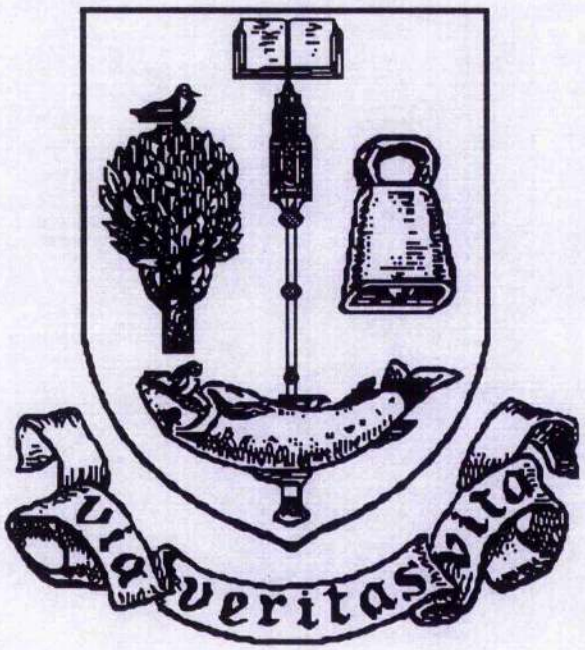
Studies on the active-site of *Escherichia coli* Shikimate dehydrogenase

A thesis
submitted to the
University of Glasgow

for the degree of
DOCTOR OF PHILOSOPHY

IN
BIOCHEMISTRY



BY
SUREKA CHACKREWARTHY

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

AUGUST, 1995

Dedicated

to my father and to the memory of my mother

ACKNOWLEDGEMENTS

The work described in this thesis would not have been possible without the help of many people. Sincere thanks to Professor J.R. Coggins for his guidance, enthusiasm and instruction during the course of this work and for his help in seeking funds for my final year of Ph. D. I would like to thank all the members of the molecular enzymology and protein chemistry research group and specially Tino Krell for his assistance with the ESMS work, Malcolm Horsburgh for his help with the sequence alignments and Andrew Elwell for his help in handling HPLC and Masslynx software. Many thanks to Mr. John Greene for technical assistance and to Dr. Brian Dunbar at the University of Aberdeen for peptide sequencing.

I would like to acknowledge the Department of Chemistry, University of Strathclyde for ESMS and LCMS facilities. I am indebted to the Government of Sri Lanka and the ORS award scheme for financial support.

Sincere thanks are extended to all my friends, whose companionship made life pleasant in Glasgow and specially to Asanga Wijayaratne for creating a homely atmosphere at Kelvinhaugh street!

Finally, special thanks are extended to my father who had the foresight to encourage me to undertake this Ph. D. and to Sunethra, Sujeeva, Dinesh, Kushan and Merrick for tolerating my long absence away from home.

List of Figures

List of Tables

Abbreviations

A	absorbance
ATP	adenosine triphosphate
Da	daltons
DAHP	3-deoxy-D-arabino-heptulosonate 7-phosphate
DEPC	diethylpyrocarbonate
DHQ	dehydroquinate
DHS	dehydroshikimate
DNA	deoxyribonucleic acid
ε	absorption coefficient
EPSP	5-enolpyruvyl-shikimate 3-phosphate
ESMS	electrospray mass spectrometry
FPLC	fast protein liquid chromatography
G3PDH	glyceraldehyde 3-phosphate dehydrogenase
GTP	guanosine triphosphate
HPLC	high pressure liquid chromatography
LADH	liver alcohol dehydrogenase
LCMS	liquid chromatography mass spectrometry
LDH	lactate dehydrogenase
MDH	malate dehydrogenase
M_R	molecular weight
NAD^+	nicotinamide adenine denucleotide (oxidised form)
$NADP^+$	nicotinamide adenine denucleotide phosphate (oxidised form)
NADH	nicotinamide adenine denucleotide (reduced form)
NADPH	nicotinamide adenine denucleotide phosphate (reduced form)
PGO	phenylglyoxal
PLP	pyridoxal phosphate
RNA	ribonucleic acid

rpm	revolution per minute
SDS	sodium dodecyl sulfate
SKDH	shikimate dehydrogenase
TFA	trifluoroacetic acid
TLCK	*N-p*-toluenesulphonyl-L-lysine chloromethylketone
TNBS	trinitrobenzene sulfonic acid
TNP	trinitrophenyl
TPCK	*N-p*-toluenesulphonyl-L-phenylalanine chloromethylketone
Tris	Tris (hydroxymethyl) aminomethane
U	units of enzyme activity
UV	ultraviolet
Vis	visible

Summary

Shikimate dehydrogenase (SKDH) catalyses the fourth step of the biosynthetic shikimate pathway, the reversible reduction of 3-dehydroshikimate to shikimate. This Thesis describes the work leading to the identificaton of three active-site residues in *E. coli* SKDH.

Group specific chemical modification with trinitrobenzene sulfonic acid (TNBS) and kinetic analyses indicated the presence of an essential lysine residue at the active-site. Electrospray mass spectrometry (ESMS) revealed that three lysine residues were modified by TNBS, and in the presence of substrate and coenzyme two were protected from modification. HPLC mass spectrometry (LCMS) identified the three modified residues as Lys-15, Lys-65 and Lys-217/219. In the presence of substrate and coenzyme Lys-65 was completely protected and Lys-15 was partially protected from modification. Sequence comparison with other known SKDH sequences allowed the identification of Lys-65 as the essential lysine residue. Experiments with methyl shikimate have provided evidence for a role for Lys-65 in substrate binding.

Arg-154 was identified as a component of the coenzyme ($NADP^+$) binding site by group specific chemical modification with phenylglyoxal (PGO). Characterisation of PGO modified SKDH by ESMS showed that a PGO modified arginine residue could have either a 1:1 or 2:1 stoichiometry. Protection experiments suggest that Arg-154 interacts with the 2' phosphate group of the adenosine moiety of the coenzyme.

Chemical modification with the group specific reagent diethylpyrocarbonate (DEPC) and kinetic analyses indicated the presence of an active-site histidine. Characterisation of DEPC modified SKDH by ESMS showed that in the presence of substrate and coenzyme two histidine residues were protected from DEPC modification. Differential peptide mapping using reverse phase HPLC identified the two protected residues as His-13 and His-253.

Sequence comparison with other SKDH sequences identified His-13 as the essential histidine. pH-dependence studies indicated a role for His-13 as a general acid/base in the catalytic reaction of SKDH.

As a preliminary step towards solving the three dimensional structure of SKDH small crystals have been obtained.

CHAPTER 1

Introduction

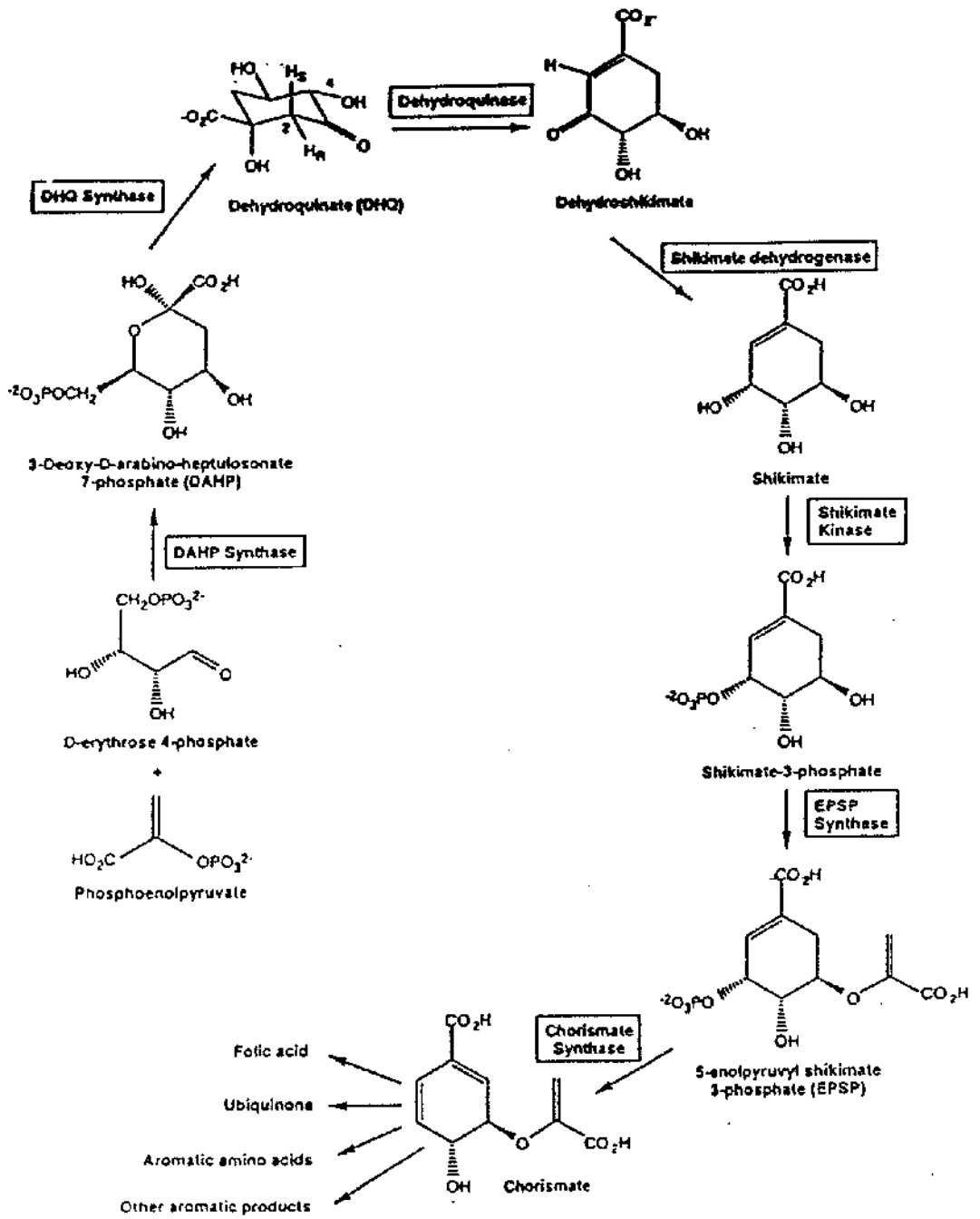


Fig. 1.1 The shikimate pathway

1.1 The shikimate pathway

1.1.1 A general introduction

In plants and in micro-organisms, the biosynthesis of all the aromatic compounds involved in primary metabolism proceeds by way of the shikimate pathway (Haslam, 1974; Weiss and Edwards, 1980; Conn, 1986; Bentley, 1990; Haslam, 1993). The pathway is named after the first isolated intermediate; shikimic acid from the fruit of the plant *Illicium religiosum* of which the Japanese name is *shikimi-no-ki* (Eykmann, 1885). Eykmann's investigations coupled with the work of Karrer and Link (1927), Fischer and Dangschat (1932, 1934) and Freudenberg *et al.* (1940) led to the elucidation of the structure and stereochemistry of shikimic acid. However, the complete pathway was elucidated only after the isolation and characterisation of pathway intermediates from mutant micro-organisms by Davis, Weiss, Sprinson and Gibson more than 30 years ago (Davis,1955; Levin and Sprinson, 1964; Gibson and Pittard, 1968).

Oxidation of glucose by glycolytic and pentose phosphate pathways yields phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4-P) respectively, the starting materials of shikimate pathway. The seven discrete enzymatic steps of the pathway begin with the condensation of these two products and proceeds via shikimate to chorismate which is the common precursor of all the aromatic amino acids and many other important compounds such as *p*-amino benzoic acid, ubiquinone and vitamin K (Haslam, 1974; Weiss and Edwards, 1980; Bentley, 1990; Haslam, 1993). The pathway is of particular importance in plants since it provides the precursors for the important structural polymer lignin, the flavinoids and secondary metabolites such as phenols and alkaloids (Coggins, 1989). It had been estimated that up to 35% of the dry weight of higher plants is generated via shikimate pathway (Boudet *et al.*, 1985), thus a substantial proportion of carbon fixed by plants is channelled through the shikimate pathway.

Chorismate was considered to be the only branch point until recently, but it is now obvious that many of the pathway intermediates function as branch points (Bentley, 1990; Haslam, 1993). Thus, a branch originating from the first intermediate of the

pathway, 3-deoxy-D-arabino-heptulosonate 7-phosphate is involved in the biosynthesis of 3-amino-5-hydroxybenzoate the proximate precursor of the *m*-C7N unit of ansamysin antibiotics (Kim *et al.*, 1992), while another branch originating from shikimate is found to be involved in the biosynthesis of the cyclohexanecarboxylate moiety of ansatrienin and the ω-cyclohexyl fatty acids (Casati *et al.*, 1987; Moore and Floss, 1994).

The shikimate pathway is studied for two main reasons:-

(a) Enzymes of the shikimate pathway are potential targets for herbicides and anti-microbial agents. Mammals lack the capacity for aromatic biosynthesis and depend on their dietary intake for nutritionally essential amino acids, vitamins and cofactors produced by this pathway. Thus, inhibitors of the enzymes of this pathway should prove to be harmless to mammals and are potential targets for herbicides and anti-microbial agents (Coggins, 1989). A detailed knowledge of the chemistry of the individual enzymes of the pathway may lead to the design of such inhibitors. Already, the herbicide glyphosate (active component of the weed killer Roundup) which is an inhibitor of the shikimate pathway enzyme EPSP synthase has a widespread application in agriculture (Amrhein *et al.*, 1980; Steinrucken and Amrhein, 1980; Boocock and Coggins, 1983; Kishore and Shah, 1988). The pathway has also attracted attention as a target for antimicrobials against the fungus *Pneumocystis carinii*, which is the principal cause of fatal pneumonia in AIDS patients in Europe and in USA (Banerji *et al.*, 1993) and also effects malnourished children (Pixley *et al.*, 1991). Recently (6S)-6-fluoro-shikimic acid, an analogue of shikimic acid was reported as a potential antibacterial agent for the use in human medicine (Davies *et al.*, 1994; Ewart *et al.*, 1995). Although the precise enzymatic target for this compound has not been confirmed this report demonstrates the vulnerability of this part of the aromatic biosynthetic pathway to potential attack by rationally designed compounds (Davies *et al.*, 1994).

(b) The shikimate pathway is a source of important secondary metabolites;

A host of important secondary metabolites such as antibiotics, mycotoxins, phenols and alkaloids are biosynthesised from shikimate-derived precursors. Genetic manipulation of the relevant biosynthetic steps may lead to the overproduction of some of these commercially important secondary metabolites.

1.1.2 Organisation of the shikimate pathway enzymes

Gene fusion and multifunctional enzymes emerged as a result of increasing complexity of the cell throughout evolution (Haslam, 1993). The shikimate pathway is an example of such an evolutionary modification in the organisation of enzymes. The seven enzymes catalysing the reaction sequence leading to the formation of chorismate are structurally and mechanistically similar in all organisms capable of aromatic biosynthesis, but their supramolecular organisation show considerable variation between species.

In *Escherichia coli* (and in other bacteria) the seven enzymes of the prechorismate pathway are monofunctional enzymes (Berlyn and Giles, 1969) and their genes are scattered throughout the genome (Pittard and Wallace; 1966). In fungi ,the enzymes catalysing the five consecutive steps involved in the conversion of DAHP to EPSP reside on a single pentafunctional polypeptide known as AROM, which is encoded by a single gene (Ahmed and Giles, 1969). The existence of a pentafunctional AROM protein has been demonstrated in *Neurospora crassa* (Lumsden and Coggins, 1977; Gaertner and Cole, 1977), *Schizosaccharomyces pombe* (Strauss, 1979) and *Euglena gracilis* (Patel and Giles, 1979). The AROM encoding genes have been cloned and characterised from *Aspergillus nidulans* (Charles *et al.*, 1986), *Saccharomyces cerevisiae* (Duncan *et al.*, 1987) and recently from *Pneumocystis carinii* (Banerji *et al.*, 1993). In plants the enzymes dehydroquinase and shikimate dehydrogenase which carry out two consecutive steps in the pathway are found as a bifunctional polypeptide whilst the other five enzymes of the pathway are monofunctional (Polley, 1978; Koshiba, 1979; Mousdale *et al.*, 1987). This bifunctional polypeptide has been cloned and characterised from *Pisum sativum* (Deka *et al.*, 1994) and *Nicotiana tabacum* (Bonner and Jensen,1994). This

organisation of shikimate pathway enzymes is illustrated in the schematic representation shown in Fig. 1.2.

The amino acid sequences of the AROM proteins studied to date are highly homologous with the monofunctional *Escherichia coli* counterparts and the most likely explanation for the origin of the gene encoding AROM is that it has evolved by the fusion of ancestral monofunctional genes like those found in *E. coli* (Hawkins, 1987). Studies on the *Aspergillus nidulans* AROM has shown that it is made up of two regions which can fold and function independently; the N-terminus containing dehydroquinate synthase and EPSP synthase domains and the C-terminus specifying shikimate kinase, dehydroquinase and shikimate dehydrogenase domains (Hawkins and Smith, 1991). In studies leading to the characterisation of the constituent domains of AROM in *A. nidulans* it has been shown that the EPSP domain is enzymatically active only when it is covalently linked to dehydroquinate synthase domain, suggesting that certain enzyme activities are stabilised by interactions between individual domains (Moore and Hawkins, 1993). However, more recent work on dehydroquinase and dehydroquinate synthase domains from the same organism has shown that they retain efficient catalytic activity in the monofunctional state (Hawkins *et al.*, 1993; Moore *et al.*, 1994). Characterisation of shikimate dehydrogenase and shikimate kinase domains have shown that they are not active as monofunctional domains, but are active as bifunctional proteins in combination with the 3-dehydroquinase domain (H. Lamb *et al.*, 1995/ manuscript submitted for publication).

1.1.3 Individual enzymes of the shikimate pathway

Detailed studies on the individual enzymes of the shikimate pathway had been hampered until recently, principally because the enzymes are present in low levels especially in plants and are difficult to purify, and secondly and importantly the substrates and coupling enzymes required for assays were not readily available (Coggins, 1989; Bentley, 1990; Haslam, 1993). However with the advances in recombinant DNA technology, all seven genes encoding *Escherichia coli* shikimate pathway enzymes have been cloned, overexpressed and sequenced (Coggins, 1989 and references cited therein;

E. coli monofunctional enzymes

P. sativum monofunctional and bifunctional enzymes

A. nidulans and *S. cerevisiae* pentafunctional AROM enzyme

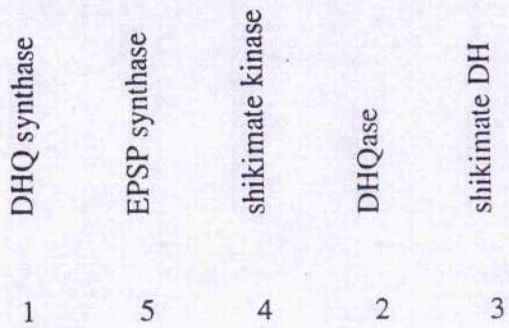


Fig. 1.2 Organisation of enzymes on the shikimate pathway

Bentley, 1990). As a result the enzymes are available in milligram quantities and studies of the enzymology of shikimate pathway has become an active area of biochemical research. The progress made over the recent years on each individual enzyme in the prechorismate part of the pathway is summarised below.

1.1.3.1 3-Deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase

DAHP synthase (EC 4.1.2.15) catalyses the first committed step in the shikimate pathway; the condensation of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to form DAHP (Step I, Fig 1.1). DAHP synthase is a regulatory enzyme and is known to posses the largest number of allosteric regulatory patterns described for any one protein (Bentley, 1990; Haslam, 1993). Four regulatory forms of this enzyme are known;

(1) DAHP synthase-O	Allosterically insensitive
(2) DAHP synthase-X	Inhibited by X
(3) DAHP synthase-X(Y)	Inhibited by both X and Y, with X being more powerful
(4) DAHP synthase-M^{2+}	Stimulated by divalent metal, M^{2+}

The precise mechanism of DAHP synthase reaction is not fully understood (Haslam, 1993). Reaction of E4P with PEP in which the enol oxygen atom was labelled with isotope ^{18}O has shown that the reaction proceeds with a C-O cleavage rather than a P-O cleavage since the isotopically labelled oxygen atom has been recovered in the inorganic orthophosphate released in the reaction (DeLeo and Sprinson, 1968; Nagano and Zalkin, 1970). A mechanism for the mode of action has been proposed in which an enzymic cysteinyl residue is involved (Ganem, 1978) as many DAHP synthases have been shown to be susceptible to inhibition by thiol modifying reagents (Staub and Denes, 1969; Huisman and Kosuge, 1974; Sugimoto and Shiio, 1980; Ray and Bauerle, 1991). In this mechanism the enzyme bound enolpyruvyl group condenses with E4P leading to the formation of DAHP and the release of free enzyme. This mechanism is consistent with the

required C-O cleavage noted earlier.

In *E.coli* three isozymes of DAHP synthase are known (Doy and Brown, 1965) which shows a high degree (41%) of sequence homology suggesting a common evolutionary origin (Bentley, 1990). These isozymes display an inhibition pattern in which each isozyme is feedback inhibited and repressed by one of the aromatic amino acids, providing the *E. coli* cell with the ability to regulate the synthetic rates in response to exogenous availability of aromatic amino acids (Bentley, 1990; Haslam, 1993). Thus, modulation of DAHP synthase isozymes is a major quantitative mechanism to control the flow of carbon into the shikimate pathway (Ogino *et al.*, 1982).

The existence of DAHP isozymes in many micro-organisms and their regulation through various types of feedback inhibition have been reported (Bentley, 1990). A different form of DAHP synthase is found in *Bacillus subtilis* , where it exists as a bifunctional protein with chorismate mutase, and is feedback inhibited mainly by prephenate and to a lesser extent by chorismate (Hasan and Nester, 1978).

In plants two isozymes of DAHP synthase are found in general, which are stimulated by Co^{+2} or Mn^{+2} (Rubin and Jensen, 1985; Ganson *et al.*, 1986, Morris *et al.*,1989). In some cases reported aromatic amino acids have an inhibitory effect on plant DAHP isozymes, whereas in some other cases they exert an activation effect (Suzich *et al.*,1985; Huisman and Kosuge, 1974; Graziana and Boudet, 1980; Pinto *et al.*, 1986). Thus, the contradictory observations of inhibition and regulatory patterns have not yet allowed general conclusions to be drawn for plant DAHP synthases. The important herbicide glyphosate is reported to inhibit the DAHP synthase-Co^{+2} from mung bean and from *Nicotiana silvestris* (Rubin *et al.*, 1982; Ganson and Jensen, 1988). More recent work on the plant DAHP synthase isozymes have shown that their synthesis is influenced by growth and environmental factors (Henstrand *et al.*, 1992; Sharma *et al.*, 1993).

1.1.3.2 3-Dehydroquinate synthase

3-Dehydroquinate (DHQ) synthase (EC 4.6.1.3) catalyses the ring closure of DAHP to form the first alicyclic intermediate of the pathway (Step II, Fig 1.1). The mechanism of 3-dehydroquinate synthase involves considerable chemistry; oxidation, β-elimination, reduction and intramolecular aldol condensation (Bartlett and Satake, 1988).

The possible overall mechanism of DHQ synthase has been elucidated by using substrate analogues (Bender *et al.*, 1989; Widlanski *et al.*, 1989a). Work done using synthetic 2-deoxy DAHP had shown that DHQ synthase catalyses only the first three steps of the normal reaction; oxidation, elimination and reduction (Bartlett and Satake, 1988; Widlanski *et al.*, 1989b). Thus, the final step of aldol cyclysation appears to be non-enzymatic. NAD^+ is essential for DHQ synthase activity and is involved in the oxidation and reduction of the hydroxyl group at C-5 position (Srinivasan *et al.*, 1963). The mechanism involves a base for deprotonation at C-6, which leads to subsequent elimination of phosphate. On the basis of work with carbacyclic phosphonate analogues of DAHP it had been shown that an enzyme group is not required and instead, the phosphate group of DAHP itself functions as the base (Widlanski *et al.*, 1989b). However, the positioning of the phosphate group appears to be crucial for proton abstraction. It had also been shown that the pyranose form of DAHP is involved in the mechanism (Widlanski *et al.*, 1989a).

3-Dehydroquinate synthase has been purified to homogeneity from *E. coli* (Frost *et al.*, 1984; Millar and Coggins, 1986) and *Phaseolus mungo* (Yamamoto, 1980) and appear to be dependent on a divalent metal ion for activity. Interestingly, DHQ synthase activity from AROM protein of *Neurospora crassa* also has a Zn^{+2} requirement (Lambert *et al.*, 1985). In *B. subtilis* the enzyme is associated with chorismate synthase and a flavin reductase in a multienzyme complex (Hasan and Nester, 1978). *E. coli* enzyme has a monomeric structure (Frost *et al.*, 1984; Millar and Coggins, 1986) whilst the plant enzyme from pea seedlings has a dimeric structure (Pompliano *et al.*, 1989).

