. (1990) *J. Biol. Chem.* **265**, 1608-1614

Evensen, G., Mathiesen, A. & Sundan, A. (1991) *J. Biol. Chem.* **266**, 6848-6852

Ely, B. & Pittard, J. (1979) *J. Bacteriol.* **138**, 933-943

Euverink, G.J.W., Hessels, G.I., Vrijbloed, J.W., Coggins, J.R. & Dijkhuizen, L. (1992) *J. Gen. Microbiol* **138**, 2449-2457

Eykmann, J.F. (1885) *Rec. Trav. Chim.* **4**, 32-54

Fiedler, E. & Schultz, G. (1985) *Plant Physiol.* **79**, 212-218

Floss, H.G., Onderka, D.K. & Carrol, M. (1972) *J. Biol. Chem.* **247**, 736-744

Frost, J.W., Bender, J.L., Kadonaga, J.T. & Knowles, J.R. (1984) *Biochemistry.* **23**, 4470-4475

Gaertner, F.M. (1987) *Methods Enzymol.* **142**, 362-366

Gaertner, F.H. & Cole, K.W. (1977) *Biochem.Biophys.Res.Comm.* **75**, 259-264

Ganem, B. (1978) *Tetrahedron* **34**, 3353-3383

Ganson, R.J., d'Amato, T.A. & Jensen, R.A. (1986) *Plant Physiol.* **82**, 203-210

Ganson, R.J. & Jensen, R.A. (1988) *Arch. Bipchem. Biophys.* **260**, 85-93

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

Gasser, C.S., Winter, J.A., Hironaka, C.M. & Shah, D.M. (1988) *J. Biol. Chem.* **263**, 4280-4289.

Gibson, T.J. (1984) In *Studies on the Epstein-Barr virus genome.* Ph.D. thesis, Cambridge University, England.

Gibson, F. & Jackman, L.M. (1963) *Nature* (London) **198**, 388-389

Gibson, M.I. & Gibson, F. (1964) *Biochem. J.* **90**, 248-256

Gibson, F. & Pittard, A.J. (1968) *Bacteriological Review* **32**, 465-492

Giles, N.H., Case, M.E., Partridge, C.W.H. & Ahmed, S.I. (1967) *Proc. Nat. Acad. Sci. USA.* **58**, 1453-1460

Giles, N.H., Case, M.E., Baum, J., Geever, R., Huiret, L., Patel, V. & Tyler, B., (1985) *Microbiol. Rev.* **49**, 338-358

Girgis, S.I., Alevizaki, M., Denny, P., Ferrier, G.J.M. & Legon, S. (1988) *Nucleic Acids Res.* **16**, 10371

Goldsbrough, P.B., Hatch, E.M., Huang, B., Kosinski, W.G., Dyer, W.E., Herrmann, K.M. & Weller, S.C. (1990) *Plant Sci.* **72**, 53-62

Gorisch, H. & Lingens, F. (1971) *Biochim.Biophys.Acta.* **242**, 630-636

Granger, S. (1989) In *Studies of the genes for the enzymes of the shikimate pathway from Pisum sativum.* Ph.D. Thesis, University of Glasgow, Scotland.

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

Grewe, R. & Haendler, J. (1966) *Biochem. Prep.* **11**, 21-26

Gubler U.& Hoffman B.J. (1983) *Gene* **25**, 263-269

Hanson, K.R. & Rose, I.A. (1963) *Proc. Natl. Acad. Sci. USA* **50**, 981-988

Hasan, N. & Nester, E.W. (1978) *J. Biol. Chem.* **253**, 4993-4998

Haslam, E. (1974) In *The Shikimate Pathway*, Butterworth, London.