1.1.3.3 3-Dehydroquinate dehydratase (3-Dehydroquinase)

3-Dehydroquinase (DHQase) (EC 4.2.1.10) catalyses the third step of the shikimate pathway; the conversion of 3-dehydroquinate to 3-dehydroshikimate (Step III, Fig 1.1). The *syn* elimination of a water molecule from 3-dehydroquinate introduces the first double bond into the aromatic ring system (Turner *et al.*, 1975). This reaction is common to two metabolic pathways; the biosynthetic shikimate pathway and the catabolic quinate pathway. Thus, two types of DHQases are found, namely type I and type II involved in shikimate and quinate pathways respectively. The two classes of enzymes were found to be unrelated at the level of primary structure (Garbe *et al.*, 1991).

The biosynthetic type I enzyme has been purified and characterised from an overproducing strain of *E. coli* (Chaudhuri *et al.*, 1986). The mechanism of action of type I DHQase from *E. coli* has been studied in detail. The *syn* elimination of elements of water proceeds via a Schiff's base intermediate formed by the reaction of substrate carbonyl group with the ε-amino group of an active site lysine followed by the abstraction of a proton by a general base (Walsh, 1979). In *E. coli* the active site lysine has been identified as Lys-170 (Chaudhuri, 1991) and His-143 as the general base involved in deprotonation (Deka *et al.*, 1992). Furthermore, Met-23 and Met-205 has been identified as likely active site residues but their role in catalysis has not been established (Kleanthous *et al.*, 1990; Kleanthous and Coggins, 1990). The type I enzymes from *Salmonella typhi* (Servos *et al.*, 1991), *A. nidulans* (Charles *et al.*, 1985) and *N. crassa* (Smith and Coggins, 1983) have been purified and the active site lysine involved in the Schiff's base formation has been identified in *N.crassa* (Chaudhuri *et al.*, 1991).

The type II enzyme is catabolic and quinate inducible. It has been characterised from *A.nidulans* (Da Silva *et al.*, 1986), *N. crassa* (Hawkins *et al.*, 1982) and from the bacteria *Streptomyces coelicolor* (White *et al.*, 1990) and *Mycobacterium tuberculosis* (Garbe *et al.*, 1991). A dual function type II enzyme has been isolated from *Amycolatropsis methanolica* which is involved in both quinate catabolism and biosynthesis of aromatic amino acids (Euverink *et al.*, 1992). Unlike type I enzymes the type II enzymes do not work through a Schiff's base intermediate (Kleanthous *et al.*, 1992). More

recently, work done on the active site of type II DHQases has resulted in the identification of an active site arginine residue which is thought to be involved in substrate binding (Krell *et al.*, 1995).

Type I and type II DHQases have different biophysical and mechanistic properties. Type I enzymes were found to be heat labile and dimeric in structure whilst type II enzymes were heat stable and dodecameric in structure (Kleanthous *et al.*, 1992). The elimination of water proceeds with opposite stereochemistries in type I and type II enzymes; *syn* and *anti* respectively (Schneier *et al.*, 1993). Crystallisation studies of DHQases have progressed over the recent years. Both a type I and a type II enzyme have been crystallised (from *S. typhi* and *M. tuberculosis* respectively) and the space groups determined (Boys *et al.*, 1992; Gourley *et al.*, 1994).

1.1.3.4 Shikimate dehydrogenase

This is the major topic of this thesis and is discussed in section 1.2.2.

1.1.3.5 Shikimate kinase

Shikimate kinase (EC 2.7.1.71) catalyses the fifth step of the shikimate pathway, the phosphate transfer from ATP to the C-3 hydroxyl group of shikimate (Step V, Fig. 1.1). Unusually for an enzyme in the middle of a metabolic pathway two isozymic forms of shikimate kinases have been described in *E. coli* and *S.typhi* (Ely and Pittard, 1979; Berlyn and Giles, 1969).

In *E. coli*, shikimate kinase II, encoded by the *aro* L gene is under specific control of *tyrR* regulator gene and is repressed by tyrosine or tryptophan (Ely and Pittard, 1979; DeFeyter *et al.*, 1986). Specific modulation of gene expression usually affects genes encoding the first enzymes in a pathway, and in the example of the shikimate pathway in *E.coli* only three DAHP synthases and shikimate kinase II are subjected to such regulation. This led to the hypothesis that an unrecognised metabolic pathway branches at shikimate (Pittard, 1987). Recent work has provided evidence for the existence of such biosynthetic branches (Moore and Floss, 1994).

The gene encoding shikimate kinase II has been cloned , sequenced and overexpressed from *E. coli* and was found to be a monomer (Millar *et al.*, 1986; DeFeyter and Pittard, 1986a). Recently, the gene encoding shikimate kinase I from *E. coli* has also been identified (*aro* K) (Lobner-Olesen and Marinus, 1992). Kinetic analysis of the *E. coli* isozymes showed that the K_m for shikimate of shikimate kinase I is 100 fold higher than that for shikimate kinase II (DeFeyter and Pittard, 1986b). A likely explanation for this is that shikimate kinase II is the one which normally functions in aromatic biosynthesis in the *E. coli* cell and shikimate kinase I functions only when intracellular shikimate levels exceed a particular threshold (DeFeyter and Pittard, 1986b; Haslam, 1993), or possibly the enzyme has a completely different metabolic role. The primary structure of both enzymes contain a region homologous with other known kinases which is likely to be a part of the ATP binding domain (Millar *et al.*, 1986; Lobner-Olesen and Marinus, 1992).

In *B. subtilis* a single shikimate kinase exists as a component of a trifunctional enzyme complex with DAHP synthase and chorismate mutase activities (Nakatsukasa and Nester, 1972; Huang *et al.*, 1975).Dissociation of shikimate kinase from this complex results in the loss of activity. A single shikimate kinase isolated from tomato shows homology with bacterial enzymes and contains an ATP binding site (Schimd *et al.*, 1992).

1.1.3.6 5-Enolpyruvyl-shikimate 3-phosphate (EPSP) synthase

EPSP synthase (EC 2.5.1.19) catalyses the formation of EPSP from shikimate 3-phosphate and PEP (Step VI, Fig 1.1). This reaction is of considerable importance since EPSP synthase is the major target for inhibition of the broad spectrum herbicide glyphosate (the active component of the weed killer Roundup) (Steinrucken and Amrhein, 1980; Boocock and Coggins, 1983). Thus, this enzyme remains the most intensively investigated among the shikimate pathway enzymes (Bentley, 1990).

The genes encoding EPSP synthases have been isolated and sequenced from bacteria (Duncan *et al.*, 1984a,b; Stalker *et al.*, 1985), fungi (Duncan *et al.*, 1987; Charles *et al.*, 1986) and from a range of plants (Gasser *et al.*, 1988; Granger, 1989; Wang

et al.,1991). Comparison of EPSP synthase sequence data from bacterial (*E. coli* , *S. typhi*), fungal (*A. nidulans*, *S. cerevisiae*) and plant species (*Petunia hybrida*, tomato) showed a consistent pattern; the sequence homology observed between the bacterial and plant enzymes was much greater (approximately 54%) than that observed between the fungal and plant or the fungal and bacterial sequences (approximately 38%) suggesting a late divergence of the plant and bacterial EPSP synthase enzymes than the plant and fungal enzymes (Gasser *et al.*, 1988).

The agronomic importance of EPSP synthase has led to the detailed investigation of the mechanism of action . The 'Sprinson mechanism' proposed two decades ago, involves the transfer of a carboxyvinyl moiety from PEP to the hydroxyl group at C-5 of shikimate 3-phosphate and proceeds via a tetrahedral intermediate (Levin and Sprinson, 1960; Levin and Sprinson, 1964). Research done since then, has provided evidence to validate the 'Sprinson mechanism'. Isotope labelling studies have shown that the C-O bond of PEP is cleaved (rather than the O-P bond) in the elimination of phosphate (Bondinell *et al.*, 1971) and kinetic experiments have revealed that EPSP synthase functions through an ordered mechanism with shikimate 3-phosphate binding first, followed by PEP (Boocock and Coggins, 1983; Anderson *et al.*, 1988a). More recently, the isolation and characterisation of the tetrahedral intermediate has provided conclusive evidence for the 'Sprinson mechanism' (Anderson *et al.*, 1988b; Leo *et al.*, 1990).

The active site of EPSP synthase has been studied by group specific chemical modification and a large number of active site residues involved in catalysis or binding have been identified. These include; arginine (Arg-28 and Arg-131 for *Petunia hybrida*) (Padgette *et al.*, 1988a), glutamate (Glu-418 for *E. coli*) (Huynh, 1988), cysteine (Cys-408 for *E. coli*) (Padgette *et al.*, 1988b), lysine (Lys-22 for *E. coli* and Lys-23 for *P.hybrida*) (Huynh *et al.*, 1988a,b) and histidine (His-385 for *E. coli*) (Huynh, 1993). Protection experiments have shown that the essential histidine and glutamate residues are located at or near the glyphosate binding site (Huynh, 1988; 1993). The cysteine residue was found to be not essential for catalysis or binding, however its protection by shikimate 3-phosphate or glyphosate from chemical modification has been attributed to its proximity

to the active site. Further confirmation of these data must await the determination of the three dimensional structure of the enzyme.

The mode of inhibition of EPSP synthase by glyphosate is not clearly understood. It has been shown that glyphosate is a competitive inhibitor of PEP (Boocock and Coggins, 1983). The inhibition is strongly pH dependent (Steinrucken and Amrhein, 1984; Rubin *et al.*, 1984), with a striking increase between pH 6.0-8.0. Based on this evidence and the pK values of the various ionizable groups of glyphosate, it was suggested that the ionised form of glyphosate is the kinetically active form and that glyphosate acts as a transition state analogue of PEP (Anton *et al.*, 1983; Steinrucken and Amrhein, 1984). The definition of a single kinetically competent intermediate in the EPSP synthase reaction has prompted the design of new inhibitors. Two phosphonate analogues of the tetrahedral intermediate has been evaluated as inhibitors of EPSP synthase from *P. hybrida* (Alberg and Bartlett, 1989). Recently, it has been shown that the 3-phosphate group contributes significantly to substrate and inhibitor recognition at the shikimate 3-phosphate site (Miller *et al.*, 1993) and this work has been followed by the synthesis of potential inhibitors which are aromatic analogues of shikimate-3-phosphate and EPSP containing replacements of the 3-phosphate group (Miller *et al.*, 1994). Moreover, a glyphosate analogue namely, N-amino glyphosate has been synthesised recently and is reported as the first successful modification of the glyphosate skeleton which exhibits inhibitor properties comparable to glyphosate (Knowles *et al.*, 1993).

1.1.3.7 Chorismate synthase

The seventh and the final step in the common part of shikimate pathway is catalysed by the enzyme chorismate synthase (EC 4.6.1.4) (Step VII, Fig. 1.1). This enzyme introduces the second of the three double bonds which are necessary for formation of the aromatic ring. The reaction requires a reduced flavin cofactor although no net overall change in the redox state of the substrate is observed.

Chorismate synthase has been characterised from three microbial sources; *E. coli* (Morell *et al.*, 1967; White *et al.*, 1988), *B. subtilis* (Hasan and Nester, 1978) and

N. crassa (Welch *et al.*, 1974; White *et al.*, 1988) and from plant sources; *P. sativum* (Mousdale and Coggins, 1986), *Corydalis sempervirens* (Schaller *et al.*, 1990) and *Euglena gracilis* (Schaller *et al.*, 1991). Recently two isozymic forms of chorismate synthase have been identified in tomato (Gorlach *et al.*, 1993). The enzymes thus far described display marked differences in their ability to generate the reduced flavin cofactor necessary for catalysis. In *N. crassa*, *B. subtilis* and *E. gracilis* chorismate synthase is bifunctional and is associated with a flavin reductase activity (diaphorase activity) which can generate the reduced flavin cofactor via the oxidation of nicotinamide nucleotide under aerobic conditions. In contrast, chorismate synthase from *E. coli* , *P. sativum* and *C. sempervirens* is monofunctional and does not contain the diaphorase avtivity. Thus, reaction proceeds only when reduced flavin is supplied exogenously under anaerobic conditions (Morell *et al.*, 1967; Gaertner, 1987). Both *E. coli* and *N. crassa* enzymes are tetrameric in structure (White *et al.*, 1988) and the deduced subunit M_R is 38,183 and 50,000 respectively. The smaller subunit size of *E. coli* may reflect a 'missing' diaphorase domain.

The need for a reduced flavin cofactor although the conversion of EPSP to chorismate is redox neutral appears unusual. This has attracted attention to the mechanism of action of chorismate synthase. Several distinctive mechanisms have been proposed, but none of these addresses the question of the role of reduced flavin in the catalytic process. However, the reaction is known to proceed with *anti* stereochemistry and involves the removal of the C-6 *pro*-R hydrogen atom and the loss of phosphate group in a 1,4 elimination process (Floss *et al.*, 1972; Hill and Newkome, 1969). Recent experiments have demonstrated that there is a spectral change associated with the flavin during the catalytic reaction (Ramjee *et al.*, 1991) and that an enzyme bound flavin mononucleotide free radical is formed stoichiometrically in the presence of the substrate analogue 6R-6-fluro EPSP (Ramjee *et al.*, 1992). Based on the observation of a secondary tritium isotope effect for the reaction catalysed by *N. crassa* chorismate synthase, it had been suggested that the mechanism probably involves a cleavage of the C6-H and C3-O bonds in distinct but partially rate determining steps (Balasubramanian *et al.*, 1995).

1.1.4 The shikimate pathway as a source of secondary metabolites

A secondary metabolite is defined as a substance that occur in a limited number of organisms, not universally, and the function of which, if known, is not that of an intermediate in metabolism (Weiss and Edwards, 1980). These metabolites are a measure of the fitness of the organism to survive and may be involved in its welfare, life and development (Haslam, 1993). In micro-organisms and plants a host of important secondary metabolites such as antibiotics and alkaloids are derived wholly or partially via the shikimate pathway. They are of commercial importance and have a widespread application in medicine, agriculture and food sciences. A brief description of these two groups of secondary metabolites are given below.

1.1.4.1 Antibiotics

An antibiotic is a chemical substance produced by a micro-organism, which in low concentration inhibits the growth of other micro-organisms. Traditionally antibiotics have been produced by fermentation, but advances in synthetic chemistry have led to some antibiotics being wholly or partially chemically synthesised.

In certain micro-organisms a range of antibiotics are biosynthesised from precursors derived from the shikimate pathway. During normal growth the biosynthesis of aromatic amino acids would not be regarded as a major metabolic route considering the low abundance of such amino acids in proteins, but following cessation of growth and initiation of aromatic containing antibiotic biosynthesis an enormous metabolic load would need to be channelled through the shikimate pathway. A few examples of shikimate derived antibiotics are summarised below.

(a) Ansamycins

This class of antibiotics are produced by different micro-organisms including *Streptomyces* and *Norcardia* sp. Ansamycins are characterised by a cyclic structure consisting of an aromatic group (benzenoid or naphthalenoid) and a polyketide derived aromatic chain that form a bridge between two non adjacent positions on the aromatic nucleus. All ansamycins contain a *m*-C_7N unit derived from 3-amino hydroxybenzoate, of which the amino analogue of DAHP is the ultimate precursor (Casati *et al.*, 1987; Kim *et al.*, 1992). Rifamycins are the only group of ansamycins commercially exploited. They are active against Gram positive bacteria and strongly inhibit RNA synthesis by inhibition of DNA dependent RNA polymerase. Rifampin is one of the most useful members of this group and is a major drug for the treatment of tuberculosis and leprosy.

(b) Chloramphenicol

Chloramphenicol is a metabolite if *Streptomyces venezuelae* and other *Streptomyces* sp. It is biosynthesised from chorismate via L-*p*-amino-phenylalanine (L-PAPA) which is the proximate precursor (Dardenne *et al.*, 1975). It is an inhibitor of protein synthesis and exerts its action by inhibiting ribosomal peptidyltransferase activity. Cloramphenicol is primarily used in the treatment of typhoid fever, bacterial meningitis and anaerobic infections.

(c) Phenazines

There are more than 50 phenazine pigments formed exclusively by bacteria (Byng and Turner, 1977; Turner and Messenger, 1986). Phenazines show antibiotic properties and an ability to intercalate with double stranded DNA. Recent studies have shown anthranilate to be the shikimate pathway intermediate from which the precursor molecule of phenazines, phenazine-1-6-dicarboxylate is derived (Essar *et al.*, 1990). Phenazines are active against Gram negative bacteria, fungi and actinomycetes.

(d) Streptonigrin

Streptonigrin is an anticancer antibiotic produced by *Sreptomyces flocculus*. The biosynthetic pathway of streptonigrin is known in considerable detail. It is derived from L-tryptophan and 4-amino anthranilate, a modified shikimate pathway intermediate, which then condenses with a four carbon acid (probably oxaloacetate) derived from the tricarboxylic acid cycle (Erickson and Gould, 1987).

(e) Vancomycin

Vancomycin is a bactericidal peptide antibiotic active against many species of Gram-positive cocci. It is a heptapeptide and contains two units of β-hydroxychlorotyrosine as structural components which are derived from L-tyrosine (Hammond *et al.*, 1982). It interferes with peptidoglycan biosynthesis by binding irreversibly to the acyl-D-alanyl-D-alanine terminus of the membrane bound peptidoglycan precursor. It is a valuable alternative therapy for serious infections caused by methicillin-resistant staphylococci, especially in patients allergic to the β-lactam agents. It is also the drug of choice for treatment of intestinal infections caused by *Clostridium difficile*.

1.1.4.2 Alkaloids

Alkaloids are basic, nitrogen containing compounds with complex molecular structures. The major source of alkaloids are higher plants but they also occur in micro-organisms, insects and animals. Alkaloids posses pharmacological activity ranging from extreme toxicity to valuable pharmaceuticals.

Morphine is an alkaloid of shikimate origin produced by the *Papaver* sp. It is an analgesic and is a major drug used in the treatment of deep-seated pain. The biosynthetic pathway of morphine is well characterised and the ultimate precursor is 3,4 dihydroxyphenylalanine (DOPA) derived from L-tyrosine. Dopamine and 3,4 dihydroxyphenylpyruvate, both derived from DOPA react to form norlaudanosoline carboxylic acid, the first alkaloid in the biosynthetic pathway, which through a series of chemical reactions is converted to morphine (Robinson, 1981).

1.2 The family of pyridine nucleotide dependent dehydrogenases

1.2.1 A general introduction

In metabolic processes, nucleotides are involved in two classes of energy transfer mechanisms; the first class involves hydrolysis of high energy phosphate bonds in triphosphates such as ATP or GTP and the second class involves electron transfer through hydrogen atoms (redox) of bases such as nicotinamide and isoalloxazine. Pyridine nucleotides belong to the latter group and contain nicotinamide as NAD^+ or as $NADP^+$ which has an extra phosphate group at the 2' hydroxyl group of the adenosine moiety (Fig 1.3). The oxidized forms NAD^+ and $NADP^+$, accept electrons in the form of a hydride ion that binds covalently to the C4 atom of the nicotinamide moiety, to give the reduced forms NADH and NADPH respectively. NADH is primarily used for the generation of ATP whereas NADPH is used almost exclusively in reductive biosynthesis.

The pyridine nucleotide coenzymes are employed by a large number of enzymes for the transfer of redox equivalents to various metabolites. The reactions catalysed by these enzymes include the reversible reduction of carbonyl bonds (dehydrogenases), Schiff bases or disulphide bonds (reductases) and aromatic carbon-oxygen bonds (hydroxylases). Among these, the dehydrogenases comprise the largest and best studied family.

Two groups of pyridine nucleotide dependent dehydrogenases can be identified. The enzymes of one group require a metal atom such as zinc or iron for activity whereas the other group do not contain any catalytically important metal atom (Aronson *et al.*, 1989; Neale *et al.*, 1986; Jornvall *et al.*, 1984). In both cases the substrate carbonyl group is reduced by the transfer of a hydride ion from NAD(P)H (Fisher *et al.*, 1953; Loewus *et al.*, 1953). In many dehydrogenases including lactate dehydrogenase (LDH), malate dehydrogenase (MDH) and glyceraldehyde-3-phosphate dehydrogenase (G3PDH) the substrate carbonyl forms a hydrogen bond with a proton from an acid catalyst, thus polarising the carbonyl group and facilitating direct hydride transfer. In all three cases the acid catalyst involved in the mechanism is a histidine side chain (Adams *et al.*, 1973; Holbrook *et al.*, 1974; Moras *et al.*, 1975). However, in the presence of a catalytically

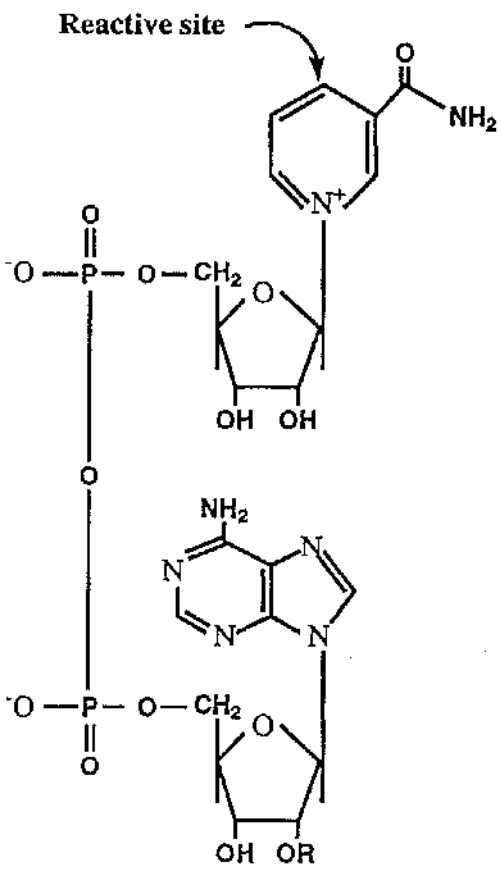


Fig. 1.3 Structure of the oxidised form of nicotinamide adenine dinucleotide (NAD^+) and of nicotinamide adenine dinucleotide phosphate ($NADP^+$). In NAD^+, R=H; in $NADP^+$, R=PO_3^{2-}

important zinc ion as in the case of liver alcohol dehydrogenase (LADH), the zinc ion takes the place of the acid catalyst, in polarizing the substrate carbonyl to facilitate hydride transfer (Eklund *et al.*, 1982). The transfer of the hydride ion from NAD(P)H is stereospecific, and two classes of dehydrogenases can be identified depending on the stereospecificity (Westheimer *et al.*, 1951; Fisher *et al.*, 1953). Class A dehydrogenases transfer the *pro-R* hydrogen (e.g. LDH, isocitrate dehydrogenase) and class B dehydrogenases transfer the *pro-S* hydrogen (e.g. G3PDH, glutamate dehydrogenase), depending on which side of the nicotinamide ring faces the substrate. Furthermore, it has been observed that class A dehydrogenases catalyse the reduction of more reactive carbonyls, while class B is associated with the less reactive carbonyls (Fersht, 1985).

In many dehydrogenases the binding of substrate and coenzyme follow a compulsory order; the coenzyme binds first to the enzyme to form a binary complex followed by substrate binding to form a ternary complex. The molecular explanation for this is the binding of the coenzyme causes a conformational change that increases the affinity of the enzyme for the other substrate. A few examples of the many dehydrogenases which follow a compulsory order in binding are LDH (Winer and Schwert, 1958; Schwert *et al.*, 1967; Zewe and Fromm, 1965; Holbrook *et al.*, 1975), alcohol dehydrogenase (Dalziel, 1963; Wratten and Cleland, 1963; 1965), ribitol dehydrogenase (Fromm and Nelson, 1962), isocitrate dehydrogenase (Nimmo, 1986) and glucose-6-phosphate dehydrogenase (Olive *et al.*, 1971). It had been shown for LDH and LADH that the ternary complex does not accumulate in the steady state but is rapidly converted to the product ternary complex which then dissociates to the product and enzyme-coenzyme binary complex (Dalziel, 1975; Holbrook and Gutfreund, 1973).

Based on the concept of structure-function relationships, structural similarities among the members of the pyridine nucleotide dependent dehydrogenase family can be expected. The three dimensional structures of many dehydrogenases have been solved and based on the structures of LDH (Adams *et al.*, 1970; 1973), G3PDH (Buehner *et al.*, 1973; 1974), MDH (Hill *et al.*, 1972; Webb *et al.*, 1973) and LADH (Branden *et al.*, 1973; Eklund *et al.*, 1974) certain generalisations concerning the structural features of the

dehydrogenase family can be made. The polypeptide chains of these four dehydrogenases fold into two clearly separated domains; a catalytic domain and a nucleotide binding domain. The structures of the catalytic domains differ in each case, but the structures of the nucleotide binding domains are remarkably similar although there is no sequence homology. Moreover, the position of this domain within the polypeptide chain varies in the four dehydrogenases. Comparison of the structures of the nucleotide binding domains of LDH, G3PDH and LADH had shown that each domain consists of two roughly identical folds, each associated with a mono-nucleotide binding (Fig. 1.4) (Rossmann *et al.*, 1974; Ohlsson *et al.*, 1974). This fold is known as the 'Rossmann fold' and consists of three parallel β-strands connected by helices in a right handed cross-over fashion (Richardson, 1976; 1977). The pyrophosphate group of the nucleotide binds in the central region of the domain at the N-terminal end of one of the connecting helices. The helix-dipole interacts favourably with the negatively charged pyrophosphate moiety and contributes significantly to the formation of nucleotide-enzyme complex (Hol *et al.*, 1978). This pyrophosphate binding helix form a part of a compact βαβ fold, and a characteristic fingerprint for the amino acid sequence of this fold had been deduced (Wierenga *et al.*, 1985). The βαβ fold is common to many pyridine nucleotide binding enzymes (Rossman *et al.*, 1975) and comprises of a universally conserved G-X-G-X-X-G (where X is any amino acid) sequence which appears at the border between the first strand of a β-sheet and the α-helix allowing the formation of a tight turn (Wierenga *et al.*, 1985; 1986). Other conserved features of this finger print region are (1) a hydrophilic residue at the N-terminus of the first β-strand, (2) a hydrophobic core composed of six small residues, and (3) a negatively charged residue at the C-terminus of the second β-strand which forms a hydrogen bond with the 2' hydroxyl group of the adenine ribose of NAD^+ (Wierenga *et al.*, 1985). In $NADP^+$ binding proteins this negatively charged residue has been replaced, presumably to accommodate the 2' phosphate group of adenine ribose (Wierenga *et al.*, 1985). It appears that this residue is an important means by which the dehydrogenases discriminate between NAD^+ and $NADP^+$ as coenzyme.

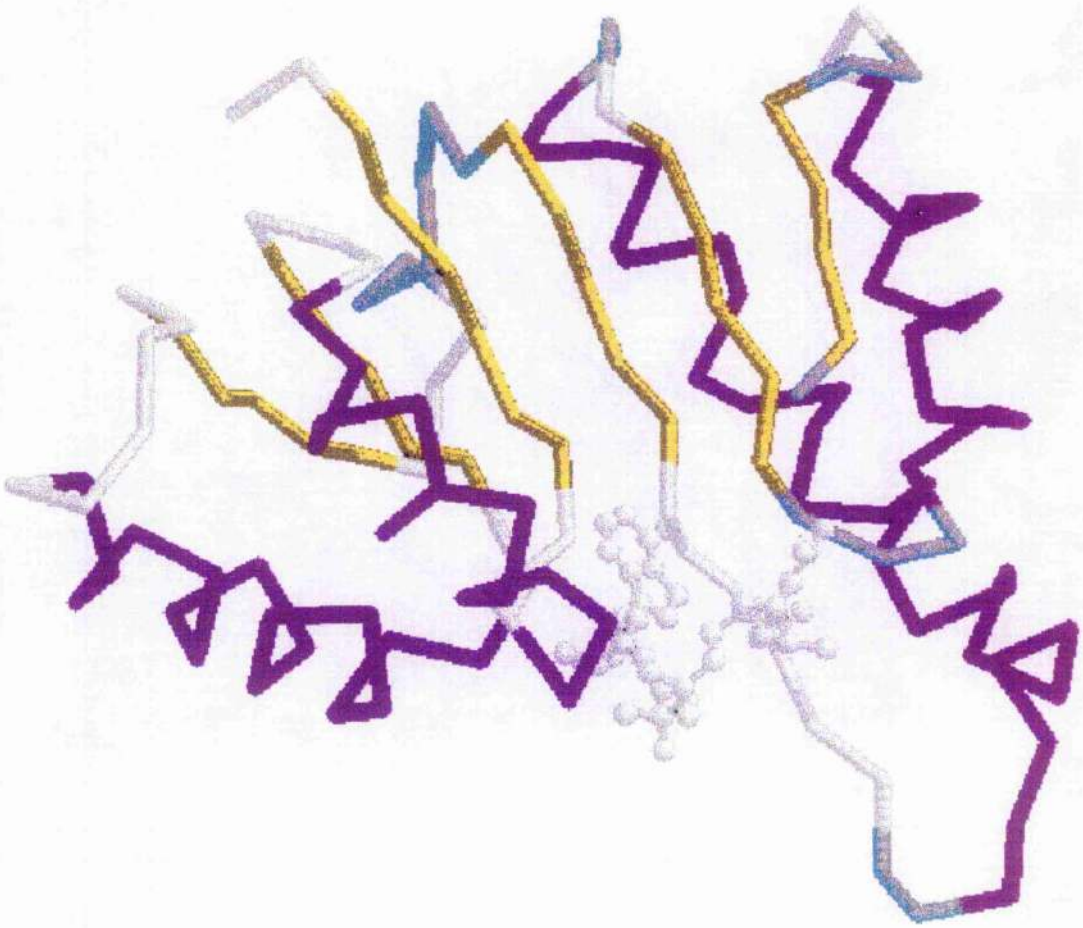


Fig. 1.4 The dinucleotide-binding domain of lactate dehydrogenase. The pyrophosphate group of NAD+ binds in the central region of the domain at the N-terminal end of one of the connecting helices.

1.2.2 Shikimate dehydrogenase

1.2.2.1 The reaction catalysed by shikimate dehydrogenase

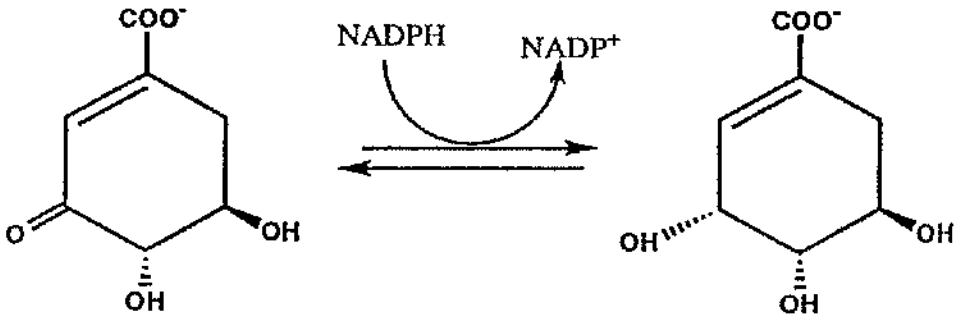


Fig. 1.5 Shikimate dehydrogenase catalyses the reversible reduction of 3-dehydroshikimate to shikimate

Shikimate dehydrogenase (SKDH) (EC 1.1.1.25) catalyses the fourth step of the shikimate pathway; the reversible reduction of 3-dehydroshikimate to shikimate (Fig. 1.5). It is an $NADP^+$-specific dehydrogenase (Yaniv and Gilvarg, 1955) which exhibits 'A'-type stereospecificity (Dansette and Azerad, 1974).

During the SKDH reaction, a proton is acquired by the carbonyl oxygen and the 4-*pro* -*R* hydrogen of the coenzyme (NADPH) is transferred as a hydride ion to the carbonyl carbon of the substrate. This proton can be considered most likely to be provided by a group capable of participating in hydrogen bonding and proton release/acceptance. In many pyridine nucleotide dependent dehydrogenases this role is fulfilled by a histidine residue (Price and Stevens, 1989). It had been observed that *E. coli* SKDH could be rapidly inactivated by histidine specific reagent diethyl pyrocarbonate and the involvement of a histidine residue in the catalytic mechanism of SKDH has been proposed (R. Syme and J.R. Coggins, unpublished data).

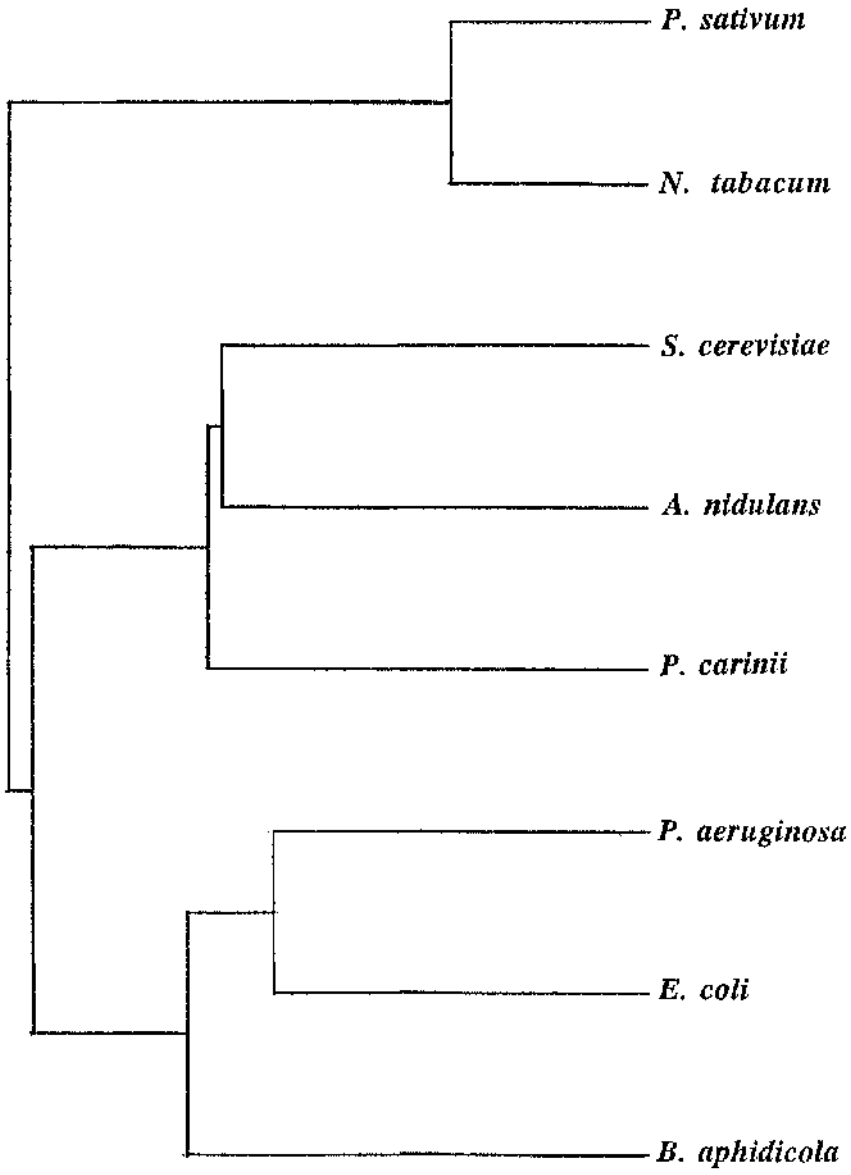


Fig. 1.6 A dendogram representing similarities between shikimate dehydrogenase sequences

1.2.2.2 Sequence similarities of shikimate dehydrogenase domains

The amino acid sequences of SKDH domains have been determined from *A. nidulans* (Charles *et al.*, 1986), *S.cerevisiae* (Duncan *et al.*, 1987), *E. coli* (Anton and Coggins, 1988), *P. carinii* (Banerji *et al.*, 1993), *P.sativum* (Deka *et al.*, 1994), *N. tabacum* (Bonner and Jensen, 1994), *Pseudomonas aeruginosa* (Hungerer *et al.*, EMBL/GenBank accession no. X85015) and *Buchnera aphidicola* (Rouhbakhsh, EMBL/GenBank accession no. U09230) and are now available for sequence comparison studies. Based on the similarity of sequence data, relationships between SKDH sequences from different organisms can be described. This is done by using GCG PileUp program (Devereux *et al.*, 1987) which creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. The two most similar sequences are aligned first producing a cluster followed by the alignment of the next most related sequence to this cluster or the next most related sequences to each other to produce another cluster. This is continued until all sequences have been included in a final pairwise alignment. A dendogram or a tree representing the clustering relationships between SKDH sequences are shown in Fig. 1.6. Accordingly, the two plant sequences (*P. sativum* and *N. tobacum*) appear to be the most closely related from all the SKDH sequences. Among the bacterial sequences, the *E. coli* sequence has more similarity with the *P. aeruginosa* sequence than with *B. aphidicola* sequence while *A.nidulans* and *S. cerevisiae* are closely related among the fungal sequences.

1.2.2.3 *E. coli* Shikimate dehydrogenase

SKDH occurs in very low levels in bacteria and it had been partially purified from wild type *E. coli* by Yaniv and Gilvarg (1955). This work was followed by Chaudhuri and Coggins (1985) and the enzyme was purified more than 10,000 fold to achieve homogeneity. As a result a very low concentration of enzyme has been available and this had hampered the detailed study of this enzyme. However, the gene encoding *E. coli* SKDH (*aro* E) has now been cloned, sequenced and placed under the control of a powerful promoter to facilitate overexpression (Anton and Coggins, 1984, 1988). As a

result milligram quantities of the enzyme can be readily obtained from 20g batches of *E. coli* cells (Chaudhuri *et al.*, 1987). This has opened the way for detailed studies on the mechanism and structure of SKDH.

Some of the kinetic properties of the partially purified *E. coli* SKDH have been reported. It had a pH optimum of 8.5 and the Michaelis constants (K_m) determined at pH 8.0 were $5.5x10^{-5}$ M and $3.1x10^{-5}$ M for shikimic acid and $NADP^+$ respectively (Yaniv and Gilvarg, 1955). However, the pure enzyme is reported to be more stable at higher pH values (Balinsky and Davies, 1961a).

E. coli SKDH is somewhat unusual in being a monomeric dehydrogenase (Chuadhuri and Coggins, 1985) since most of the dehydrogenases are larger proteins with multimeric structures. The complete polypeptide chain consists of 272 amino acid residues, and a deduced molecular weight of 29,414 Da (Anton and Coggins, 1988). One interesting feature of the amino acid sequence is the occurrence between residues 121 and 151, the amino acid 'fingerprint' characteristic of an ADP binding βαβ-fold (Anton and Coggins, 1988). The final position in the fingerprint must not be an acidic residue in $NADP^+$ linked dehydrogenases (Wierenga *et al.* 1985, 1986) and in the case of *E. coli* SKDH this position is occupied by a threonine residue. These observations strongly suggest that this fold is a part of the nucleotide binding domain of the enzyme.

In order to investigate enzyme-substrate interactions, the specificity of *E. coli* SKDH towards analogues of 3-dehydroshikimic acid (natural substrate), which lack the C-4 and C-5 hydroxyl groups and the carbon-carbon double bond has been studied (Bugg *et al.*, 1988). The 5-deoxy analogue was an excellent substrate (k_{cat} = 75 s^{-1}) compared to 3-dehydroshikimic acid (k_{cat} = 100 s^{-1}) indicating that the removal of the C-5 hydroxyl group has little effect on specificity. The removal of the C-4 hydroxyl group in the dideoxy analogue (lacking both C-4 and C-5 hydroxyl groups) decreased the k_{cat} value to 0.06 s^{-1} indicating the importance of C-4 hydroxyl group in enzyme-substrate interactions. This is consistent with the observation by Balinsky and Davies (1961b) that the most effective aromatic inhibitors of the reverse reaction of *P. sativum* SKDH possessed a para hydroxyl group. It has been suggested that this hydroxyl group forms a

hydrogen bond with a charged group at the enzyme active site (Dennis and Balinsky, 1972). The dideoxy-dihydro analogue lacking both hydroxyl groups and the double bond was also reduced slowly (k_{cat} = 0.09 s^{-1}). Furthermore, the enzyme behaves enantioselectively with respect to racemic substrates with a preference for the *S*-configuration at C-1 and C-4.

1.2.2.4 Isozymes of shikimate dehydrogenase

The existence of isozymic forms of SKDH in a variety of plant species has been reported. Early work has shown that SKDH from peas and *Phaseolus mungo* consists of two isozymic forms (Koshiba, 1978). Similarly, SKDH purified from tomato fruit resolved into two fractions indicating the presence of two isozymic forms which differ in biochemical properties (Lourenco and Neves, 1984). The major fraction had K_m values of $3.8x10^{-5}$ M and $1x10^{-5}$ M for shikimic acid and $NADP^+$ respectively. No activity was detected when $NADP^+$ was replaced with NAD^+. Among a number of inhibitors tested only protocatechuic acid functioned as a competitive inhibitor ($K_i = 7.7x10^{-5}$M) and several divalent metal ions (Hg, Zn and Cu) were also found to be inhibitory. SKDH purified from cucumber pulp also contained two isozymic forms and the biochemical properties were very similar to the properties observed for the tomato enzyme (Lourenco *et al.*, 1991).

In Ponderosa pine needles SKDH was a monomer and there were three allozymes, produced by three different alleles, in the two populations surveyed (Linhart *et al.*, 1981). Any given individual tree had a phenotype of one or at most two of the three allozymes.

Plant isozymes have a wide application as genetic markers in agriculture. SKDH isozyme systems have been used in cultivar identification and breeding studies of a variety of agronomically important plant species such as soybean (Yu and Kiang, 1993), banana (Bhat *et al.*, 1992), lychees (Degani *et al.*, 1995) and strawberry (Williamson *et al.*, 1995).

1.2.2.5 Quinate-shikimate dehydrogenase

This enzyme is distinct from the biosynthetic SKDH and catalyses the first step in the catabolic quinate pathway; the oxidation of quinate and shikimate to 3-DHQ and 3-DHS respectively. In both prokaryotes and eukaryotes quinate and shikimate are converted to protocatechuate by the quinate pathway, which is then metabolised by the β-ketoadipate pathway to succinate and acetyl Co-A (Tresguerres *et al.*, 1970). Quinate -shikimate dehydrogenase has been purified and characterised from *N. crassa* (Barea and Giles, 1978), *A. nidulans* (Hawkins *et al.*, 1993) and from the Gram-positive bacterium *Rhodococcus rhodochrous* (Bruce and Cain, 1990). Its existence has been demonstrated in *Aspergillus niger* (Cain, 1972) and the gene encoding this enzyme in *Acinetobacter calcoaceticus* has recently been sequenced (Elsemore and Ornston, 1994).

The coenzyme specificity of catabolic quinate-shikimate dehydrogenases show variation among different organisms investigated unlike the biosynthetic SKDH which is $NADP^+$ specific. The fungal enzymes and the enzyme from *R. rhodochrous* are NAD^+ specific (Cain, 1972; Bentley, 1990; Bruce and Cain, 1990), but the enzyme from *A. calcoaceticus* is reported to be pyrrolo-quinoline quinone (PQQ) dependent (Elsemore and Ornston, 1994).

1.3 Objectives of the project

Among the shikimate pathway enzymes, SKDH is one enzyme which has not been subjected to extensive study. As a result of the sequencing and overexpression of *E. coli aro E* gene, the amino acid sequence and relatively large amounts of the *E. coli* enzyme had been made available. This has facilitated detailed studies on the structure and mechanism of this enzyme. As a part of this study, the main objective of my Ph.D. project was to study the active site of *E. coli* SKDH.

The practical aims were:

(1) To identify residues involved in substrate and coenzyme binding using chemical and biophysical methods (group specific chemical modification, electrospray mass spectrometry

and HPLC) and to further investigate their interactions with the enzyme using substrate analogues.

(2) To identify the proposed histidine residue involved in the catalytic mechanism as a general acid/base facilitating substrate reduction/oxidation.

(3) To attempt crystallisation of SKDH as a preliminary step towards solving the three dimensional structure.

Finally, it is taken into consideration that this enzyme is a potential target for antimicrobial and herbicide action. Knowledge of the active site and the identification of enzyme-substrate interactions could lead to the design of such potential inhibitors.

CHAPTER 2

Methodology

2.1 Group specific chemical modification

Chemical modification is of special importance in probing active-site structure of enzymes. Of the 20 natural amino acids, only those possessing a polar side chain are normally the object of chemical modification. If the reaction of an enzyme with a group specific reagent results in the inactivation of the enzyme, it may indicate that an active site residue involved in catalysis or binding is modified (Means and Feeney, 1971).

However, there are several limitations to this technique (Eyzaguirre, 1987). Most of the commonly used group specific reagents are not absolutely specific for a certain amino acid side-chain. The specificity of a reagent can be influenced by experimental variables such as pH and buffers. Furthermore, an enzyme may posses many residues of a certain amino acid, which can potentially react with a group specific reagent. Therefore, absolute specificity for one single amino acid is difficult to achieve. However, active site residues often show markedly different pK_a values from the free amino acids, and are much more reactive. This increases their selectivity towards modifying reagents. When more than one residue is modified differential labelling (modification first in the absence and then in the presence of substrate) can be used to identify active site residues (Means and Feeney, 1971). In the presence of substrate vulnerable residues at or near the active site may no longer be accessible to the modifying reagent and therefore will be protected from modification.

Interpretation of chemical modification data needs considerable caution. Inactivation of an enzyme following chemical modification could result from a conformational change in the enzyme. Another possibility is steric hindrance at the active site due to the presence of the modifying group, specially if bulky reagents are used. Thus, although chemical modification is an effective tool, it is necessary to confirm the function assigned to the modified residue by other means. Site-directed mutagenesis is one approach, but the protein crystal structure would be the ultimate means by which the active site structure is confirmed.

2.2 Electrospray Mass Spectrometry

Mass spectrometry is an indispensable analytical tool that has been used for some time in biochemical research for the analysis of thermally stable small molecules; large or thermally labile molecules did not survive the desorption and ionization processes intact. This problem had been solved recently with the introduction of electrospray mass spectrometry (ESMS) (Fenn *et al.*, 1989) and matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) (Karas and Hillenkamp, 1988), thereby allowing mass spectrometry to take on new roles in analysis of proteins, peptides and other biopolymers.

ESMS is a sensitive and an efficient technique. It offers picomole-to-femtomole sensitivity, and can be used to measure molecular masses up to >200 kDa with an accuracy in the order of $\pm$ 0.01% (Siuzdak, 1994). It can be used to monitor events such as phosphorylation, glycosylation, oxidation, disulphide bond formation and protein folding (Carr *et al.*, 1991). Moreover, it can be used to monitor chemical modification of proteins (Bordier *et al.*, 1992; Krell *et al.*, 1995) since the resolving power of ESMS is sufficiently high to distinguish between modified and unmodified protein molecules without pre-separation.

The basics of ESMS are shown in Fig. 2.1. The solution of analyte molecules is introduced through a silica capillary tube into a strong electrical field where it disperses into a fine spray of highly charged droplets. Dry gas, heat or both are applied to the highly charged droplets causing the solvent to evaporate. Evaporation causes the droplet size to decrease, while surface charge density increases. Multiply charged ions are transferred to the gas phase as a result of their expulsion from the droplet. In contrast to all other mass spectrometric ionization methods, the change from the liquid into gas phase is thought to occur very gently with at least one solvation shell surrounding the protein (Mann and Wilm, 1995). Ions are then directed into the quadrupole analyser through a vacuum region where their mass to charge ratios (m/z) are measured. Mass is determined by ascertaining the charged states of the peaks and a mass is calculated to each peak in the series; alternatively, the series is deconvoluted by a computer algorithm into a single peak (Mann *et al.*, 1989).

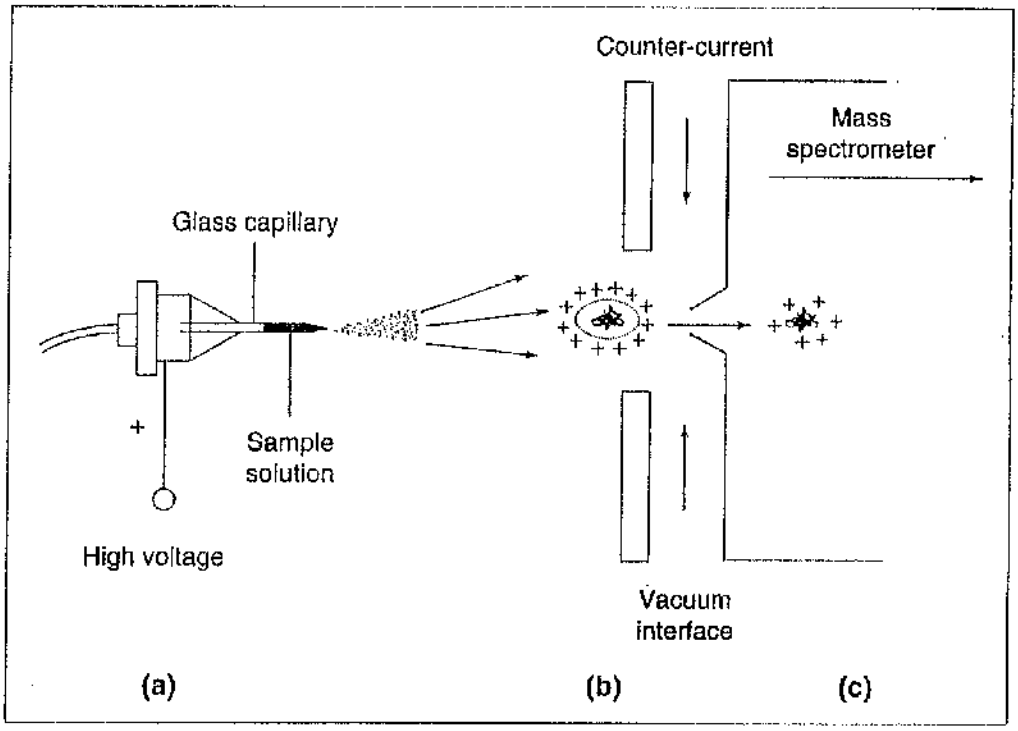


Fig. 2.1 The three principal steps of electrospray mass spectrometry. (a) Formation of small, highly charged droplets by electrostatic dispersion of a solution under the influence of a high electric field. (b) Desorption of protein ions from the droplets into the gas phase. (c) Mass analysis of the ions in a mass spectrometer.

The main technical limitation of ESMS is the selectivity in buffers (Mann and Wilm, 1995). Ionic buffers disturb the spraying process and compete with the analyte molecules for charges, thus such buffers should be avoided in ESMS work. Furthermore, involatile substances such as salts and detergents can disturb the desorption process by forming a solid core around the droplet from which the ions cannot escape. However, this problem can be overcome by desalting the sample prior to analysis.