Hauptmann, R.M., della-cioppa, G., Smith, A.G., Kishore, G.M. & Widholm, J.M. (1988) *Mol. Gen. Genet.* **211**, 357-363

Henderson, R. (1970) *J. Mol. Biol.* **54**, 341-354

Hill, R.K. & Newkome, G.R. (1969) *J. Am. Chem. Soc.* **91**, 5893-5894

Holbrook, J.J. & Ingram, V.A. (1973) *Biochem. J.* **131**, 729-738

Horiike, K. & McCormick, D.B. (1979) *J. Theor. Biol.* **79**, 403-414

Horiike, K., Tsuge, H. & McCormick, D.B. (1979) *J. Biol. Chem.* **254**, 6638-6643

Huisman, O.C. & Kosuge,T. (1974) *J. Biol. Chem.* **249**, 6842-6848

Huynh, Q.K. (1987) *Arch. Biochem. Biophys.* **258**, 233-239

Huynh, Q.K. (1988) *J. Biol. Chem.* **263**, 11631-11635

Huynh, Q.K., Kishore, G.M. & Bild, G.S. (1988) *J. Biol. Chem.* **263**, 735-739

Jensen, R.A. (1986a) *Physiol. Plant.* **66**, 164-168

Jensen, R.A. (1986b) In *The Shikimic Acid Pathway: Recent Advances in Phytochemistry* (E.E. Conn, ed.) Vol.20, Plenum Press, New York.

Karlin-Neumann, G.A. & Tobin, E.M. (1986) *EMBO J.* **5**, 9-13

Keith, B., Dong, X., Ausubel, F.M. & Fink, G.R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8821-8825

Kishore, G.M. & Shah, D.M. (1988) *Ann. Rev. Biochem.* **57**, 627-663

Kitz, R. & Wilson, I.B. (1962) *J. Biol. Chem.* **237**, 3245-3249

Kleanthous, C., Cullis, P.M. & Shaw, W.V. (1985) *Biochemistry* **24**, 5307-5313

Kleanthous, C., Campbell, D.G. & Coggins, J.R. (1990) *J. Biol Chem.***265**, 10929-10934

Kleanthous, C. & Coggins, J.R. (1990) *J. Biol. Chem.* **265**, 10935-10939

Kleanthous, C., Reilly, M., Cooper, A., Kelly, S., Price, N.C. & Coggins, J.R. (1991) *J. Biol. Chem.* **266**, 10893-10898

Kleanthous, C., Deka, R., Davis, K., Kelly, S.M., Cooper, A., Harding, S.E., Price, N.C., Hawkins, A.R. & Coggins, J.R. (1992) *Biochem. J.* **282**, 687-695

Klee, H.J., Muskopf, Y.M. & Gasser, C.S. (1987) *Mol.Gen.Genet.* **210**, 437-441

Knoth, K., Roberds, S., Poteet, C. & Tamkun, M. (1988) *Nucleic Acids Res.* **16**, 10932

Koshiba, T. (1979) *Plant Cell Physiol.* **20**, 667-670

Kuriyan, J., Krishna, T.S.R., Wong, L., Guenther, B., Pahler, A., Williams, C.H.Jr. & Model, P. (1991) *Nature* **352**, 172-174

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

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

Lambert, J.M., Boocock, M.R. & Coggins, J.R. (1985) *Biochem. J.* **226**, 817-829

Last, R.L., Bissinger, P.H., Mahoney, D.J., Radwanski, E.R. & Fink, G.R. (1991) *Plant Cell* **3**, 345-358

Lee, C.C., Wu, X., Gibbs, R.A., Cook, R.G., Muzny, D.N. & Caskey, C.T. (1988) *Science* **239**, 1288-1291