2.3 HPLC Mass Spectrometry

Electrospray ionization demands a constant delivery of liquid and is therefore easily coupled to a liquid-based separation system such as HPLC (Covey *et al.*, 1991). This form of mass spectrometry is known as LCMS. The ability to directly analyse compounds from aqueous/organic solvents has made ESMS an excellent detector for HPLC.

LCMS has a wider application in the protein chemistry field. It is especially well suited for the analysis of proteolytic digests of proteins (Meyer *et al.*, 1993). During the HPLC separation of a proteolytic digest, peptides eluting from the HPLC column are directly introduced into the mass spectrometer and mass spectra are recorded every few seconds. All the ions detected are summed up in a total ion chromatogram and peptides are identified by their masses using a software package. However, the use of site-specific proteases for proteolysis is of importance in the correct identification of peptides.

LCMS has been successfully employed to locate sites of phosphorylation (Meyer *et al.*, 1993) and di-sulphide bond formation (Jaquinod *et al.*, 1993). Furthermore, it can be used to identify sites of group specific chemical modification (Krell *et al.*, 1995), thereby adding a new dimension to active-site studies of enzymes.

CHAPTER 3

Experimental procedures

3.1 Materials and reagents

3.1.1 Chemicals and biochemicals

Centricon-10 concentrators were from Amicon Ltd., Stonehouse, Gloucestershire, UK.

Ampicillin, benzamidine, phenylmethylsulfonyl fluoride (PMSF), shikimic acid, diethylpyrocarbonate (DEPC), hydroxylamine-hydrochloride ($NH_2OH.HCl$), phenylglyoxal (PGO), bromophenol blue, Coomasie brilliant blue, N,N,N',N' -tetramethylethylene diamine (TEMED) were obtained from Sigma Chemical Co., Poole, Dorset, UK.

Bactotryptone and yeast extract were obtained from Difco, Detroit, Michigan, USA.

Dithiothreitol (DTT), NAD^+, $NADP^+$ and Tris buffer were obtained from Boehringer Mannheim, Lewes, Sussex, UK.

Trinitrobenzene sulfonic acid (TNBS), imidazole and 2-mercaptoethanol were obtained from BDH Chemicals, Poole, Dorset, UK.

Acrylamide, bisacrylamide, SDS and analytical grade formic acid, were obtained from FSA Laboratory Supplies, Loughborough, Leicestershire, UK.

Ultrapure grade guanidium hydrochloride (GdnHCl), urea and isopropyl-β-D-thiogalactoside (IPTG) were obtained from BRL,Gibco Ltd., Paisley, Scotland, UK.

HPLC grade water, acetonitrile and trifluoroacetic acid (TFA) were obtained from Rathburn Chemicals Ltd., Walkerburn, Scotland, UK. HPLC grade sodium phosphate was from BDH Chemicals.

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

3.1.2. Enzymes and proteins

Bovine serum albumin (BSA), deoxyribonuclease I, subtilisin and a molecular weight standard marker kit were obtained from Sigma Chemical Co.

TPCK treated bovine trypsin and TLCK treated bovine α-chymotrypsin were obtained from Worthington Biochemical Corp., Freehold, New Jersey, USA.

3.1.3. Chromatography media

DEAE-Sephacel, Sephacryl S-200 (superfine grade), 2'5' ADP-Sepharose 4B and Sephadex G-25 were obtained from Pharmacia, Milton Keynes, Buckinghamshire, UK.

3.1.4. Pre-packed media

A pre-packed Mono Q column was obtained from Pharmacia and utilised on a Pharmacia FPLC System. Vydac C4 reverse phase columns were purchased from Sigma-Aldrich Chromatography, Poole, Dorset, UK. and attached to a Shimadzu HPLC system (Shimadzu Corp., Kyoto, Japan). A narrow bore Delta-Pak™ HPI C4 reverse phase column was purchased from Waters Chromatography, Watford, Hertfordshire, UK. and used on a LCMS system (Shimadzu HPLC apparatus coupled to a VG BioTech Platform single quadrupole mass spectrometer from VG BioTech Ltd., Altrincham, Cheshire, UK.).

3.2 General laboratory methods

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)

3.2.1 pH measurement

pH measurements were made with a Radiometer Model 26 pH meter (Copenhagen, Denmark), calibrated at room temperature.

3.2.2. Protein estimation

Protein concentrations were estimated by the method of Bradford (1976), with BSA as a standard.

3.2.3 Lyophilization

Protein and peptide solutions were collected either into Reacti Vials (Pierce), polypropylene tubes (eppendorfs) or 25-100 ml acid washed round bottom flasks and the contents frozen by immersing the vessels in a dry-ice/ethanol mixture before lyophilization on a FTS Systems (Stone Ridge,New York, USA.) Flexi-Dry freeze dryer.

3.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel electrophoresis was performed by the method of Laemmli (1970), with a 5% stacking gel and a 15% running gel. The ratio of acrylamide : bisacrylamide in all PAGE experiments was 30 ; 0.8 and polymerisation was induced by the addition of 0.03% (v/v) TEMED and 0.05% (w/v) ammonium persulphate. After electrophoresis gels were stained for protein by the Coomassie method (Section 3.3.1).

3.3.1 Staining for protein

Gels were stained with Coomassie blue for 30-45 minutes at 40°C for visualising of protein. The Coomassie reagent was 0.1% (w/w) Coomassie brilliant blue G250 in 50% (v/v) methanol, 10% glacial acetic acid ; destaining was carried out in 10% methanol (v/v), 10% glacial acetic acid at 40°C until the background was fully destained and bands clearly visible.

3.4 SKDH assay

Although the reaction in the direction of shikimic acid is the one that takes place biosynthetically 3-dehydroshikimate and NADPH are less readily available reagents than shikimic acid and $NADP^+$. Therefore, the reaction was followed in the reverse direction using shikimic acid and $NADP^+$ according to the method of Balinsky and Davies

(1961a). SKDH activity was assayed by monitoring the reduction of $NADP^+$ at 340 nm ($\varepsilon = 6.2 \times 10^3\ M^{-1}cm^{-1}$) (Horecker and Kornberg, 1948) during the oxidation of shikimate to 3-dehydroshikimate at 25°C. The standard assay mixture in a total volume of 1ml, contained a final concentration of 2 mM $NADP^+$ (coenzyme), 4 mM shikimate (substrate) and 100 mM Na_2CO_3 pH 10.6. All assays for chemical modification were performed using a Gilford/Unicam model 252 UV/Vis spectrophotometer equipped with an external slave recorder.

3.5 Cell culture and enzyme purification

3.5.1 Growth of *E.coli* AB2834/pIA321

Cells of *E. coli* strain AB2834/pIA321 were grown in L-broth at 37°C on a orbital shaker. L-broth containing 10 g/litre Bactotryptone, 5 g/litre yeast extract and 10 g/litre NaCl was sterilised by autoclaving at 15 psi. D-glucose and ampicillin were added to a final concentration of 0.1% (w/v) and 100 mg/litre respectively prior to inoculation. At first cells were grown in small feeder cultures (3x100 ml of media in 250 ml conical flasks) for approximately 20 hr. These feeder cultures were used to inoculate each of 8x500 ml of fresh media in 2 litre conical flasks. One hour after inoculation isopropyl 1-thio-β-D-galactoside (IPTG) was added to a final concentration of 0.5 mM. Incubation was continued until growth ceased (A_{595} of approximately 5.2) and the cells were then harvested by centrifugation using a Beckman 6L centrifuge (4500 rpm for 15 minutes). Cells were stored at -20°C. A typical 4 litre culture gave 26 g wet weight of cells.

3.5.2 Purification of SKDH

Shikimate dehydrogenase was purified to homogeneity from the *E. coli* AB2834/pIA321 overproducing strain according to the procedure of Chaudhuri *et al.* (1987), except that a final step of chromatography on a Mono Q column using a Pharmacia FPLC system was included as in the procedure described for the purification of wild type enzyme by Chaudhuri and Coggins (1985).

The specific activity of the purified enzyme was in the range of 950-1050 units/mg.

3.5.3 Determination of SKDH concentration

The extinction coefficient of purified SKDH was determined by amino acid analysis, and was found to be 0.72 (mg/ml) at 280 nm. This value was used in the quantification of SKDH in all experiments.

3.6 Characterisation of TNBS mediated inactivation of SKDH

3.6.1 TNBS inactivation reaction

Inactivation was carried out by incubating SKDH (2-3 μM monomer concentration) in the dark with different concentrations of TNBS in 50 mM borate/NaOH buffer pH 9.2 at 25°C. Aliquots (10 μl) were withdrawn from the reaction mixture (1 ml) at various time intervals and assayed for residual activity as described in section 3.4.

3.6.2 Substrate protection against TNBS inactivation

Protection experiments were carried out by incubating SKDH with increasing concentrations of $NADP^+$ (coenzyme) or shikimic acid (substrate) prior to addition of TNBS. In a separate experiment SKDH was incubated with a mixture of both $NADP^+$ and shikimic acid together in a final concentration of 0.5 mM each prior to addition of TNBS. Aliquots were removed at time intervals and were assayed for residual enzyme activity. All other conditions were as described in section 3.6.1.

3.6.3 Absorption spectrum of TNBS modified SKDH

SKDH was 90% inactivated by treatment with 35-fold molar excess of TNBS in 50 mM borate/NaOH buffer, pH 9.2 at 25°C. The excess reagent was removed by dialysis against several changes of buffer. Absorption spectrum of the trinitrophenylated enzyme was recorded over the wavelength range 280nm-500nm by using a double beam spectrophotometer (Phillips PU 8700) with a reference cell containing the enzyme which had been carried through the same treatment except that TNBS was omitted.

3.6.4 Preparation of TNBS modified SKDH

SKDH (0.3mg) was modified to different extents by treatment with 5-fold and 30-fold molar excess of TNBS over protein monomer in 50 mM borate/NaOH buffer pH 9.2 at 25°C. After one hour of incubation the sample treated with 5-fold molar excess TNBS was 82% active and >75% activity was lost in the sample treated with 30-fold molar excess TNBS.

In a parallel experiment SKDH was modified in the presence of substrate and coenzyme. SKDH (0.3mg) was first incubated with 2mM $NADP^+$ and 2mM shikimic acid and this was followed by the addition of 30-fold molar excess of TNBS. After one hour of incubation 86% of the activity remained in the protected sample.

In all three cases the reactions were stopped by gel filtration on Sephadex G-25, equilibrated in 10 mM ammonium bicarbonate pH 9.0. Fractions containing enzyme activity were pooled and concentrated using centricon-10 (Amicon) cells by centrifuging at 5000xg. In order to remove low molecular weight contaminants all three samples were washed thoroughly with HPLC grade water by centrifugation using centricon cells. Samples were stored at -20°C and used for electrospray mass spectrometry, reverse phase HPLC and on-line LCMS.

3.6.5 Electrospray mass spectrometry (ESMS) and stoichiometry of labelling

Stoichiometry of labelling of the samples described in section 3.6.4 was determined using ESMS. Samples were analysed on a VG BioTech platform single quadrupole mass spectrometer (2-3000 amu range) fitted with a pneumatically assisted electrospray source and controlled via the VG Masslynx software. Carrier solvent [1:1(v/v) acetonitrile:water with formic acid added to a final concentration of 0.2%(v/v)] infusion was controlled at 10 µl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, Mass., USA.). Samples were dissolved in carrier solvent at a concentration of 20 pmol/µl, centrifuged at 5000xg for 2 min and then 10-20µl samples were injected. The MaxEnt deconvolution procedure (Ferridge *et al.*, 1992) was applied for quantitative

analysis of raw data using 1.0 Da peak width and 1.0 Da channel resolution.

3.6.6 Proteolysis of TNBS modified SKDH

150 µg from each sample (described in section 3.6.4) were digested with chymotrypsin prior to analysis by reverse phase HPLC and LCMS. First the samples were denatured by adding one volume of 0.5% ammonium bicarbonate buffer pH 8.0 containing 8M GdnHCl, and subsequent incubation for 15 min at 30°C. The solution was then diluted with 0.5% ammonium bicarbonate buffer pH 8.0, to a final GdnHCl concentration of 2M. Chymotrypsin (5% of the mass of substrate) was added at this stage. The digests were performed for one hour at 30°C and stopped by freezing the samples in aliquots (2x75 µg) in dry ice/ethanol. All digests were stored at -70°C.

3.6.7 Reverse phase HPLC of peptides

Chymotryptic digests of SKDH modified with 30-fold molar excess TNBS in the absence and presence of $NADP^+$ and shikimate were fractionated using a Vydac 214TP reverse phase C4 column, (4.6 x250 mm) (Sigma-Aldrich) equilibrated in buffer A [0.1% trifluoroacetic acid in 2:98 (v/v) acetonitrile:water]. 75 µg from each sample were injected and the peptides were eluted with a split gradient of buffer B [0.1% trifluoroacetic acid in 70:30 (v/v) acetonitrile:water] increasing to 70% B in 45 minutes and to 100% B in 50 minutes, with a flow rate of 1 ml/min. The peptides were monitored spectrophotometrically at three different wavelengths simultaneously using a photodiode array detector (Shimadzu). Peptides were detected by absorbance at 214 nm and TNBS modified peptides were detected by their absorbance at 346 nm and 420 nm.

3.6.8 LCMS

LCMS of the chymotryptic digests of TNBS modified SKDH [82% active sample, >75% inactive sample(unprotected) and 86% active sample(protected)] were carried out by coupling a Shimadzu HPLC system to VG Biotech platform single quadrupole mass spectrometer, using a Delta-Pak HPI C4 column (2.0x150 mm)(Waters)

and the solvent system mentioned in section 3.6.7. The flow rate applied was 0.3 ml/min, and 75 µg from each sample were used for analysis. After injecting the sample the column was washed for 8 min to remove GdnHCl using buffer A. The column was developed using a split gradient of buffer B, increasing to 70% in 45 minutes and to 100% in 50 minutes. Peptides eluted were directly introduced into the mass spectrometer with the drying gas flow maximised and the source temperature set at 100°C. The absorption profile of the eluted peptides were recorded at 214 nm and mass spectra were recorded at 4 sec intervals.

In a control experiment a chymotryptic digest of unmodified SKDH (native), digested as described in section 3.6.6, was run under identical conditions as those used for TNBS modified samples and spectra were recorded.

3.6.9 Identification of peptides

All the peptides expected to be generated in a chymotryptic digest of SKDH and their masses were predicted from the known amino acid sequence data using a program in VGMassLynx software. Peptides were identified by the mass/charge values generated by this program. Peptides containing modified residues were identified by the increase in mass due to the attachment of trinitrophenyl groups.

3.7 Characterisation of PGO mediated inactivation of SKDH

3.7.1 PGO inactivation reaction

SKDH (2-3 µM monomer concentration) was incubated with different concentrations of PGO in 100 mM sodium bicarbonate-carbonate buffer, pH 9.4 at 25°C. Aliquots (10 µl) were removed from the reaction mixture (1 ml) at different time intervals and were assayed for residual enzyme activity as described on section 3.4.

3.7.2 Substrate protection against PGO inactivation

SKDH was incubated with $NADP^+$, with shikimate or with a combination of both $NADP^+$ and shikimate prior to the addition of PGO. The concentrations used were 0.2 mM $NADP^+$, 0.4 mM shikimate and 2.5 mM PGO. All other conditions were as described in section 3.7.1.

3.7.3 Preparation of PGO modified SKDH

SKDH (0.3 mg) in 100 mM sodium bicarbonate-carbonate buffer, pH 9.4 at 25°C was modified to different extents by incubating with 0.5 mM and 1.5 mM final concentration of PGO (freshly made up 50 mM stock solution in water). Inactivation was monitored by removing aliquots from the reaction mixture for enzyme assay. After 30 min of incubation, samples were 90% and 60% active respectively. In parallel experiments SKDH (0.3 mg) was treated with a final concentration of 3.0 mM PGO in the presence and absence of 2.0 mM $NADP^+$. The sample incubated in the absence of $NADP^+$ was approximately 35% active and in the presence of $NADP^+$ 93% of the activity was retained after 30 min.

In each case reaction was stopped by gel filtration on Sephadex G-25, equilibrated with 10 mM ammonium bicarbonate pH 9.0. To remove low molecular weight contaminants the samples were washed by diluting with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators. Samples were stored at -20°C and used for electrospray mass spectrometry, reverse phase HPLC and on-line LCMS.

3.7.4 Stoichiometry of incorporation of PGO

The stoichiometry of incorporation of PGO into SKDH was determined by ESMS using an aliquot from each sample prepared in section 3.7.3. ESMS was performed as described in section 3.6.5.

3.7.5 Proteolysis of PGO modified SKDH

Tryptic digests of PGO modified SKDH (samples described in section 3.7.3) were prepared using 160 µg from each sample. First, the samples were denatured in 0.5% ammonium bicarbonate (pH 8.0) containing 8M urea at 37°C for one hour. Samples were diluted to a final urea concentration of 2M with buffer and trypsin (5% of the mass of substrate) was added immediately. The digests were performed for 4 hr at 37°C with constant stirring. This was followed by a second addition of trypsin (2% of the mass of substrate) and digestion was continued for a further 2 hr in order to obtain a complete tryptic digest. Digests were frozen in dry ice/ethanol in aliquots (2x80 µg) and were stored at -70°C prior to analysis by reverse phase HPLC or LCMS.

3.7.6 Reverse phase HPLC of tryptic digests of PGO modified SKDH in the absence of presence of $NADP^+$

Tryptic digests of SKDH modified with 3mM PGO in the absence of $NADP^+$ (35% active sample) and in the presence of $NADP^+$ (93% active sample) were fractionated using a Vydac 214TP reverse phase column (4.6x250 mm) (Sigma-Aldrich). Identical quantities of protein (80 µg) from each sample was injected. Solvent system and all other conditions for elution of peptides were as described in section 3.6.7 except that the peptide maps were monitored at 214 nm, the usual wavelength for peptide detection.

3.7.7 LCMS

Tryptic digests of SKDH modified with 3.0 mM PGO in the absence and presence of $NADP^+$ (different aliquots of the samples analysed by RP-HPLC in section 3.7.6) were analysed by LCMS using a Delta-Pak HPI C4 column (2.0x150 mm) (Waters). 80 µg of sample was analysed in each case. Experimental details for LCMS are as described in section 3.6.8. In a control experiment a tryptic digest of native SKDH was analysed under identical conditions.

3.7.8 Identification of peptides

Masses of all tryptic peptides of SKDH were calculated from the known amino acid sequence using VGMassLynx software. Peptides were identified by mass/charge values generated by this program and PGO modified peptides were identified by the increase in mass.

3.8 Characterisation of DEPC mediated inactivation of SKDH

3.8.1 Determination of DEPC concentration

The concentration of commercially available DEPC may be variable owing to hydrolysis. Therefore the concentration of stock DEPC was quantitatively determined by reaction with 10 mM imidazole in 50 mM potassium phosphate buffer, pH 7.5. The increase in absorbance at 240 nm (ε=3200 M^{-1} cm^{-1}) (Miles, 1977) due to the formation of N-carbethoxy imidazole was read against a reference cell containing absolute alcohol and 10 mM imidazole in buffer. The reagent was freshly diluted with ice cold absolute alcohol before each experiment.

3.8.2 DEPC inactivation reaction

SKDH (2-3 μM monomer concentration) was incubated with DEPC in 50 mM sodium phosphate buffer, pH 7.0 at 25°C. The final concentration of ethanol in the reaction mixture (1 ml) was kept below 5% (v/v) so as not to affect enzyme activity. Aliquots (100 μl) were removed at time intervals from the reaction mixture and quenched with 200 μl 20 mM imidazole and diluted with 50 mM sodium phosphate buffer, pH 7.0 and then the residual enzyme activity was determined as described in section 3.4.

3.8.3 Substrate protection against DEPC inactivation

SKDH was incubated with $NADP^+$ or with shikimate or with a combination of both $NADP^+$ and shikimate prior to addition of DEPC. All other conditions were as described in section 3.8.2.

3.8.4 Carbethoxylation and decarbethoxylation

A double beam spectrophotometer (Philllips/PU8700) equipped with a thermostatically controlled cuvette holder and a data storage facility was used to determine the time dependent N-carbethoxylation of histidine residues by increase in the absorbance at 240 nm in 1 ml reaction volume in 50 mM sodium phosphate buffer, pH 7.0. Reference spectra of 1.0 mg of enzyme in buffer and ethanol were recorded over the wavelength range 220-350 nm and then subtracted from sample spectra recorded using 1.0 mg of enzyme in buffer and DEPC (1.0 mM) to generate difference spectra. The number of modified histidine residues as a function of time were calculated using the molar extinction coefficient for N-carbethoxyhistidine (ε_{240}= 3200 M^{-1} cm^{-1})(Miles, 1977).

Decarbethoxylation was carried out by adding an equal volume of neutral hydroxylamine to both reference and sample cuvettes when the enzyme was approximately 80% inactivated due to carbethoxylation. Spectra were recorded after one, two and three hours of addition of neutral hydroxylamine.

Carbethoxylation and decarbethoxylation were correlated to activity by assaying aliquots withdrawn from the reaction mixtures at various times during the experiment.

3.8.5 pH dependence of DEPC inactivation

The pH dependence of DEPC inactivation was determined in 50 mM sodium phosphate buffer over the pH range 5.5 - 7.1 at 25°C. All other conditions were as described in section 3.8.2.

3.8.6 Preparation of DEPC modified SKDH

SKDH (0.2 mg) was incubated with 30-fold molar excess of DEPC in 50 mM sodium phosphate buffer pH 7.0 at 25°C in the absence and presence of $NADP^+$ and shikimate (final concentration 5.0 mM each). After 30 minutes of incubation >70% activity was lost from the sample incubated in the absence of $NADP^+$ and shikimate (unprotected sample) whereas 65% of the activity retained in the protected sample.

Reactions were stopped by gel filtration on Sephadex G-25, equilibrated with 10 mM sodium phosphate buffer, pH 7.0. Samples were washed thoroughly with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentration cells. Samples were stored at -20°C.

3.8.7 Determination of stoichiometry of incorporation of DEPC

The stoichiometry of incorporation of DEPC into SKDH was determined by ESMS using an aliquot from each of the samples prepared in section 3.8.6. ESMS was performed as described in section 3.6.5.

3.8.8 Proteolysis of DEPC modified SKDH

100 µg each of DEPC modified SKDH in the absence and presence of $NADP^+$ and shikimate (samples prepared in section 3.8.6) were digested using subtilisin. The samples were denatured in 50 mM sodium phosphate buffer pH 7.0 containing 8M GdnHCl for 15 minutes at 30°C. The solution was then diluted to a final GdnHCl concentration of 2M with buffer. Subtilisin (5% of the mass of substrate) was added at this stage and the digests were performed for 2 hours at 30°C with continuous stirring. This was followed by immediate freezing of digests in dry ice/ ethanol and storage at -70°C prior to fractionation by RP-HPLC.

3.8.9 Reverse phase HPLC of peptides

The digests were fractionated using a Vydac 214TP reverse phase C4 column (4.6x250 mm) (Sigma) equilibrated in buffer A (10 mM sodium phosphate, pH 6.4). Identical quantities (100 µg) of protein were injected from each sample. The column was developed with a split gradient of buffer B (10 mM sodium phosphate, pH 6.4 in 70:30 acetonotrile:water) increasing to a 70% B in 45 minutes and then to 100% B in 50 minutes. Peptide maps were monitored at 214 nm which is the common wavelength for peptide detection and at 240 nm for the detection of DEPC modified peptides. Peptides protected from DEPC modification by substrate and coenzyme were collected manually and

freeze dried using acid washed glassware [Glass Reacti Vials (Pierce) soaked overnight in 6N HNO_3 and rinsed exhaustively with distilled water] and stored at -20°C.

3.8.10 Sequence analysis of isolated peptides

Isolated peptides were sequenced on an Applied Biosystems Model 470A Gas-phase sequencer with on-line detection of amino-acid thiohydantoins by a model 120A analyser. The instrument was operated by Dr. B. Dunbar at The Protein Sequencing Facility, University of Aberdeen.

CHAPTER 4

The essential lysine residue at the active-site of *Escherichia coli* shikimate dehydrogenase

4.1 Introduction

The epsilon amino group of lysine in it's unprotonated form is a highly reactive nucleophile and permits selective modification with chemical reagents. The average pK_a of this group is around 10 but this can be significantly altered by microenvironmental effects on the surface of the protein molecule as in the case of acetoacetate decarboxylase, where a value as low as 5.9 has been reported for a reactive lysine (Schmidt and Westheimer, 1971).

At the active sites of enzymes lysines may be functionally significant residues involved in catalysis or binding. Many cases have been reported where the positive charge of the epsilon amino group has mediated the binding of negatively charged substrates via charge interactions; in glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* Lys-21 interacts with the phosphate group of glucose-6-phosphate (Milhausen and Levy, 1975; Lee and Levy, 1992) while in human glutathione transferase P1-1, Lys-45 interacts with the glycine carboxylate of glutathione (Widersten *et al.*,1992). Furthermore in many pyridine nucleotide dependent enzymes lysine side chains have been implicated in coenzyme binding, such as Lys-21 and Lys-262 in human aldose reductase (Wilson *et al.*, 1992) and Lys-183 of lobster G3PDH (Buehner *et al.*, 1973). In class I aldolases and in type I dehydroquinases lysines play a catalytic role in reacting with carbonyl groups of substrates to form Schiff base intermediates (Horecker *et al.*, 1970; Chaudhuri *et al.*, 1991). In PLP dependent enzymes, the coenzyme binds in an imine linkage to the ε-amino group of an active-site lysine residue, thus converting the enzyme into an active aldimine form to facilitate catalysis (Braunstein, 1970).

In the studies leading to the characterisation of the active site of *E.coli* SKDH, involvement of a lysine residue in the mechanism of action was postulated. In preliminary experiments inactivation of SKDH with methyl benzimidate (MB) and with 2,4,6-trinitrobenzenesulfonic acid (TNBS), both lysine specific reagents, suggested the presence of an important lysine residue. The present chapter describes experiments leading to the identification of an essential lysine at the active site of this enzyme.

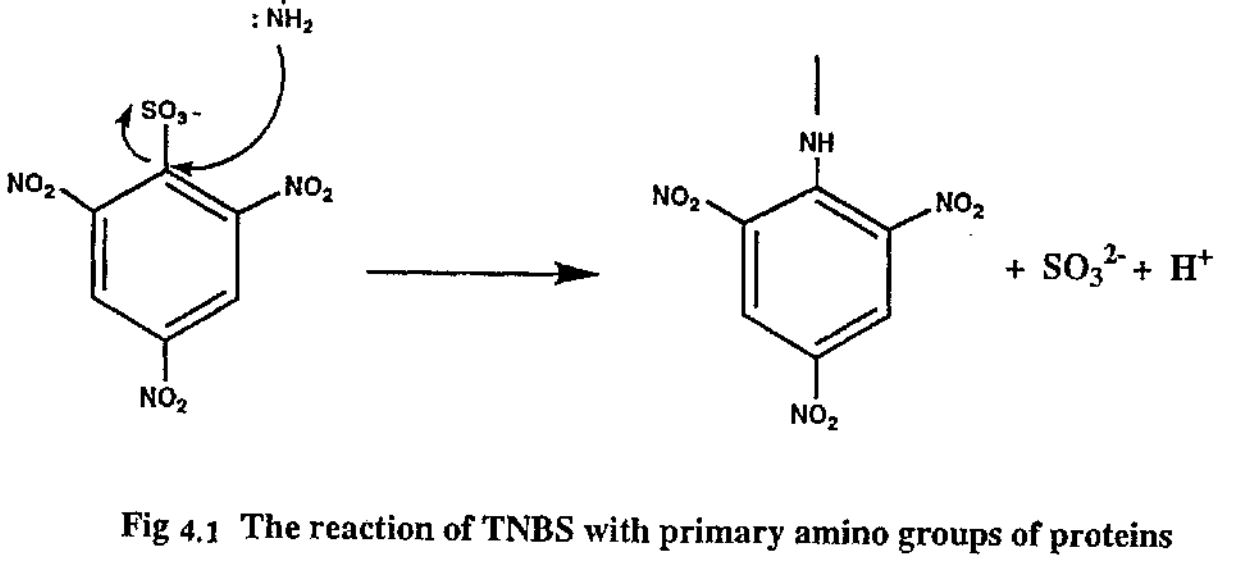


Fig 4.1 The reaction of TNBS with primary amino groups of proteins

4.2 Chemical modification with TNBS

4.2.1 Background

TNBS was originally introduced by Okuyama and Satake (1960) as a reagent to selectively react with primary amino groups of amino acids and proteins. In the absence of available primary amines it may react with sulfhydryl groups (Kotaki *et al.*, 1964) but no reaction has been detected with the hydroxyl groups of serine, threonine nor tyrosine nor with guanido or imidazole groups (Okuyama and Satake, 1960; Satake *et al.*, 1960). The reaction yields a trinitrophenyl (TNP)-amino derivative with the displacement of a sulphite ion from TNBS (Fig 4.1).

The ultraviolet absorption spectrum of this derivative shows a characteristic absorption maximum at 345-348 nm and a broad shoulder at 420 nm (Okuyama and Satake, 1960; Kotaki *et al.*, 1964). Sulfite ions formed during the reaction associate with TNP-amino groups to form complexes, altering the absorption spectrum by increasing it at 420 nm and decreasing at 346 nm. However in the presence of excess sulfite, absorbance of the sulfite complex measured at 420 nm ($\varepsilon = 2.0 \times 10^4\ M^{-1}\ cm^{-1}$) is the most sensitive index for quantification of the reaction (Lundblad and Noyes, 1984).

The reaction of TNBS with amines is strongly pH dependent, since only the unprotonated form reacts (Goldfarb, 1966; Freedman and Radda, 1968). Reactivity to TNBS is a sensitive measure of the basicity of the amino group which may be influenced by adjacent charged groups.

4.2.2 Kinetics of TNBS inactivation of SKDH

In pseudo first-order kinetics [I] >> [E] (I and E represents inhibitor and enzyme respectively) and the rate of the reaction (v) is given by,

$$v = k_{obs}\,[E] \quad \text{..... equation (1)}$$

which can be rewritten as,

$$-d[E_a] / dt = k_{obs}[E_a] \quad \text{..... equation (2)}$$

where $[E_a]$ is the enzyme concentration at time t and k_{obs} is the pseudo first-order rate constant. Integration of equation (2) with respect to $[E_a]$ between times 0 and t gives a relationship between the natural logarithm of the fractional activity of the enzyme and the time;

$$\ln[E_a]_t / [E_a]_0 = -k_{obs}\, t \quad \text{..... equation (3)}$$

and conversion to a common logarithmic form gives the following equation.

$$\log [E_a]_t /[E_a]_0 = -k_{obs}/2.303\, t \quad \text{..... equation (4)}$$

Therefore logarithm of the fraction of activity remaining against time gives a straight line in a pseudo first-order plot (Eyzaguirre, J., 1987).

Incubation of *E. coli* SKDH with TNBS at 25°C in 50 mM borate/NaOH buffer pH 9.2, resulted in a rapid loss of enzyme activity. In control experiments, in the absence of TNBS no loss of activity was detected. The fraction of activity remaining (A_t/A_0, where A_t is the activity at a given time and A_0 is the initial activity *i.e.* at 0 time) was calculated as the percentage of activity remaining at a given time (**t**). The plot of logarithm of percent remaining activity against time at various concentrations of TNBS showed pseudo first-order kinetics (Fig 4.2 A). The rate of inactivation was dependent on TNBS concentration.

The pseudo first-order rate constants (k_{obs}) for inactivation at various TNBS concentrations were calculated using the following expression:

$$k_{obs} = \ln 2/ t_{1/2} = 0.693/ t_{1/2} \quad \text{..... equation (5)}$$

where $t_{1/2}$ (half life) is the time after which the activity has decreased to half of its original activity, *i.e.* time when $A_t=A_0/2$. The half life of inactivation at various concentrations of TNBS were measured directly from the pseudo first-order plot. The results are summarized in Table 4.1.

Table 4.1

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

[TNBS] mM	$t_{(1/2)}$ min	k_{obs}/min
0.030	50.1	0.014
0.050	31.2	0.022
0.100	16.5	0.042
0.150	11.5	0.060
0.175	9.7	0.071

The relationship between the rate constant and the inhibitor concentration in pseudo first-order kinetics may be expressed as:

$$k_{obs} = k\,[I]^n \quad \text{..... equation(6)}$$

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

The plot of pseudo first-order rate constants (k_{obs}) as a function of TNBS concentration (Fig. 4.2 B) yielded a straight line passing through the origin which indicated

Fig. 4.2 Kinetics of inactivation of SKDH with TNBS

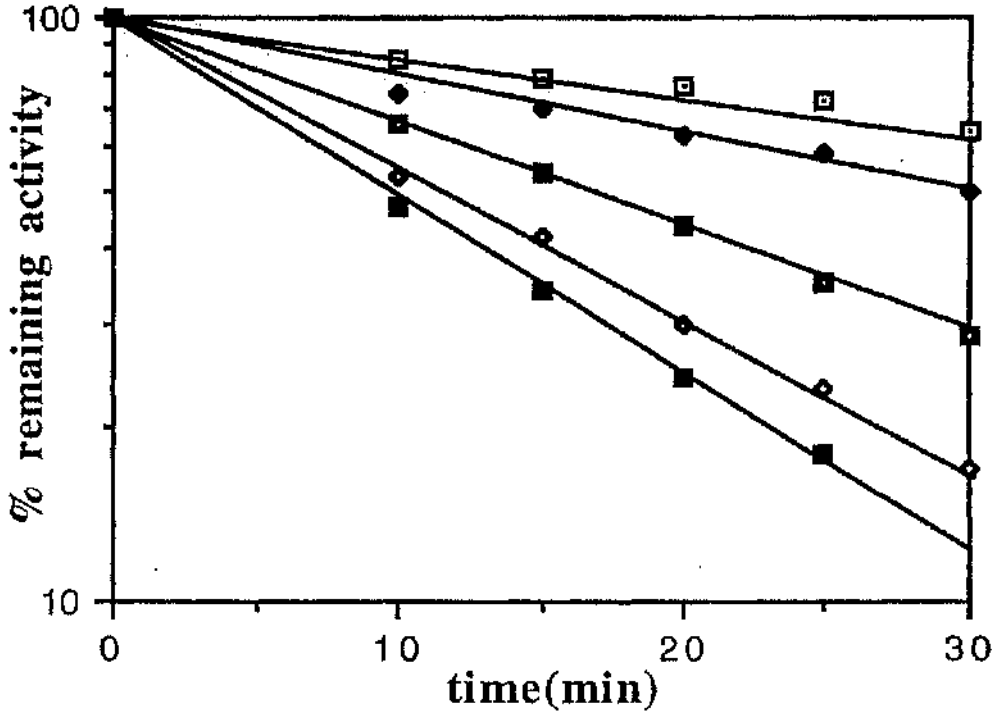
Fig. 4.2 A. Pseudo first-order plots for inactivation

SKDH (2-3 μM) was incubated with increasing concentrations of TNBS in 50 mM borate/NaOH buffer, pH 9.2 at 25°C. Aliquots were removed at time intervals and assayed for residual activity as described in section 3.4.The concentrations of TNBS used were 0.03 mM(▣), 0.05 mM(◆), 0.10 mM(◘), 0.15 mM(◈) and 0.175 mM(■).

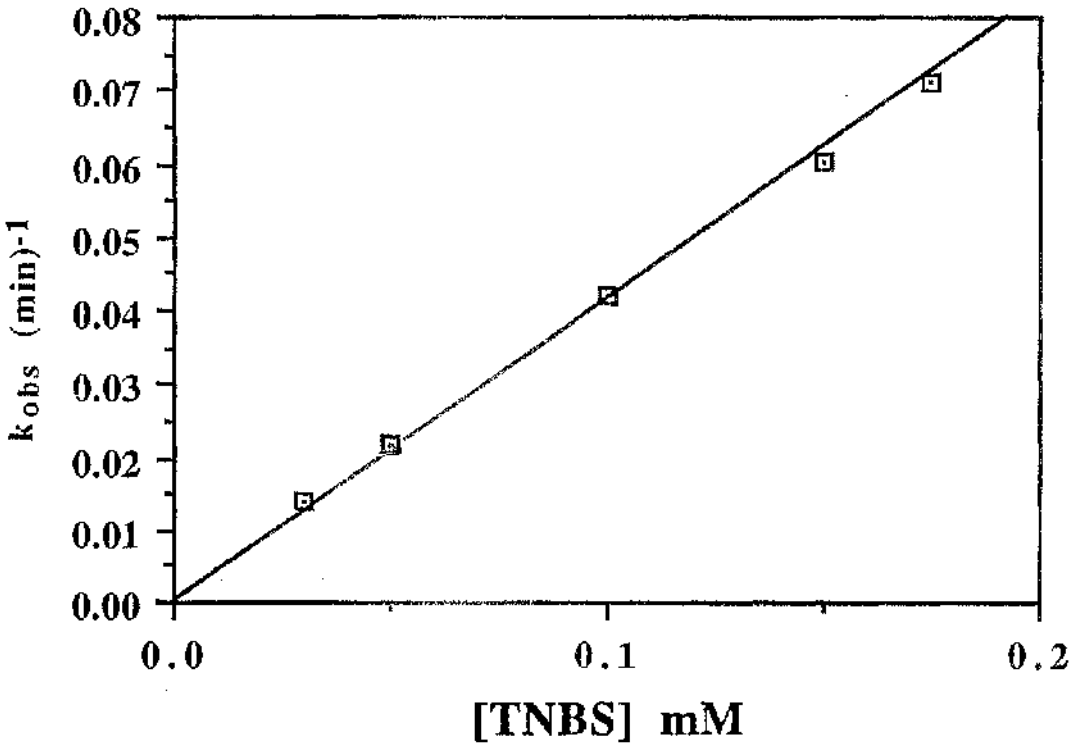
Fig. 4.2 B. Determination of the second-order rate constant of inactivation

Pseudo first-order rate constants (k_{obs}) calculated in part A were replotted against TNBS concentration. The second-order rate constant (k) calculated from the slope of this plot was 405 $M^{-1}min^{-1}$.

A


100
10
% remaining activity
0
10
20
30
time(min)

B


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
k_{obs} (min)-1
0.0
0.1
0.2
[TNBS] mM

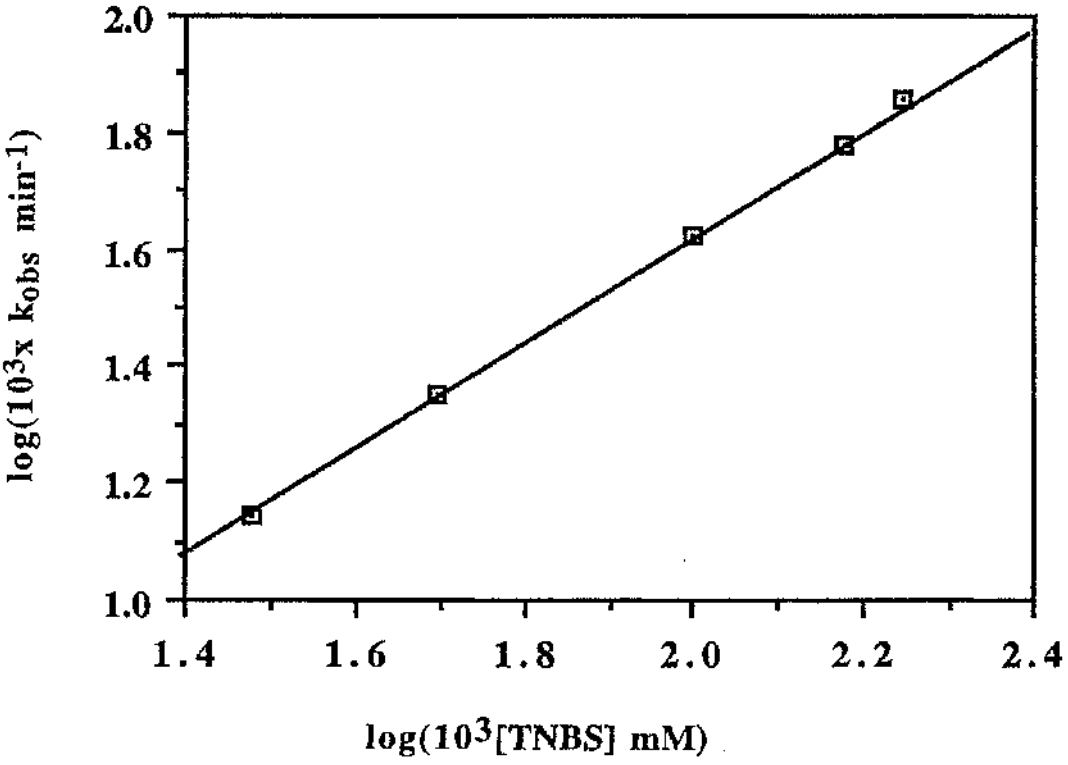


Fig. 4.2 C. Determination of the order of the reaction

The order of reaction with respect to TNBS calculated from the slope of this plot was 1.07, which indicates that TNBS mediated inactivation of SKDH results from the modification of a single lysine residue.

that the inactivation is a bimolecular process and a dissociable enzyme-TNBS complex was not formed prior to inactivation (Kitz and Wilson, 1962; Church *et al.,* 1985). The second-order rate constant (k) calculated from the slope was 405 M^{-1} min^{-1}.

The kinetic order of the reaction can be estimated from the logarithmic form of equation (6):

$$\log k_{obs} = n \log [I] + \log k \quad \text{..... equation (7)}$$

where n is the kinetic order of the reaction or minimal number of reagent molecules needed to inactivate a single molecule of active enzyme unit (Levy *et al.*, 1963). A plot of log k_{obs} against log [TNBS] yielded a straight line with a slope near unity (1.07) (Fig. 4.2 C).This indicates that the TNBS mediated inactivation of SKDH results from the modification of a single lysine residue.

4.2.3 Substrate protection against TNBS inactivation

Retention of enzyme activity following modification in the presence of substrate may indicate that inactivation is active site directed (Means and Feeney, 1971). In the presence of substrate vulnerable residues at or near the active site may no longer be accessible to the modifying reagent and provides a criterion for the identification of active site residues.

SKDH activity could be protected against TNBS inactivation by prior addition of $NADP^+$ (coenzyme) (Fig.4.3 A) or shikimate (substrate) (Fig.4.3 C). In both cases the % protection offered increased with concentration as illustrated in Tables 4.2 and 4.3 respectively. % protection was calculated by,

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

Protection occurred at low concentrations of $NADP^+$ and higher concentrations of shikimate. The Ks values for $NADP^+$ and shikimate at pH 9.2 determined from these data are 70 μM (Fig.4.3 B) and 240 μM (Fig.4.3 D) respectively. These data indicate that $NADP^+$ binds more tightly to the free enzyme than shikimate. However maximum protection occurred in the presence of both $NADP^+$ and shikimate (Fig. 4.3 E) suggesting that the ternary complex is more stable than the binary complexes. It is interesting that shikimate alone does afford some protection, which is enhanced in the presence of $NADP^+$. This may indicate that the proper binding of shikimate to the enzyme occurs in the presence of $NADP^+$.

Table 4.2

Effect of $NADP^+$ concentration on SKDH inactivation by TNBS

[$NADP^+$]	$t_{1/2}$ (min)	k_{obs}/min	% Protection
0.0 mM	**5.2**	**0.133**	**0**
0.1 mM	**13.2**	**0.052**	**61**
0.2 mM	**19.0**	**0.036**	**73**
0.3 mM	**30.0**	**0.023**	**83**
0.4 mM	**36.0**	**0.019**	**86**

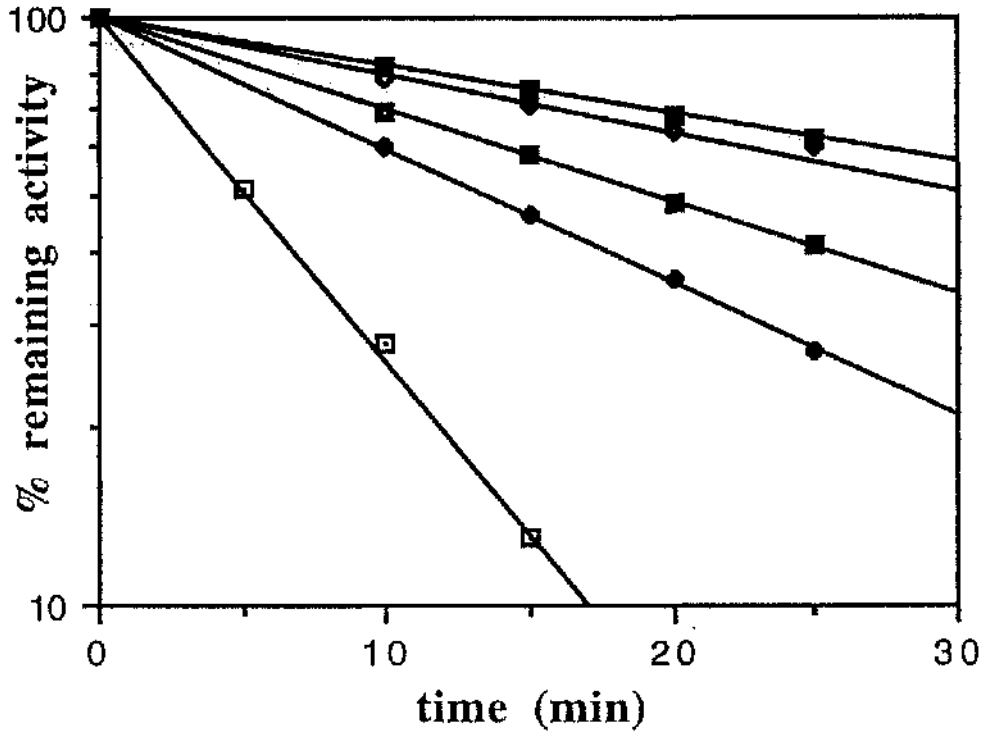
Fig. 4.3 A. Protection with $NADP^+$ against TNBS inactivation

SKDH (2-3 μM) was incubated with 0.25 mM TNBS in the presence of increasing concentrations of $NADP^+$. Aliquots were withdrawn at time intervals and were assayed for enzyme activity. $NADP^+$ concentrations used were 0mM (□), 0.1mM(◆), 0.2mM(□), 0.3mM(◆) and 0.4mM(■).

Fig. 4.3 B. Determination of Ks for $NADP^+$

The half life of values calculated from part (A) were plotted against $NADP^+$ concentrations. The Ks value for $NADP^+$ determined from the intercept of this plot is 70 μM.

A


100
10
% remaining activity
0
10
20
30
time (min)

B

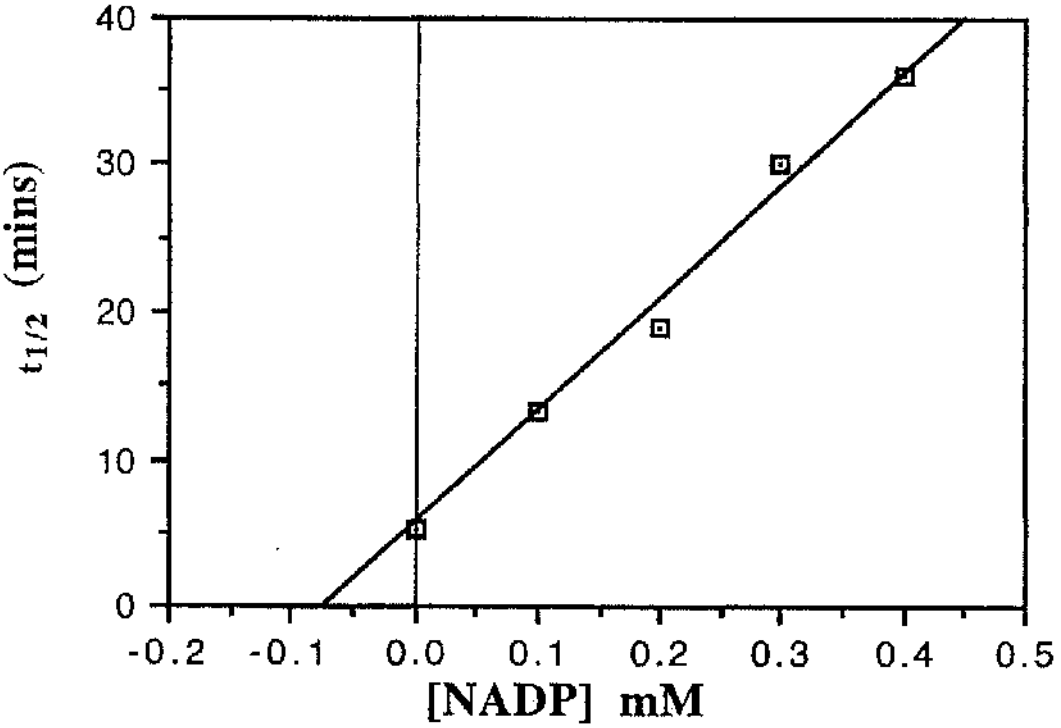

40
30
20
10
0
$t_{1/2}$ (mins)
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
0.5
[NADP] mM

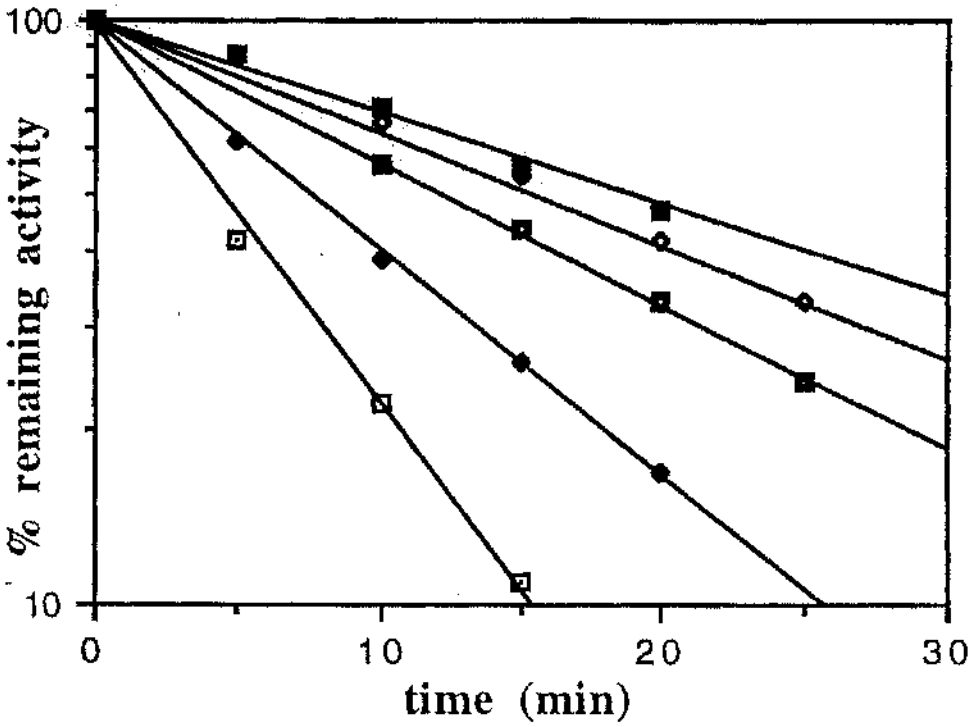
Fig. 4.3 C. Protection with shikimate against TNBS inactivation

SKDH (2-3 μM) was incubated with 0.25 mM TNBS in the presence of increasing concentrations of shikimate. Aliquots were withdrawn at time intervals and were assayed for enzyme activity. The concentrations of shikimate used were 0mM(□), 0.2mM(◆), 0.4mM(▣), 0.6mM(◇) and 0.8mM(■).