LeMarechal, P. & Azerad, R. (1976) *Biochimie* **58**, 1123-1128

Levin, J.G. & Sprinson, D.B. (1964) *J. Biol. Chem.* **239**, 1142-1150

Levy, H.M., Leber, P.D. & Ryan, E.M. (1963) *J. Biol. Chem.* **238**, 3654-3659

Lewendon, A. & Coggins, J.R. (1983) *Biochem. J.* **213**, 187-191

Lingner, J., Kellermann, J. & Keller, W. (1991) *Nature* **354**, 496-498

Løbner-Olesen, A. & Marinus, M.G. (1992) *J. Bacteriol.* **174**, 525-529

Loosemore, M.J. & Patt, R.F. (1976) *FEBS Lett.* **72**, 155-158

Lourenco, E.J. & Neves, V.A. (1984) *Phytochemistry* **23**, 479

Lumsden, J. & Coggins, J.R. (1977) *Biochem. J.* **161**, 599-607

Lundblad, R.L. & Noyes, C.M. (1984) In *Chemical Reagents for Protein Modification*, Vol. I & II, CRC Press, Boca Raton, Florida.

Mack, D.H. & Sninsky, J.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6977-6981

Maskell, D.J., Morrisey, P. & Dougan, G. (1988) *J. Bacteriol.* **170**, 2467-2473

Matsudaria, P. (1987) *J. Biol. Chem.* **262**, 10035-10038

Mattox, W., Ryner, L. & Baker, B. S. (1992) *J. Biol. Chem.* **267**, 19023-19026

McCue, K.F. & Conn, E.E. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7374-7377

Means, G.E. & Feeney, R.E. (1973) In *Chemical modification of proteins*, Holden-Day, Inc., San Francisco, California.

Melchior, W.B., Jr. & Fahrney, D. (1970) *Biochemistry* **9**, 251-258

Messing, J. (1983) *Methods Enzymol.* **101**, 20-24

Miles, E.W. (1977) *Methods Enzymol.* **47**, 431-442

Millar, G. & Coggins, J.R. (1986) *FEBS Lett.* **200**, 11-17

Millar, G., Lewendon, A., Hunter, M.G. & Coggins, J.R. (1986) *Biochem.J.* **237**, 427-437

Miran, S.G., Chang, S.H. & Raushel, F.M. (1991) *Biochemistry* **30**, 7901-7907

Mitsuhashi, S. & Davis, B.D. (1954) *Biochim. Biophys. Acta.* **15**, 54-61

Morrel, H., Clark, M.J., Knowles, P.F. & Sprinson, D.B. (1967) *J. Biol. Chem.* **242**, 82-90

Morris, P.F., Doong, R.-L. & Jensen, R.A. (1989) *Plant Physiol.* **89**, 10-14

Mousdale, D.M.& Coggins, J.R. (1984) *Planta* **160**, 78-83

Mousdale, D.M. & Coggins, J.R. (1985) *Planta* **163**, 241-249

Mousdale, D.M. & Coggins, J.R. (1986) *FEBS Lett.* **205**, 328-332

Mousdale, D.M. & Coggins, J.R. (1992) In *Methods in Plant Biochemistry* (P.J. Lea, ed.) Vol.9 pp.1-23. Academic Press, New York and London.

Mousdale, D.M., Campbell, M.S. & Coggins, J.R. (1987) *Phytochemistry* **26**, 2665-2670

Muday, G.K. & Herrmann, K.M. (1992) *Plant Physiol.* **98**, 496-500

Muhlrad, A., Hegyi, G., & Toth, G. (1967) *Biochem. Biophys. Acta.* **2**, 19-26

Mullis, K.B. & Faloona, F.A. (1987) *Methods Enzymol.* **155**, 335-350

Nazfiger, E.D., Widholm, J.M., Steinrucken, H.C. & Killmer, J.L. (1984) *Plant Physiol.* **76**, 571-574

Nickbarg, E.B., Davenport, R.C., Pestko, G.A. & Knowles, J.R. (1988) *Biochemistry* **27**, 5948-5960

Nimmo, G.A. & Coggins, J.R. (1981) *Biochem. J.* **197**, 427-436

Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101-107

Ogino, T., Garner, C., Markley, J.L. and Herrmann, K.M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5828-5832

Onderka, D.K. & Floss, H.G. (1969) *J. Am. Chem. Soc.* **91**, 5894-5896

Ovadi, J., Libor, S. & Elodi, P. (1967) *Biochim. Biophys. Acta.* **2**, 455-458

Padgette, S.R., Huynh, Q.K., Aykent, S., Sammons, R.D., Sikorski, J.A. & Kishore, G.M. (1988a) *J. Biol. Chem.* **263**, 1798-1802

Padgette, S.R., Smith, C.E., Huynh, Q.K. & Kishore, G.M. (1988b) *Arch. Biochem. Biophys.* **266**, 254-262

Paravicini, G., Braus, G. & Hutter, R. (1988) *Mol. Gen. Genet.* **214**, 165-169

Patel, V.B. & Giles, N.H. (1979) *Biochim. Biophys. Acta* **567**, 24-34

Patil, R.V. & Dekker, E.E. (1990) *Nucleic Acids Res.* **18**, 3080

Pinto, J.E.B.P., Suzich, J.A. & Herrmann, K.M. (1986) *Plant Physiol.* **82**, 1040-1044

Pinto, J.E.B.P., Dyer, W.E., Weller, S.C. & Herrmann, K.M. (1988) *Plant Physiol.* **87**, 891-893

Pittard, J. & Wallace, B.J. (1966) *J. Bacteriol.* **92**, 1070-1075.

Pittard, J. (1987) In *Escherichia coli and Salmonella typhimurium : Cellular and Molecular Biology* (F.C. Neidhardt, ed.), Vol.1 American Society for Microbiology, Washington, D.C.

Polley, L.D. (1978) *Biochim. Biophys. Acta* **526**, 259-266

Pompliano, D.L., Reimer, L.M., Myrvold, S. & Frost, J.W. (1989) J. *Am. Chem. Soc.* **111**, 1866-1871

Powell, H.A., Kerby, N.W., Rowell, P., Mousdale, D.M. & Coggins, J.R. (1992) *Planta* **188**, 484-490

Ramakrishnan, V & Gerchman, S.E. (1991) *J. Biol. Chem.* **266**, 880-885

Ramjee, M.N., Balasubramanian, S., Abell, C., Coggins, J.R., Davies, G.M., Hawkes, T.R., Lowe, D.J. & Thorneley, R.N.F. (1992) *J. Am. Chem. Soc.* **114**, 3151-3153

Ray, J.M. & Bauerle, R. (1991) *J. Bacteriol.* **173**, 1894-1901

Riordan, J.F. (1979) *Mol. Cell. Biochem.* **26**, 71-92

Robertus, J.D., Kraut, J., Alden, R.A. & Birktoft, J.J. (1972) *Biochemistry* **11**, 4293-4303

Rotenberg, S.L. & Sprinson, D.B. (1970) *Proc. Nat. Acad. Sci. USA.* **67**, 1669-1672

Rotenberg, S.L. & Sprinson, D.B. (1978) *J. Biol. Chem.* **253**, 2210-2215

Rubin, J.L., Gaines, C.G. & Jensen, R.A. (1982) *Plant Physiol.* **70**, 833-839

Rubin, J.L. & Jensen, R.A. (1985) *Plant Physiol.* **79**, 711-718

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988) *Science* **239**, 487-491

Salamon, I.I. & Davis, B.D. (1953) *J. Am. Chem. Soc.* **75**, 5567-5571

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Press, New York

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467

Sanger, F., Carlson, A.R., Barrell, B.G., Smith, A.J.H. & Roe, B. (1980) *J. Mol. Biol.* **143**, 161-178

Schaller, A., Windhofer, V. & Amrhein, N. (1990) *Arch. Biochem. Biophys.* **282**, 437-442

Schaller, A., Afferden, M.V., Windhofer, V., Bulow, S., Abel, G., Schmid, J. & Amrhein, N. (1991) *Plant Physol.* **97**, 1271-1279

Schmid, J., Schaller, A., Leibinger, U., Boll, W. & Amrhein, N. (1992) *The Plant J.* **2**, 375-383