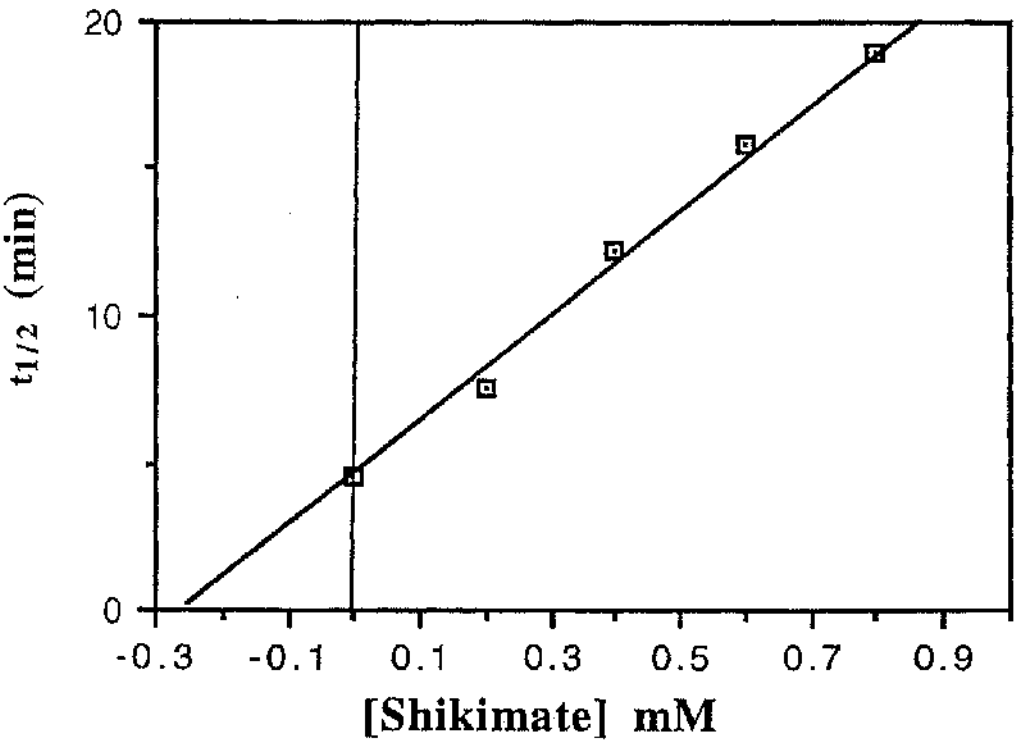
Fig. 4.3 D. Determination of Ks for shikimate

The half-life values calculated from part (A) were plotted against shikimate concentrations. The value of Ks for shikimate determined from the intercept of this plot is 240 μM.

C


100
10
% remaining activity
0
10
20
30
time (min)

D


20
10
0
$t_{1/2}$ (min)
-0.3
-0.1
0.1
0.3
0.5
0.7
0.9
[Shikimate] mM

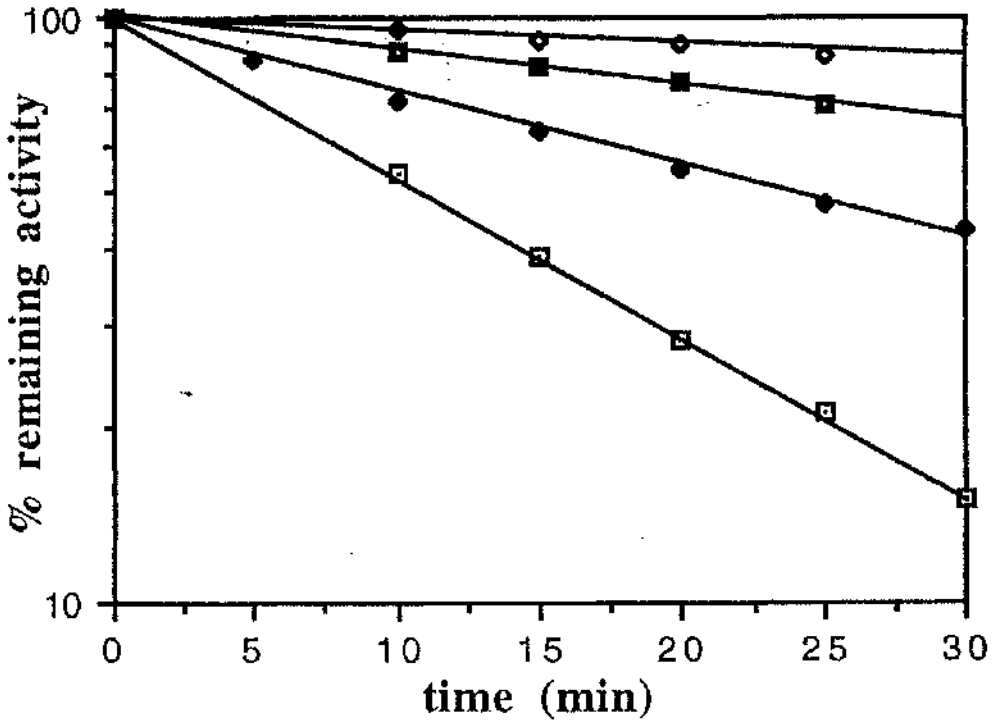


Fig. 4.3 E. Protection against TNBS inactivation in the presence of a mixture of $NADP^+$ and shikimate

SKDH (2-3 μM) was incubated with 0.175mM TNBS alone(▫), in the presence of 0.5mM shikimate (◆), 0.5 mM $NADP^+$ (▪) and a mixture of 0.5mM $NADP^+$ and shikimate (◇). Pseudo first-order rate plots shown in this figure indicates that maximum protection against TNBS inactivation is obtained in the presence of both NADP+ and shikimate.

Table 4.3

Effect of shikimate concentration on SKDH inactivation by TNBS

[Shikimate]	$t_{1/2}$(min)	k_{obs}/min	% Protection
0.0 mM	4.6	0.151	0
0.2 mM	7.5	0.092	39
0.4 mM	12.2	0.057	62
0.6 mM	15.8	0.044	71
0.8 mM	19.0	0.036	76

4.2.4 Characterization of TNBS-modified SKDH

TNBS preferentially reacts with primary amino groups of proteins but it may also react with sulfhydryl groups forming S-TNP derivatives (Kotaki *et al.*, 1964). There are three cysteine residues in the amino acid sequence of SKDH and the possibility of a cysteine residue reacting with TNBS could not be ruled out without investigation.

The UV absorption spectrum of TNBS-modified SKDH (90% inactivated) was identical with that of a TNP-amino derivative (Fig.4.4), with a characteristic absorption maximum at 345-348 nm and a pronounced shoulder at 420 nm (Okuyama and Satake, 1960; Kotaki *et al.*, 1964). This excluded the possibility of a cysteine residue reacting since S-TNP cysteine has no maximum in this region and shows intense end absorption at low wavelengths (Kotaki *et al.*, 1964; Hollenberg *et al.*, 1971). Furthermore, the S-TNP linkage is labile in alkaline solution (Kotaki *et al.*, 1964) and would not remain intact at pH 9.2, the pH at which the spectrum was recorded.

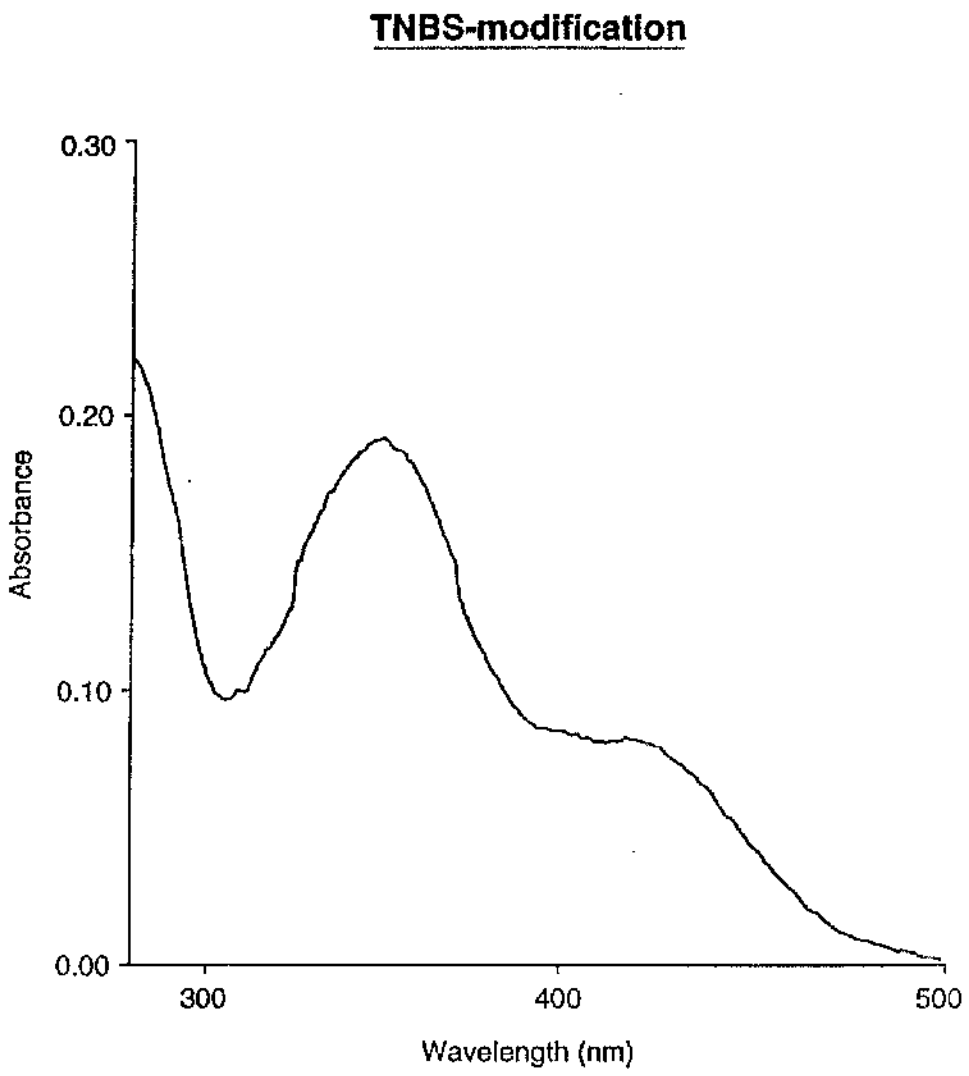


Fig. 4.4 Absorption spectrum of TNBS modified SKDH

Absorption spectrum of TNBS modified SKDH recorded against a reference cell containing the unmodified enzyme in buffer, revealed an absorption maximum at 346-348 nm and a shoulder at 420 nm characteristic of a TNP-amino derivative.

4.2.5 Stoichiometry of incorporation of TNBS into SKDH

SKDH was inactivated to different extents by treatment with TNBS in order to measure the stoichiometry of incorporation using ESMS. An aliquot from each sample prepared in section 3.6.4 was subjected to ESMS in the positive ion mode as described in section 3.6.5. The MaxEnt deconvolution procedure was applied for quantitative analysis of raw data using 1.0 Da peak width and 1.0 Da channel resolution (Ferridge *et al.*, 1992).

The electrospray profile of the partially modified sample (82% active after treatment with 5-fold molar excess TNBS) revealed only one adduct with a molecular mass of 29,625 Da (Fig. 4.5 A). The molecular mass of native SKDH deduced from the amino acid sequence is 29,414 Da. Attachment of a trinitrophenyl group as a result of reaction between TNBS and lysine would increase the mass by 211 Da. Therefore 29,625 Da (29,414+211) relates to an enzyme species containing a single modified lysine.

SKDH treated with 30-fold molar excess TNBS in the absence of substrate and coenzyme was extensively modified and >75% activity was lost. Electrospray spectrum of this sample showed three adducts with masses of 29,625 Da (29,414+211) ; 29,836 Da (29,414+422) and 30,047 Da (29,414+633) (Fig. 4.5 B). These adducts correspond to enzyme species with one, two and three modified lysines respectively.

In the presence of substrate and coenzyme incorporation of TNBS was greatly reduced compared with the previous sample although the same amount of reagent (30-fold molar excess) was used. The protected sample was 86% active and the electrospray spectrum revealed only a single adduct with a mass of 29,625 Da (29414+211) relating to enzyme species with one modified lysine residue (Fig. 4.5 C). This clearly indicates that two lysine residues recognised by TNBS are no longer available for modification in the presence of substrate and coenzyme. Thus, by the criteria of protection by substrate and stoichiometry of modification it appears that two lysine residues are functionally important in the catalytic mechanism of SKDH.

Fig.4.5 Determination of the stoichiometry of incorporation of TNBS into SKDH by ESMS

A. Partially modified (82% active) SKDH with TNBS

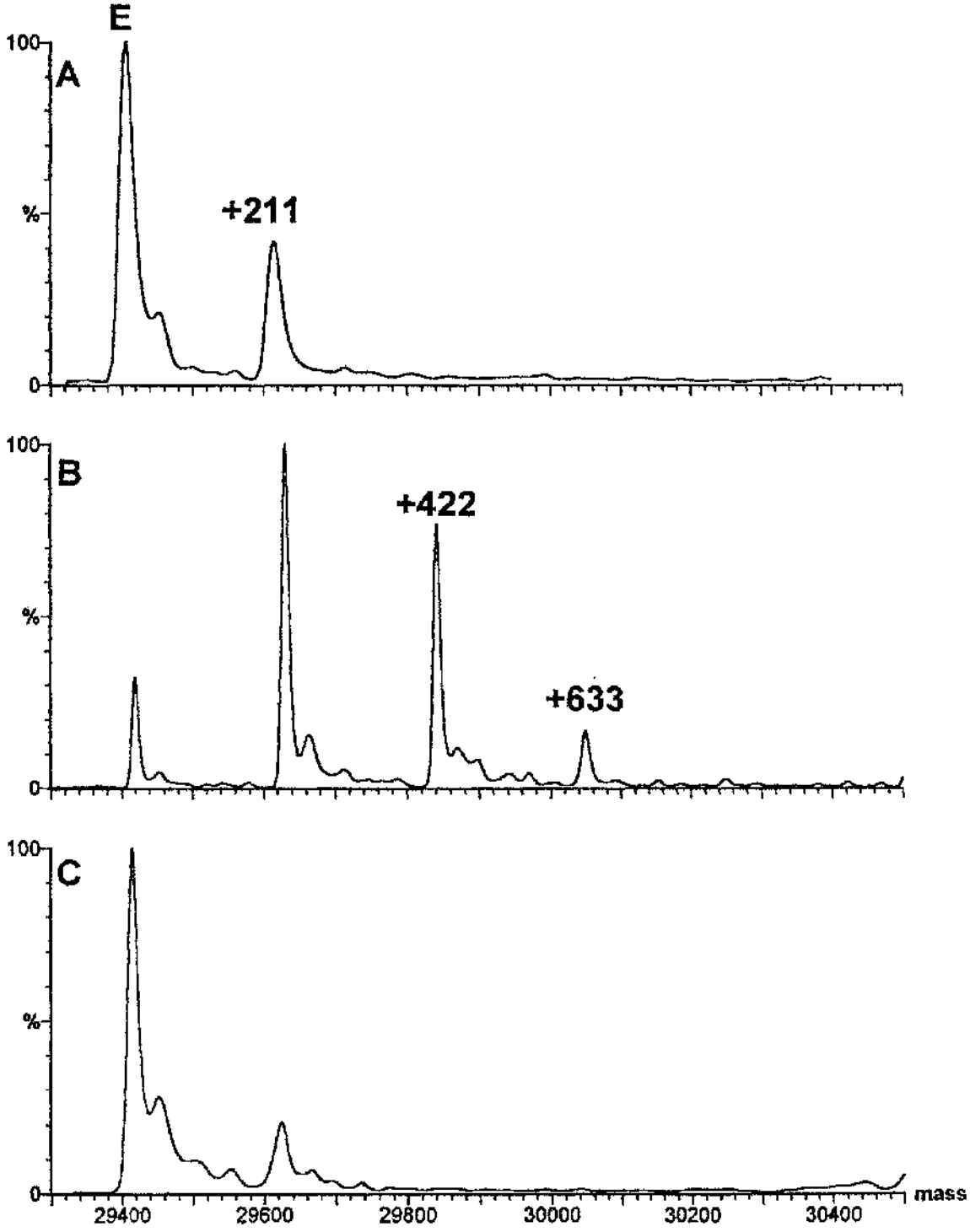
ESMS profile of the partially modified sample revealed two species of different molecular masses. The species with a molecular mass of 29,414 Da relates to unmodified SKDH (E) and the species with +211 mass difference (M_R=29,625 Da) relates to SKDH with one modified lysine.

B. Extensively modified (>75% inactive) SKDH with TNBS

ESMS profile of the extensively modified SKDH revealed three modified species having mass differences of +211 (M_R=29,625 Da), +422 (M_R=29,836 Da) and +633 (M_R=30,047 Da) indicative of one, two and three modified lysines respectively.

C. SKDH modified with TNBS in the presence of shikimate and $NADP^+$(86% active)

ESMS profile of the protected sample revealed only one adduct with a mass difference of +211 (M_R=29,625 Da) relating to enzyme species with one modified lysine residue. Accordingly, two lysines are protected from TNBS modification in the presence of shikimate and $NADP^+$.


E
100
A
%
+211
0
100
B
%
+422
+633
0
100
C
%
0
29400
29600
29800
30000
30200
30400
mass

4.3 Identification of substrate protected lysines by differential peptide mapping

Stoichiometric analysis of TNBS modified SKDH indicated that up to three lysines can be labelled in an extensively modified sample, and in the presence of $NADP^+$ and shikimate two out of the three are protected from modification. In order to identify the lysines protected by substrate differential peptide mapping approach was employed. This technique allows identification of regions of the protein which are modified by a reagent but are protected in the presence of substrate. TNBS labelled peptides can be detected by their absorbance at 346 nm and 420 nm and therefore it was not necessary to use radioactive TNBS to identify the modified peptides.

4.4 Differential peptide mapping of TNBS modified SKDH

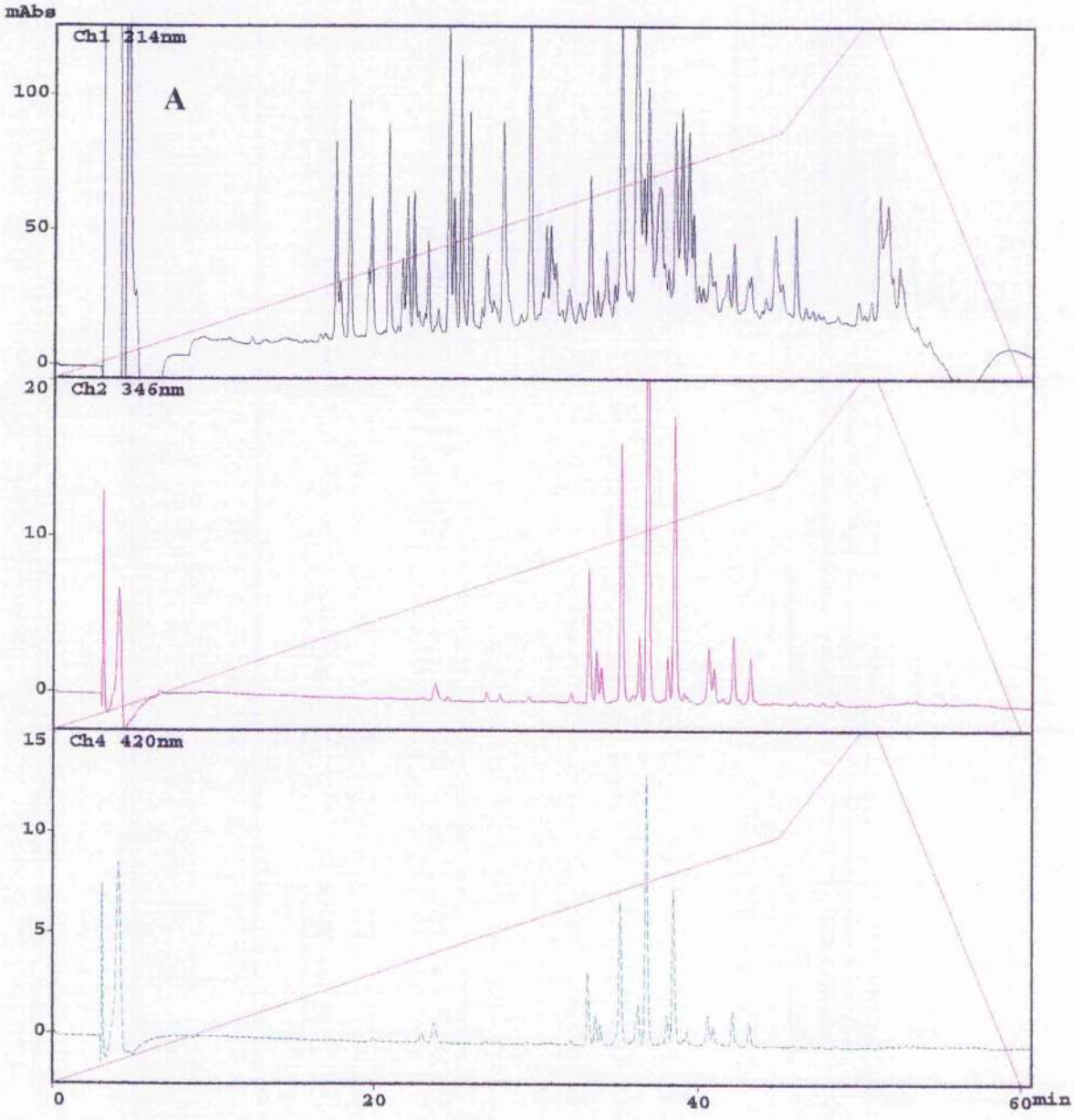
Differential peptide mapping was done using chymotryptic digests of SKDH modified with a 30-fold molar excess of TNBS in the absence and presence of $NADP^+$ and shikimate. The digests were fractionated by reverse phase HPLC using a Vydac 214TP C4 column as described in section 3.6.7. Peptides were eluted with a split acetonitrile gradient and monitored at three different wave lengths; 214 nm for detection of peptides, 346 nm and 420 nm for the detection of peptides containing TNP-lysine residues.