Schmidt, C.L., Danneel, H.-J., Schultz, G. & Buchanan, B.B. (1990) *Plant Physiol.* **93**, 758-766

Servos, S., Chatfield, S., Hone, D., Levine, M., Dimitriadis, G., Pickard, D., Dougan, G., Fairweather, N. & Charles, I.A. (1991) *J. Gen. Microbiol.* **137**, 147-152

Shah, D.M., Horsch, R.B., Klee, H.J., Kishore, G.M., Winter, J.A., Tumer, N.E., Hironaka, C.M., Sanders, P.R., Gasser, C.S., Aykent, S.A., Siegel, N.R., Rogers, S.G. & Fraley, R.T. (1986) *Science* **233**, 478-481

Shneier, A., Kleanthous, C., Deka, R., Coggins, J.R. & Abell, C. (1991) *J. Am. Chem. Soc.***96**, 175-184

Shneier, A., Harris, J., Kleanthous, C., Coggins, J.R., Hawkins, A.R. & Abell, C. (1992) *Bioorg. Med. Chem. Lett.* (in the press).

Smart, C.C., Johanning, D., Muller, G. & Amrhein. N. (1985) *J. Biol. Chem.* **260**, 16338-16346

Smith, D.D.S. & Coggins, J.R. (1983) *Biochem. J.* **213**, 405-415

Smith, C.M., Pratt, D. & Thomson, G.A. (1986) *Plant Cell Rep.* **5**, 298-301

Sprinson, D.B. (1960) *Adv. Carbohydrate Chem.* **15**, 235-270

Srinivasan, P.R. & Sprinson, D.B. (1959a) *J. Biol. Chem.* **234**, 716-722

Srinivasan, P.R. & Sprinson, D.B. (1959b) *J. Biol. Chem.* **238**, 3170-3175

Srinivasan, P.R., Rothschild, J. & Sprinson, D.B. (1963) *J. Biol. Chem.* **238**, 3176-3182

Srinivasan, P.R., Shigeura, H.T., Sprecher, M., Sprinson, D.B. & Davis (1956) *J. Biol. Chem.* **220**, 477-497

Stalker, D.M., Hiatt, W.R. & Comai, L. (1985) *J. Biol. Chem.* **260**, 4724-4728

Stallings, W.C., Abdel-Meguid, S.S., Lim, L.W., Shieh, H.-S., Dayringer, H.E., Leimgruber, N.K., Stegeman, R.A., Anderson, K.S., Sikorski, J.A., Padgette, S.R. & Kishore, G.M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5046-5050

Staub, M. & Denes, G. (1969a) *Biochim.Biophys.Acta.* **178**, 588-598

Staub, M. & Denes, G. (1969b) *Biochim.Biophys.Acta.* **178**, 599-608

Steinrucken, H.C. & Amrhein, N. (1980) *Biochim.Biophys.Res.Comm.* **94**, 1207-1212

Steinrucken, H.C. & Amrhein, N. (1984) *Eur. J. Biochem.* **143**, 341-346

Steinrucken, H.C., Schulz, A., Amrhein, N., Porter, C.A. & Fraley, R.T. (1986) *Arch. Biochem. Biophys.* **244**, 169-178

Sugimoto, S. & Shiio, I. (1980) *J. Biochem.* **87**, 881-890

Suzich, J.A., Dean, J.F.D. & Herrmann, K.M. (1985) *Plant Physol.* **79**, 765-770

Suzich, J.A., Ranjeva, R., Hasegawa, P.M. & Herrmann, K.M. (1984) *Plant Physiol.* **75**, 369-371

Takahashi, K. (1968) *J. Biol. Chem.* **243**, 6171-6179

Takeuchi, M., Asano, N., Kameda, Y. & Matsui, K. (1986) *J. Biochem.* **99**, 1571-1577

Tsou, C.-L. (1962) *Sci. Sin.* (Engl. Ed) **11**, 1535-1558

Turner, M.J., Smith., B.W. & Haslam, E. (1975) *J. Chem. Soc.* Perkin **1**, 52-55

Vangrysperre, W., Ampe, C., Kersters-Hilderson, J. & Tempsa, P. (1989) *Biochem. J.* **263**, 195-199

Walker, J.E., Saraste, M., Runswick, M.J. & Gray, M.J. (1982) *EMBO J.* **1**, 945-951

Walsh, C. (1979) In *Enzymatic Reaction Mechanisms*, Freeman, San Franscico, California.

Wang, Y., Jones, J.D., Weller, S.C. & Goldsbrough, P.B. (1991) *Plant Mol. Biol.* **17**, 1127-1138

Warshel, A., Naray-Szabo, G., Sussman, F. & Hwang, J.K. (1989) *Biochemistry* **28**, 3631-3637

Weiss, B., Jacquemin-Sablon, A., Live, T.R., Fareed, G.C. & Richardson, C.C. (1968) *J. Biol. Chem.* **243**, 4543-4555

Weiss, U. & Edwards, J.H. (1980) In *The Biosynthesis of Aromatic Compounds*, Wiley, New York.

Weiss, U., Davis, B.D. & Mingioli, E. (1953) *J. Am. Chem. Soc.* **75**, 5572-5576

Weiss, U. & Mingioli, E.S. (1956) *J. Am. Chem. Soc.* **78**, 2894-2898

Welch,G.R., Cole, K.W. & Gaertner, F.H. (1974) *Arch. Biochim. Biophys.* **165**, 505-518

Wells, J.A., Cunningham, B.C., Graycar, T.P. & Estell, D.A. (1986) *Phil. Trans. R. Soc. Lond.* A **317**, 415-423

Wells, M.A (1973) *Biochemistry* **12**, 1086-1093

White, P.J., Millar, G. & Coggins, J.R. (1988) *Biochem. J.* **251**, 313-322

White, P.J., Mousdale, D.M. & Coggins, J.R. (1987) *Biochem. Soc.Trans.***15**, 144-145

White, P.J., Young, J., Hunter, I.S., Nimmo, H.G. & Coggins, J.R. (1990) *Biochem. J.* **265**, 735-738

Widlanski, T., Bender, S.L. & Knowles, J.R. (1989) *Biochemistry* **28**, 7572-7582

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103-119

Yaniv, H. & Gilvarg, C. (1955) *J. Biol. Chem.* **213**, 787-795

APPENDIX

The following publications contain some of the results obtained in the present work.

1. Shneier, A., Kleanthous, C., Deka, R., Coggins, J.R. and Abell, C. (1991) Observation of an imine intermediate on dehydroquinase by electrospray mass spectrometry. *J. Am. Chem. Soc.* **96**, 175-184.

2. Kleanthous, C., Deka, R., Davis, K., Kelly, S.M., Cooper, A., Harding, S.E., Price, N.C., Hawkins, A.R. and Coggins, J.R. (1992) A comparision of the enzymological and biophysical properties of two distinct classes of dehydroquinase enzymes. *Biochem. J.* **282**, 687-695.

3. Boys, C.W.G., Fawcett, S.M., Swayer, L., Moore, J.D., Charles, I.G., Hawkins, A.R., Deka, R., Kleanthous, C. and Coggins, J.R. (1992) The crystallisation of a type I 3-dehydroquinase from *Salmonella typhi. J. Mol. Biol.* **227**, 352-355.

4. Deka, R.K., Kleanthous, C. and Coggins, J.R. (1992) Identification of the essential histidine residue at the active site of *E. coli* dehydroquinase. *J. Biol. Chem.* **267**, 22237- 22242.

5. Deka, R.K., Anton, I.A., Dunbar, B. and Coggins, J.R. (1993) The characterisation of the shikimate pathway enzyme dehydroquinase from *Pisum sativum. Planta* (in press).