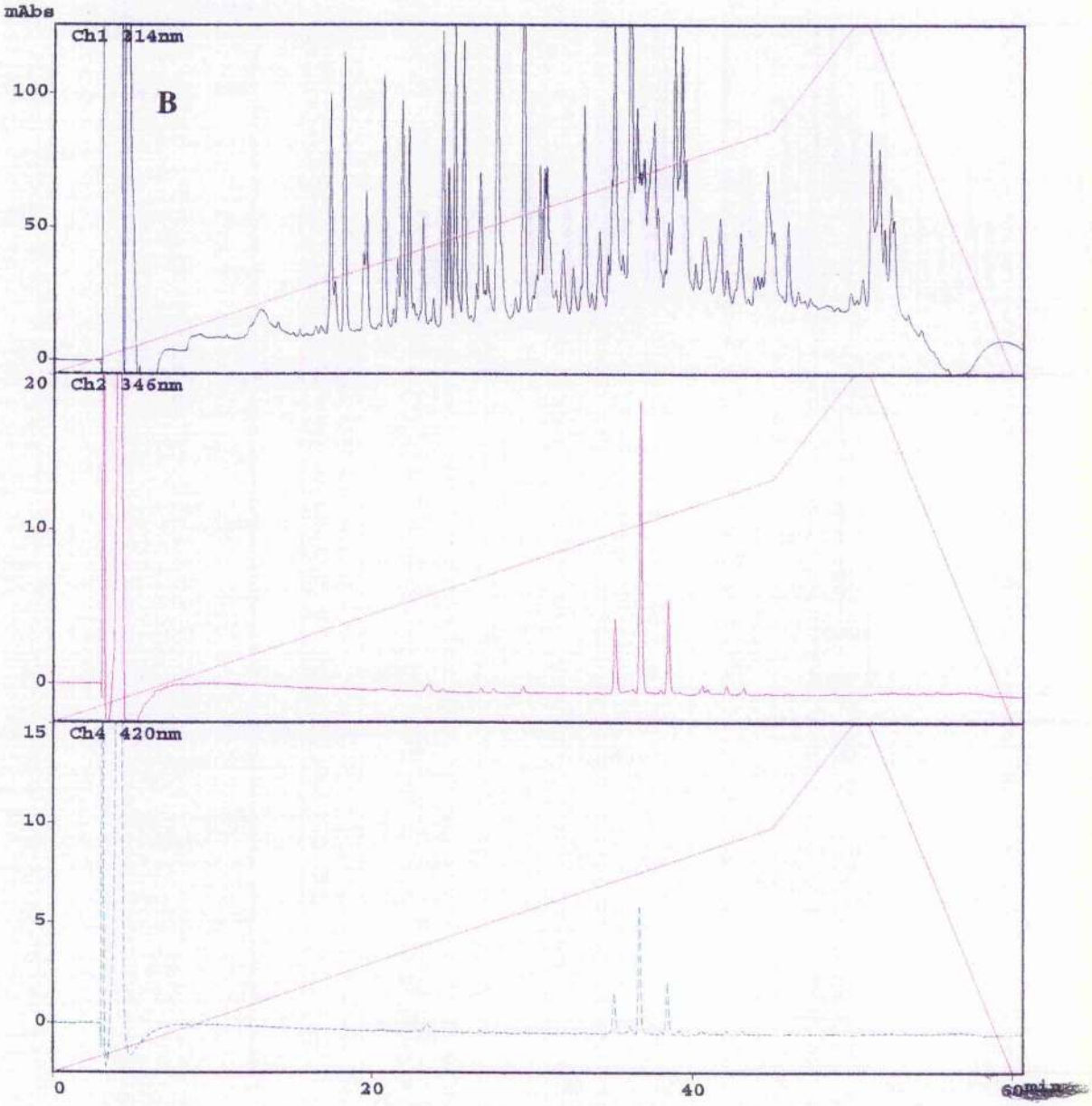
The HPLC profile at 346 nm and 420 nm of the unprotected sample, revealed many peptides containing modified residues, located in the hydrophobic region of the peptide map (Fig. 4.6.A). However, as revealed by ESMS data the unprotected sample contained only three modified residues. Therefore it appears that the modified residues are contained in large hydrophobic peptides which are further cleaved into partial peptides upon prolonged incubation. The HPLC profile of the protected sample showed significant protection by the disappearance of a number of peaks (Fig. 4.6.B) at 346 nm and 420 nm although ESMS revealed only two residues were protected. Therefore identification of any two particular peptides relating to the two protected residues proved difficult. In order to overcome this problem and to identify the substrate protected lysines a different approach where the HPLC was coupled to the mass spectrometer (LCMS) was taken.

Fig. 4.6 Differential peptide mapping by RP-HPLC

HPLC profiles of chymotryptic digests of TNBS modified SKDH in the absence (A) and in the presence (B) of shikimate and $NADP^+$

Experimental details are given in section 3.6.7. Identical quantities of protein were injected (100 μg) and peptide maps were monitored at 214 nm for the detection of peptides, 346 nm and 420 nm for the detection of peptides containing modified lysine residues. mAbs represents milli absorbance units.


mAbs
Ch1 214nm
A
100
50
0
20
Ch2 346nm
10
0
15
Ch4 420nm
10
5
0
0
20
40
60min


mAbs
Ch1 214nm
B
100
50
0
20
Ch2 346nm
10
0
15
Ch4 420nm
10
5
0
0
20
40
60min

4.5 Localization of TNBS labelled lysines by LCMS

In an attempt to identify TNBS labelled lysines in the absence and presence of substrate the approach of LCMS was employed. In this technique individual masses of the peptides eluting from the HPLC column are measured by the mass spectrometer and peptides are identified by their masses. The presence of a TNP group as a result of modification would increase the theoretical mass of a peptide by +211 Da. Chymotrypsin specifically hydrolyses the carboxy terminal amide bonds of phenylalanine, tyrosine and tryptophan. Accordingly, in a chymotryptic digest the eight lysines in the SKDH sequence would be localized on seven peptides, one containing two lysines. Therefore by scanning each lysine containing peptide for a mass difference of +211 Da, peptides containing modified lysines could be identified. The presence of a signal for a mass difference of +211 Da in the unprotected sample and its absence in the substrate protected sample would allow the identification of the protected lysine residues.

4.6 LCMS of TNBS modified SKDH

LCMS of the chymotryptic digests of TNBS modified SKDH were performed as described in section 3.6.8 using a Delta-Pak HPI C4 column, and the column was developed using a split acetonitrile gradient.

In the control experiment, a chymotryptic digest of unmodified SKDH (native) was analysed. Out of the seven lysine containing peptides six were identified by scanning their masses.

The partially modified sample (82% active after treatment with 5-fold molar excess TNBS) contained one modified lysine as revealed by ESMS. The chymotryptic digest of this sample was analysed on LCMS. Each lysine containing peptide was scanned for a mass difference of +211 Da. As a result of this scan one peptide with an added mass of 211 Da was identified. This peptide is referred to as peptide C1 (Table 4.4). The HPLC trace and the mass scan of the doubly protonated ion of peptide C1+211 are shown in Fig. 4.7.1 (A) and (B) respectively.

Table 4.4

Peptide	Sequence	M+2H/2 Da
C1	^{52}SAGGK GANVT VPFKE EAF^{69}	1011.01

(M= mass of peptide + 211 Da, H= 1.00 Da)

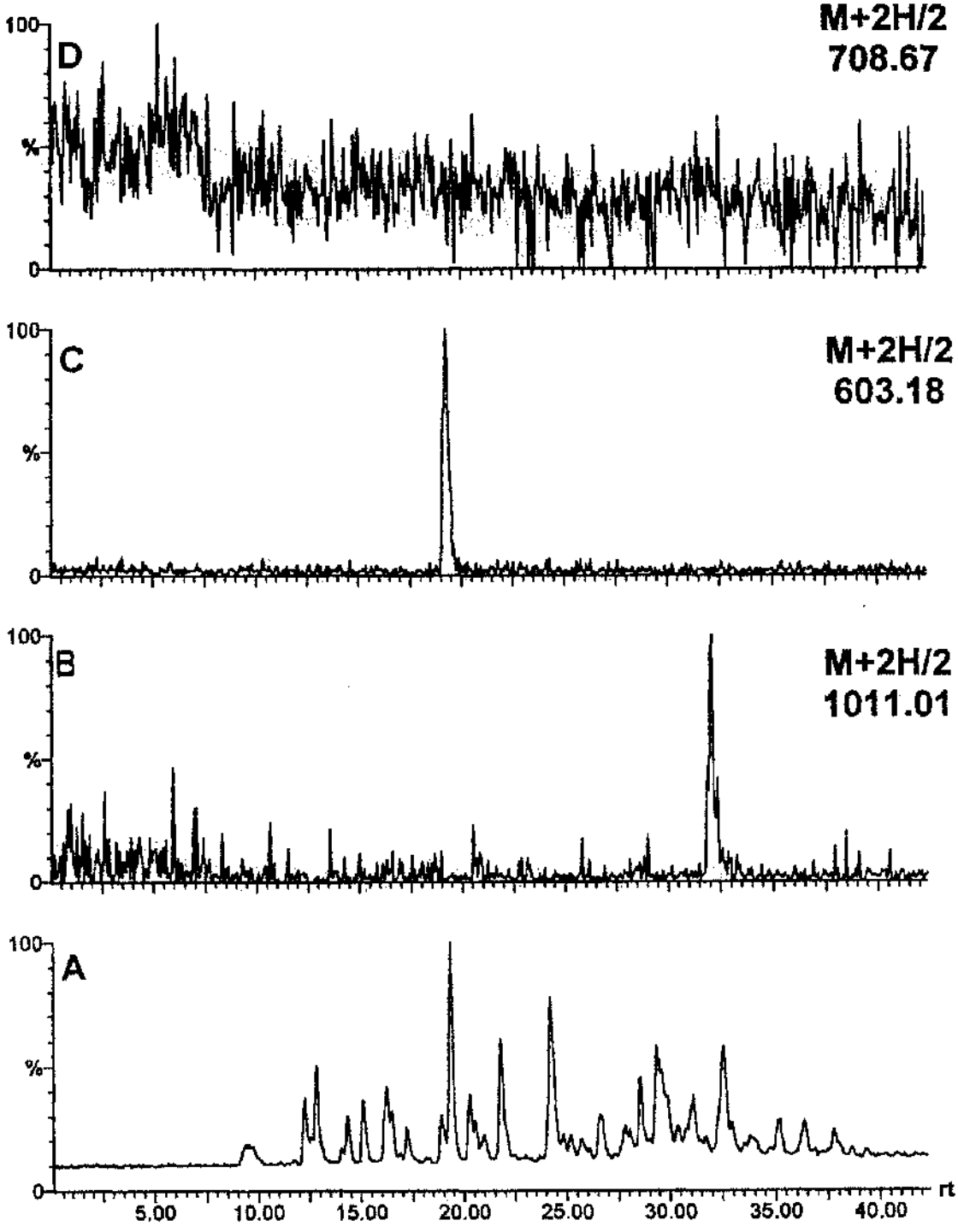
Considering the specificity of chymotrypsin, peptide C1 consists of two partial peptides referred to as peptide C1(A) - SAGGK GANVT VPF and peptide C1(B) - KEEAF. Modification of Lys-65 has resulted in the loss of a chymotrypsin sensitive site between ^{64}Phe-Lys^{65} due to the bulky nature of the trinitrophenyl group. Thus, the appearance of peptide C1 can be explained. Also, the presence of a peak corresponding to the mass of C1(A) [Fig.4.7.1(C)] and the absence of a peak corresponding to the mass of C1(A)+211 [Fig.4.7.1(D)] provides further evidence that the lysine residue modified in peptide C1 is Lys-65.

The sample treated with 30-fold molar excess of TNBS in the absence of substrate and coenzyme was >75% inactive, and contained three modified lysines as revealed by ESMS. Mass scans following LCMS of the chymotryptic digest of this sample resulted in the identification of three peptides with added masses of 211 Da. These peptides are referred to as C1 (already identified in the 82% active sample), C2 and C3 (Table 4.5). The HPLC trace and the mass scans of the doubly protonated ions of C1+211, C2+211 and C3+211 of this sample are shown in Fig. 4.7.2(A)-(D) respectively.

Fig. 4.7 LCMS

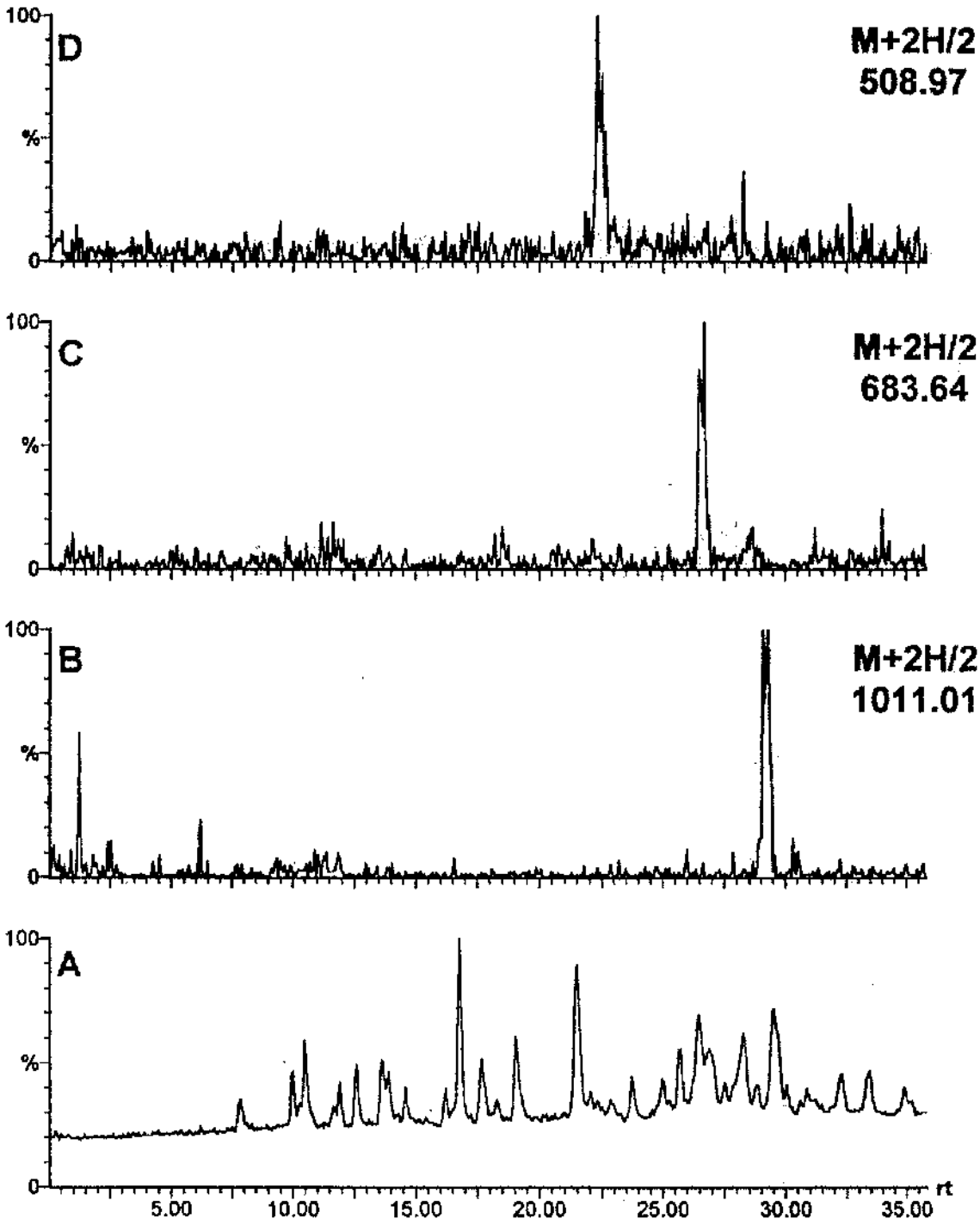
4.7.1 LCMS of the partially modified SKDH (82% active) with TNBS

HPLC trace (at 214 nm) of the chymotryptic digest of partially modified SKDH **(A)** , mass scans for the doubly protonated ions of peptides C1+211 **(B)** , C1(A) **(C)** and C1(A)+211 **(D)** illustrating that the lysine residue modified in peptide C1 is Lys-65 and not Lys-56.


100
D
%
0
M+2H/2
708.67
100
C
%
0
M+2H/2
603.18
100
B
%
0
M+2H/2
1011.01
100
A
%
0
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
rt

4.7.2 LCMS of the extensively modified SKDH (>75% inactive) with TNBS

HPLC trace (at 214 nm) of the chymotryptic digest of extensively modified SKDH (**A**), mass scans for the doubly protonated ions of peptides C1+211 (**B**), C2+211 (**C**) and C3+211 (**D**). These data indicate that the three modified lysine residues are Lys-65 , Lys-15 and Lys-217/219 contained in peptides C1, C2 and C3 respectively.


D
M+2H/2
508.97
C
M+2H/2
683.64
B
M+2H/2
1011.01
A
100
%
0
5.00
10.00
15.00
20.00
25.00
30.00
35.00
rt

4.7.3 LCMS of SKDH modified with TNBS in the presence of substrate and coenzyme (86% active)

HPLC trace (at 214 nm) of the chymotryptic digest of TNBS modified SKDH in the presence of substrate and coenzyme **(A)**, mass scans for the doubly protonated ions of peptides C1+211 **(B)**, C2+211 **(C)** and C3+211 **(D)** indicating that the two protected lysines by substrate and coenzyme are Lys-65 and Lys-15 contained in peptides C1 and C2 respectively.

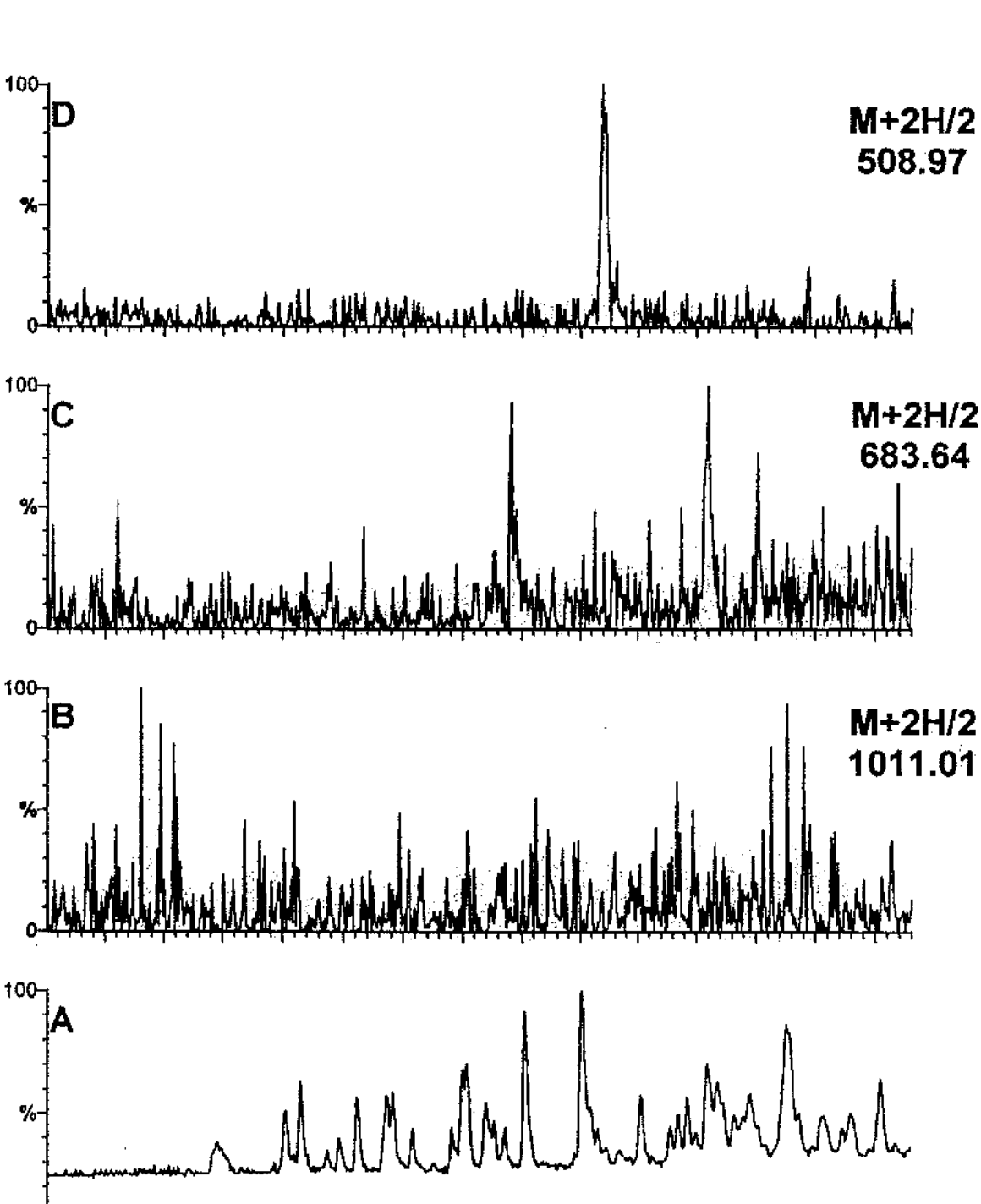

100
D
%
0
M+2H/2
508.97
100
C
%
0
M+2H/2
683.64
100
B
%
0
M+2H/2
1011.01
100
A
%
0
5.00
10.00
15.00
20.00
25.00
30.00
35.00
rt

Table 4.5

Peptide	Sequence	M+2H/2 Da
C1	^{52}SAGGK GANVT VPFKE EAF^{69}	1011.01
C2	^{8}GNPIA HSKSP F^{18}	683.64
C3	^{216}QKGKT PF^{222}	508.97

(M=mass of peptide+211, H=1.00 Da)

From these results it can be concluded that the three lysines modified in the >75% inactive sample are Lys-65, Lys-15 and Lys-217/219.

The sample modified in the presence of $NADP^+$ and shikimate was 86% active and contained only one modified residue as shown by ESMS. This residue was identified as Lys-217/219 by mass scans following LCMS of the chymotryptic digest of this sample. In comparison with the unprotected sample a weak signal was obtained for the mass of C2+211 suggesting that Lys-15 was substantially protected by $NADP^+$ and shikimate. No signal was obtained for the mass of C1+211 indicating complete protection. The HPLC trace and the mass scans of the doubly protonated ions of peptides C1+211, C2+211 and C3+211 in the protected sample are shown in Fig. 4.7.3 (A)-(D). From these results it can be concluded that Lys-65 is completely protected whereas, Lys-15 is partially protected by $NADP^+$ and shikimate.

4.7 The essential lysine residue at the active site of *E. coli* SKDH

ESMS and LCMS data together indicate that Lys-65 is completely protected and Lys-15 is partially protected by substrate and coenzyme from TNBS modification. Therefore it appears that both these residues are in the active site of SKDH. However, analysis of kinetic data (section 4.2.2) suggests that only one essential lysine is responsible for the TNBS mediated inactivation of SKDH. This may be explained by comparison of

```
E..coli        7FGNPI AHSKS PFIHQ21..... 57GANVT VPFKE EAF69
P.aeruginosa    FGNPI GHSKS PLIHR ..... GANVT VPFKE EAY
B.aphidicola    FGNPI DHSQS PKIHN ..... GANVT APFKK EAY
 A.nidulans     FGSPI SQSRS PALHN .....GGSVT IPLKL DIM
P.carinii       FGKPI KHSQS PNIHN .....GASVT IPLKT NIS
S.cerevisiae    VGKPI GHSRS PILHN .....GAAVT IPLKL DIM
P sativum       IGKPV SHSKS PILFN ..... GFSVT IPHKE SAL
N.tabacum       IGKPV SHSKS PLLYN .....GSAVT IPHKE AIV
```

Fig. 4.8 Sequence alignments around Lys-15 and Lys-65 of *E. coli* SKDH with the corresponding regions of SKDH sequences from other organisms. Lys-65 is conserved in all sequences and is shown in *boldfaces*. This alignment is taken from the GCG PileUp program (Devereux *et al.*, 1987) mounted on the University of Glasgow UNIX system.

SKDH sequences from different organisms (Fig. 4.8). Of the two protected residues Lys-65 is a conserved residue. In fact out of the eight lysines in the *E. coli* sequence Lys-65 is the only conserved lysine among all known SKDH sequences. Lys-15 is conserved in three species, namely *P. aeruginosa*, *N. tabacum* and *P. sativum* but is replaced by Arg in *A. nidulans* and *S. cerevisiae* sequences and by Gln in *B. aphidicola* and *P. carinii* sequences. It is possible that Lys-15 is present in the vicinity of the actve-site, and the binding of substrate and coenzyme at the active-site have blocked a bulky molecule like TNBS from reacting with Lys-15. Thus, considering the data from the kinetic analysis and the complete protection of Lys-65 by substrate and coenzyme together with its conservation among all known sequences it is reasonable to conclude that Lys-65 is essential to the catalytic mechanism of SKDH.

4.8 The role of Lys-65

The observation that shikimate affords protection against TNBS inactivation and that this protection is enhanced in the presence of $NADP^+$, suggested a possible role for Lys-65 in shikimate binding. Lysine residues can mediate the binding of negatively charged substrates. Thus, an ionic interaction between the carboxylate group of shikimate and the ε-amino group of Lys-65 was a strong possibility. However, the C-4 hydroxyl group of shikimate has been shown to be important in substrate binding, and a hydrogen bonding interaction between this hydroxyl group and the enzyme active site has been suggested (Bugg *et al.*, 1988). Therefore, the possibility of Lys-65 interacting with the C-4 hydroxyl group could not be ruled out without investigation.

In order to investigate the above mentioned possibilities protection experiments against TNBS inactivation were repeated using methyl shikimate, an analogue of shikimate. In methyl shikimate the carboxyl group is blocked by converting to a methyl ester. Therefore, if Lys-65 is interacting with the carboxyl group of shikimate, methyl shikimate would not be able to afford any protection against TNBS modification, whereas an interaction between Lys-65 and the C-4 hydroxyl group of shikimate would not be affected by methyl shikimate. Results of this experiment are shown in Fig.4.9.

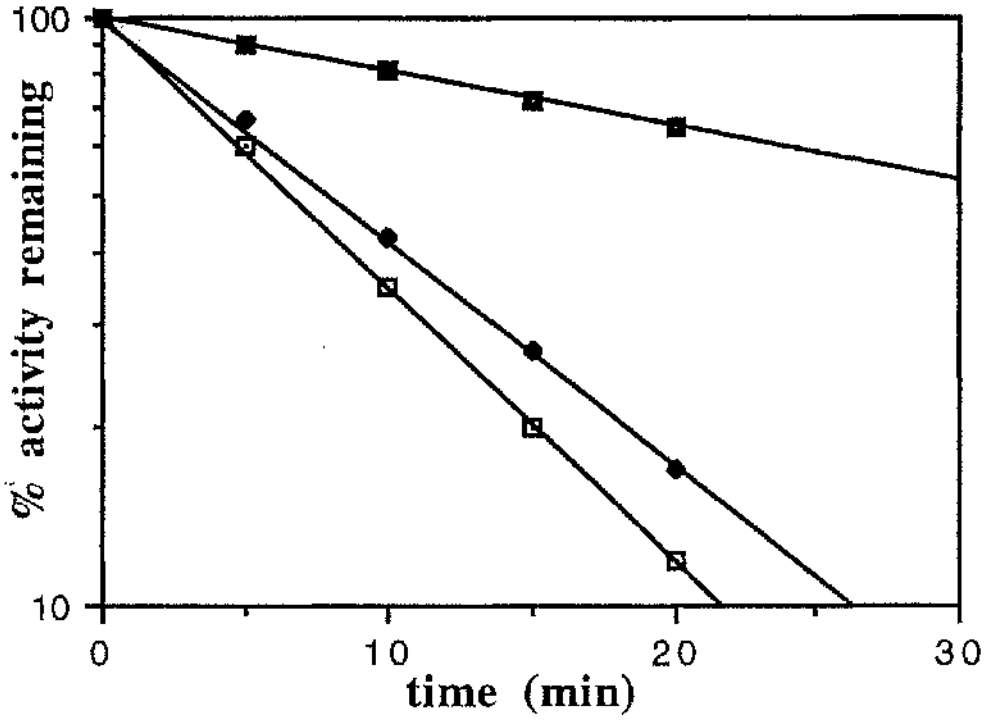


Fig. 4.9 Protection with methyl shikimate and shikimate against TNBS inactivation

SKDH was incubated with 0.20 mM TNBS alone (▫), or in the presence of 0.8 mM methyl shikimate (◆), or 0.8 mM shikimate (■). Aliquots were withdrawn at time intervals and were assayed for remaining activity.

The results indicate that methyl shikimate is not able to afford any protection against TNBS modification, while shikimate in same concentration affords significant protection. Thus it can be concluded that Lys-65 plays a role in substrate binding at the enzyme active site, via an ionic interaction between the ε-amino group and the carboxyl group of shikimate.

4.9 Discussion

In order to investigate the involvement of a lysine residue at the active site of *E. coli* SKDH and to propose a possible role in the mechanism of action, chemical modification with lysine specific reagents were attempted. In preliminary experiments SKDH was treated with methyl benzimidate, an imidoester. Imidoesters react in alkaline solution with amino groups of proteins to form imidoamides (Hunter and Ludwig, 1962; 1972). Prolonged incubation of SKDH with methyl benzimidate resulted in the inactivation of the enzyme. Inactivation was a slow process with a rate constant of 12 M^{-1} hr^{-1}. Chemical modification with TNBS also resulted in the loss of activity, in a rapid inactivation process. Substrate and coenzyme afforded protection against both methyl benzimidate and TNBS inactivation. These data suggested the presence of a lysine residue at the enzyme active site. To investigate this and to identify the active site lysine, further experiments were carried out using TNBS to chemically modify the protein. TNBS was preferred to methyl benzimidate since the reaction with methyl benzimidate was very slow and required high reagent concentrations.

Incubation of SKDH with TNBS resulted in a rapid loss of enzyme activity following pseudo first-order kinetics. The second-order rate constant of inactivation was 405 M^{-1} min^{-1} indicative of a highly reactive group at pH 9.2. The kinetic order of the reaction with respect to TNBS was 1.07 indicating that inactivation resulted due to the modification of one residue per active site. Sulfhydryl groups may react with TNBS but are less reactive than amino groups (Kotaki *et al.*, 1964). However this possibility could not be dismissed without further study. The UV absorption spectrum of TNBS modified SKDH provided the evidence to rule out this possibility. It contained a maximum at 346-348 nm

with a shoulder at 420 nm characteristic of a TNBS modified amino group (Okuyama and Satake, 1964). The observation that S-TNP groups are unstable at alkaline pH and break down to form sulfhydryl groups provides further evidence that the inactivation at pH 9.2 is due to modification of an amino group.

In protection experiments, the protection afforded by shikimate alone against TNBS inactivation implies the formation of a enzyme-substrate complex. The maximum protection obtained in the presence of both shikimate and $NADP^+$ would account for the formation of a strong enzyme-coenzyme-substrate ternary complex.

ESMS data clearly indicated the stoichiometry of incorporation. The accuracy of the measured masses of the modified species highlights ESMS as a technique of high resolution. In >75% inactivated sample three lysines were modified, and in the presence of $NADP^+$ and shikimate two of these three residues were protected suggesting that they were important residues involved in catalysis or binding.

Attempts to identify the two protected lysines by differential peptide mapping proved more difficult than anticipated. Appearance of the three modified lysines on a number of partial peptides as indicated by the HPLC trace at 346 and 420 nm and the disappearance of most of these peptides in the protected sample made identification of any two peptides corresponding to the protected residues difficult. Chymotrypsin specifically cleaves at the carboxyl terminal bonds of Phe, Tyr and Trp (Beynon and Bond, 1990) but with prolonged incubation, and upon complete opening of the structure following denaturation large peptides may be secondarily cleaved to generate partial peptides. However the approach of LCMS to identify the three modified lysines by the increase in mass was straight forward and efficient. The modified lysines were identified as Lys-65, Lys-15 and Lys-217/219. The attachment of a TNP group delayed the retention time of the modified peptides by approximately 13 min compared with the native peptides. This is presumably due to the increased hydrophobicity caused by the TNP group. The two lysines protected by $NADP^+$ and shikimate were identified as Lys-65 and Lys-15.

Although protection studies showed that two lysines were protected by substrate and coenzyme, kinetic data revealed that only one lysine was essential. This was in

agreement with sequence data since Lys-65 is the only conserved lysine in all known SKDH sequences. From these results it can be reasonably concluded that Lys-65 is the essential lysine at the active site of *E. coli* SKDH.

In order to investigate a possible role for Lys-65 protection experiments were carried out using methyl shikimate, a substrate analogue of shikimate. Methyl shikimate was not able to afford any protection against TNBS inactivation, whereas shikimate in the same concentration afforded significant protection. These data clearly indicated that the ε-amino group Lys-65 is involved in an ionic interaction with carboxylate group of shikimate.

This work has identified Lys-65 of *E. coli* SKDH as an essential residue at the binding site of shikimate. The identification of this residue opens the way for further experiments such as site-directed mutagenesis aimed more directly, at probing it's role.

CHAPTER 5

A functional arginine residue at the coenzyme binding site of *Escherichia coli* shikimate dehydrogenase

5.1 Introduction

The guanidino group of arginine is highly basic and has been proposed to play an essential role in the active sites of enzymes binding anionic substrates and coenzymes (Riordan, 1979). The high pKa (~12-13) of the guanidino functional group necessitates alkaline reaction conditions to generate an effective nucleophile in chemical modification of arginine residues. However it had been shown that the pK_a of arginyl residues at anion binding sites is lower than that of other arginine residues due to the strong electric potential of the anion binding sites (Patthy and Thesz, 1980) which would also explain their remarkable specificity towards dicarbonyl reagents used in chemical modification.

Within the family of pyridine nucleotide dependent dehydrogenases, many enzymes contain arginyl residues as components of their coenzyme binding sites; *e.g.* Arg-101 in dogfish M4 lactate dehydrogenase (Adams *et al.*, 1973) and Arg-47 in horse liver alcohol dehydrogenase (Eklund *et al.*, 1974). It has been shown by X-ray crystallographic studies that these arginine residues interact with the pyrophosphate moiety of NAD^+ facilitating it's binding to the enzyme (Eklund *et al.*, 1974; Adams *et al.*, 1973). The guanidino group is ideally suited for interaction with phosphate by virtue of its coplanar structure brought about by the partial double bond character of each of the C-N bonds (Cotton *et al.*, 1973). This unique chemistry enables the guanidino group to form multiple hydrogen bonds with phosphate thus highlighting a major biological function of arginine in interacting with phosphorylated metabolites (Riordan *et al.*, 1977).

In preliminary experiments treatment of SKDH with the arginine specific reagent phenyl glyoxal (PGO) resulted in a rapid loss of enzyme activity. $NADP^+$ protected the enzyme effectively against inactivation but shikimate did not offer any protection. These observations led to the investigation of the involvement of arginine residue(s) at the coenzyme binding site of SKDH, firstly to provide further evidence for the view that arginine residues are generally involved in coenzyme binding in the pyridine nucleotide-dependent dehydrogenases and secondly to locate the particular arginines involved.

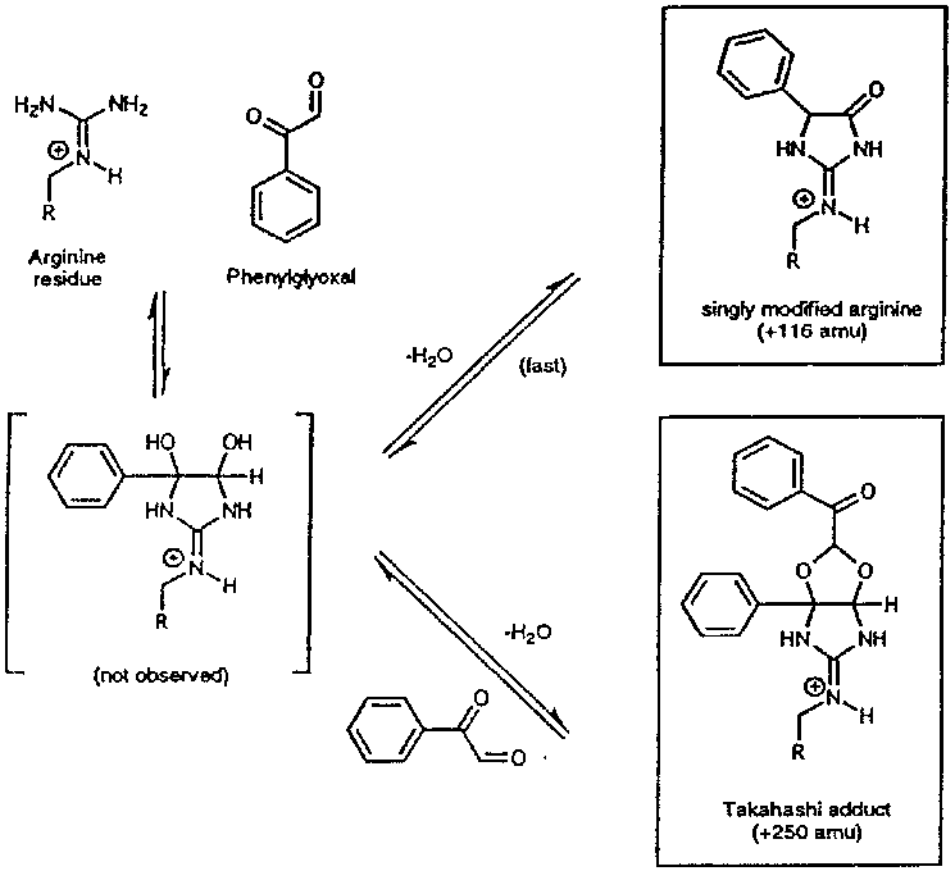


Fig. 5.1 Reaction of phenylglyoxal with arginine residues in proteins

5.2 Chemical modification of arginine with Phenylglyoxal (PGO)

5.2.1 Background

PGO was first introduced by Takahashi in 1968 as a reagent to react with guanidino groups of arginine under mild conditions (pH 7-8) to give a product that contains two PGO moieties per guanidino group (Takahashi, 1968). This product is relatively stable at acidic pH's and dissociates to regenerate arginine on prolonged incubation in neutral and alkaline media.

PGO also reacts with α-amino groups with a significant rate, probably via Schiff base formation followed by transamination, to give α-keto acyl derivatives. Upon prolonged treatment with high concentrations of PGO ε-amino groups may also react (Takahashi, 1968). The reaction of arginine with PGO is pH dependent and is greatly accelerated in bicarbonate-carbonate buffer systems (Cheung and Fonda, 1979). The specific effect of bicarbonate is thought to be due to complex formation between bicarbonate and the guanidino group, which lowers the pK_a and thus promotes nucleophilic attack by the guanidino group at the carbonyl carbon of phenylglyoxal (Riordan, 1979).

The stoichiometry of the reaction between PGO and arginine has been variable in different cases studied. It has been quoted as 2:1 (Takahashi, 1968; Lange *et. al.*, 1974) or 1:1 (Borders and Riordan, 1975; McKee and Nimmo, 1989) and has been suggested to depend on the microenvironment of the arginine residue involved (Riordan, 1979; Borders and Riordan, 1975). However more recent work, done using electrospray mass spectrometry to monitor PGO modification by measuring the increase in mass has illustrated that initial PGO modification leads to the formation of a 1:1 adduct which can very rapidly dehydrate to form a stable product or react with a second PGO molecule to form a 2:1 adduct (Krell *et al.*, 1995). This reaction scheme is shown in Fig. 5.1.

5.2.2 Kinetics of PGO modification of SKDH

Treatment of *E.coli* SKDH with PGO at 25°C caused a time-dependent loss of enzyme activity. The inactivation became more rapid as the pH was increased and a pH of 9.4 in sodium carbonate-bicarbonate buffer was adopted for further study. In the control experiment SKDH was incubated in the absence of PGO and no loss of activity was detected. The fraction of activity remaining was calculated as the percentage of activity remaining at a given time. The plot of logarithm of percent activity remaining against time at various PGO concentrations yielded straight lines indicating that the inactivation followed pseudo first-order kinetics (Fig. 5.2 A). The rate of inactivation was dependent on reagent concentration. The $t_{1/2}$ values and pseudo first-order rate constants (k_{obs}) for the inactivation process are summarized in table 5.1. Treatment of kinetic data and calculation of pseudo first-order rate constants were done as described in section 4.2.2.

Table 5.1

Half life ($t_{1/2}$) and pseudo first-order rate constants (k_{obs}) for the inactivation of SKDH at different concentrations of PGO

[PGO] mM	$t_{1/2}$ min	k_{obs}/min
0.50	99.5	0.007
1.00	22.4	0.031
1.50	14.8	0.047
2.00	11.1	0.062
2.50	7.3	0.095
3.00	5.3	0.131

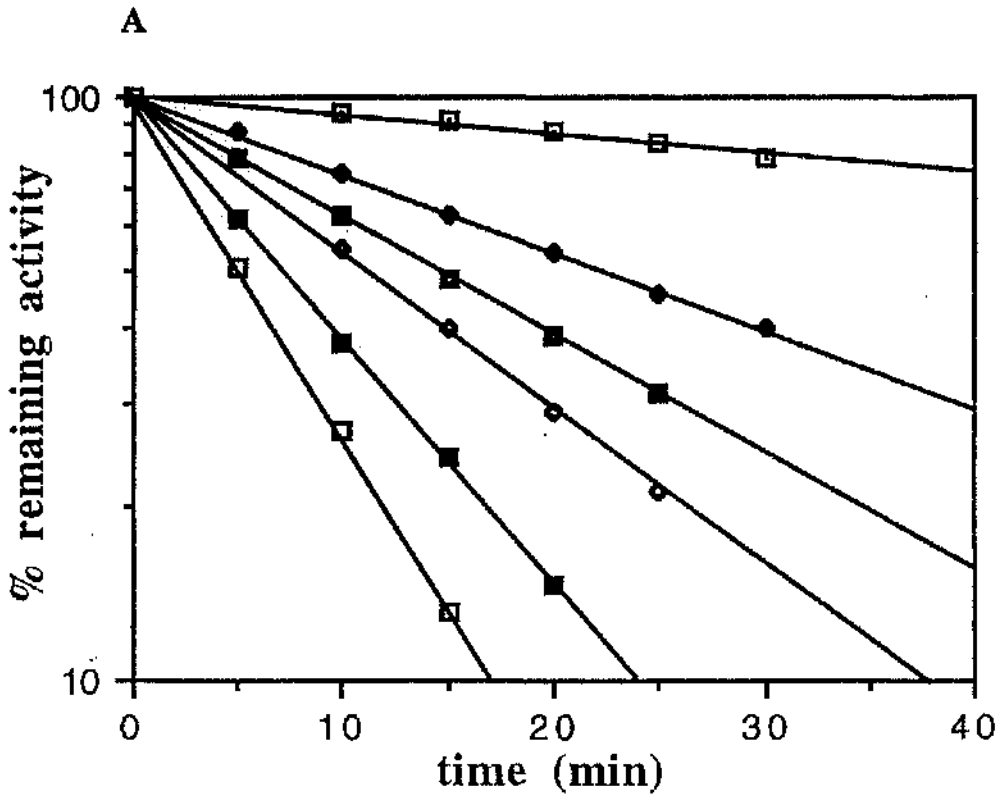
Fig 5.2 Kinetics of inactivation of SKDH by PGO

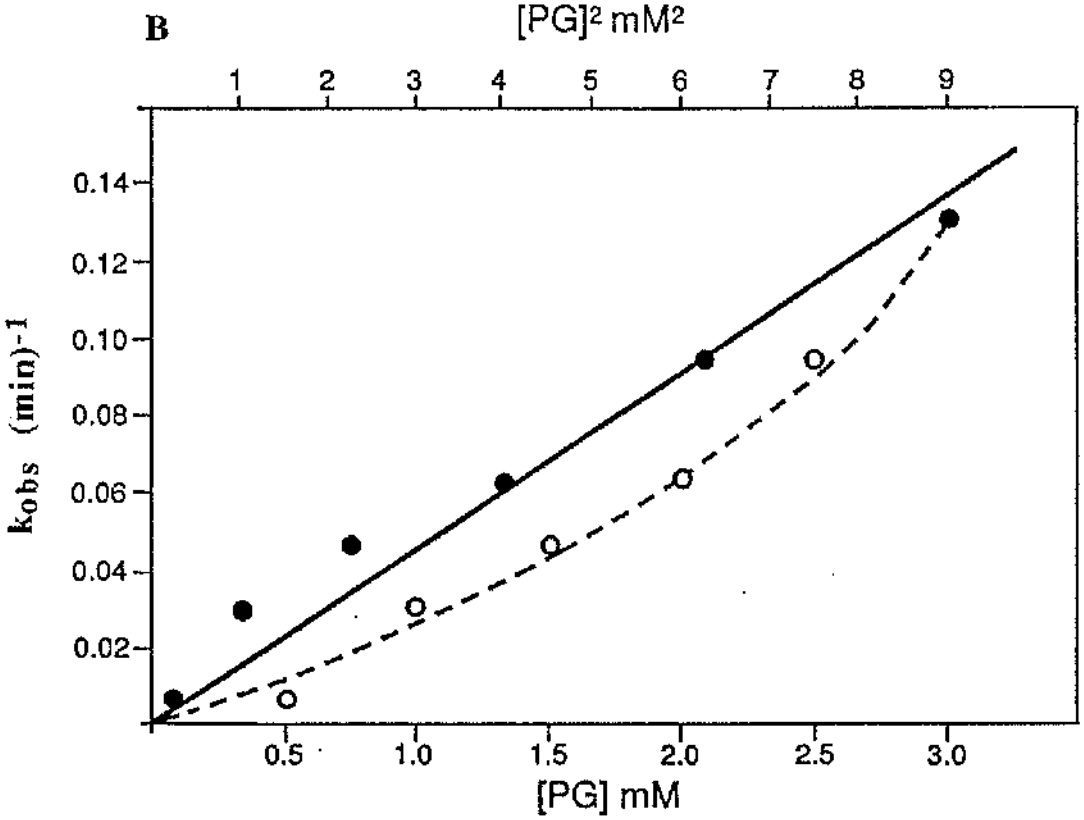
Fig 5.2 A. Pseudo first-order plots for inactivation

SKDH (2-3 μM) was incubated with increasing concentrations of PGO in 100 mM sodium bicarbonate-carbonate buffer, pH 9.4 at 25°C. Aliquots were removed at time intervals and were assayed for residual enzyme activity as described in section 3.4. The concentrations of PGO used were 0.5 mM(▫), 1.0 mM(♦), 1.5 mM(▫), 2.0 mM(◆), 2.5 mM(■) and 3.0 mM(▫).

Fig 5.2 B Relationship between pseudo first-order rate constants and PGO concentration

Pseudo first-order rate constants calculated in part A were plotted against PGO concentration (○) or the square of PGO concentration (●). These data indicate that the reaction is second-order with respect to PGO.


A
100
10
% remaining activity
0
10
20
30
40
time (min)


B
[PG]² mM²
1
2
3
4
5
6
7
8
9
0.14
0.12
0.10
0.08
0.06
0.04
0.02
kobs (min)-1
0.5
1.0
1.5
2.0
2.5
3.0
[PG] mM

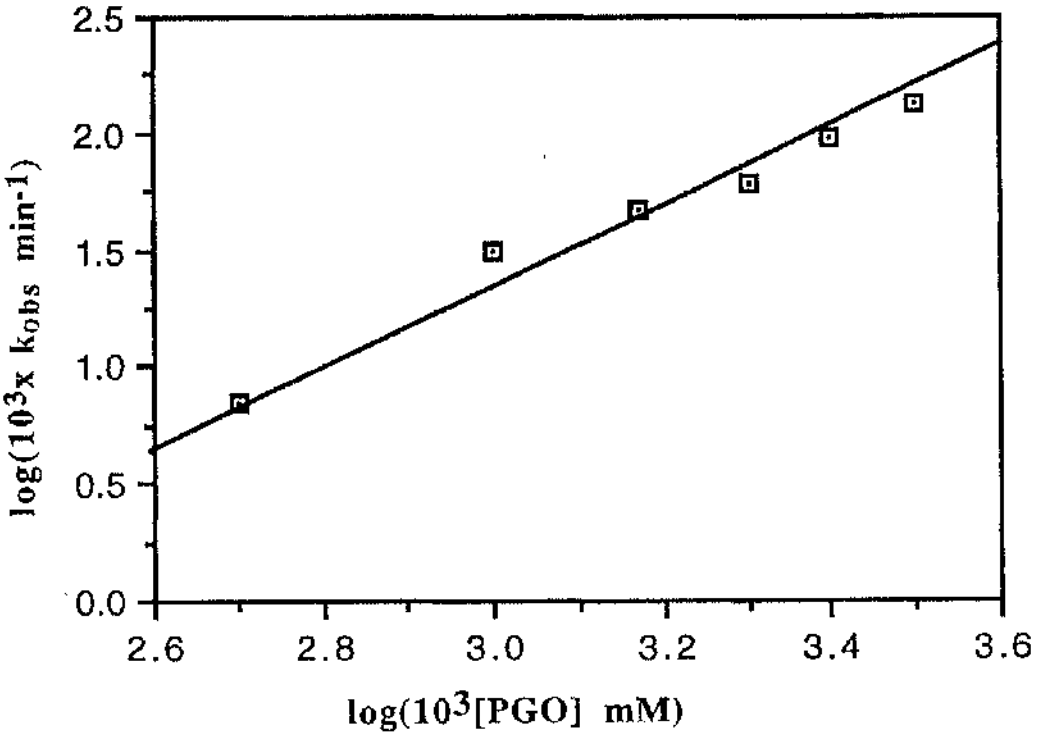


Fig. 5.2 C Determination of the order of the reaction

A plot of log (pseudo first-order rate constants) against log (PGO concentration) yielded a straight line with a slope of 1.7 indicating that approximately two molecules of PGO are needed to inactivate the enzyme.

A plot of pseudo first-order rate constants against PGO concentration is markedly nonlinear (Fig. 5.2 B). However if the data are replotted by using the square of the PGO concentration a straight line is obtained (Fig. 5.2 B). This indicates that the inactivation is first order with respect to enzyme and second-order with respect to PGO. Similar observations have been reported for L-lactate monooxygenases (Peters *et al.*, 1981; Soon *et al.*, 1978).

Consistent with this is the kinetic order calculated using equation (7) in section 4.2.2. A plot of log k_{obs} against log [PGO] yielded a straight line with a slope of 1.7 (Fig. 5.2 C), indicating that two molecules of PGO are needed to inactivate a single molecule of active enzyme unit (Levy *et al.*, 1963).

5.2.3 Substrate protection against PGO inactivation

Incubation of SKDH with PGO in the presence of shikimate and $NADP^+$ resulted in the retention of enzyme activity (Fig. 5.3) indicating that PGO is reacting with a arginine(s) at or near the active site. The percentages of protection afforded were calculated as described in section 4.2.3. Shikimate alone afforded 12% protection whereas $NADP^+$ alone increased this to a significant 78%. However in the presence of both shikimate and $NADP^+$ this was slightly increased to 81%.

These data indicate that $NADP^+$ protects the enzyme against PGO inactivation effectively and therefore PGO appears to be reacting with an arginine or arginines involved in binding of $NADP^+$ to the enzyme. This is in good agreement with other pyridine nucleotide dependent dehydrogenases where functional arginines are found at coenzyme binding sites (Adams *et al.*, 1973; Eklund *et al.*, 1974). The protection afforded by shikimate, alone or with $NADP^+$ is insignificant and may provide evidence to rule out the possibility of arginine residues being involved in shikimate binding or catalysis.

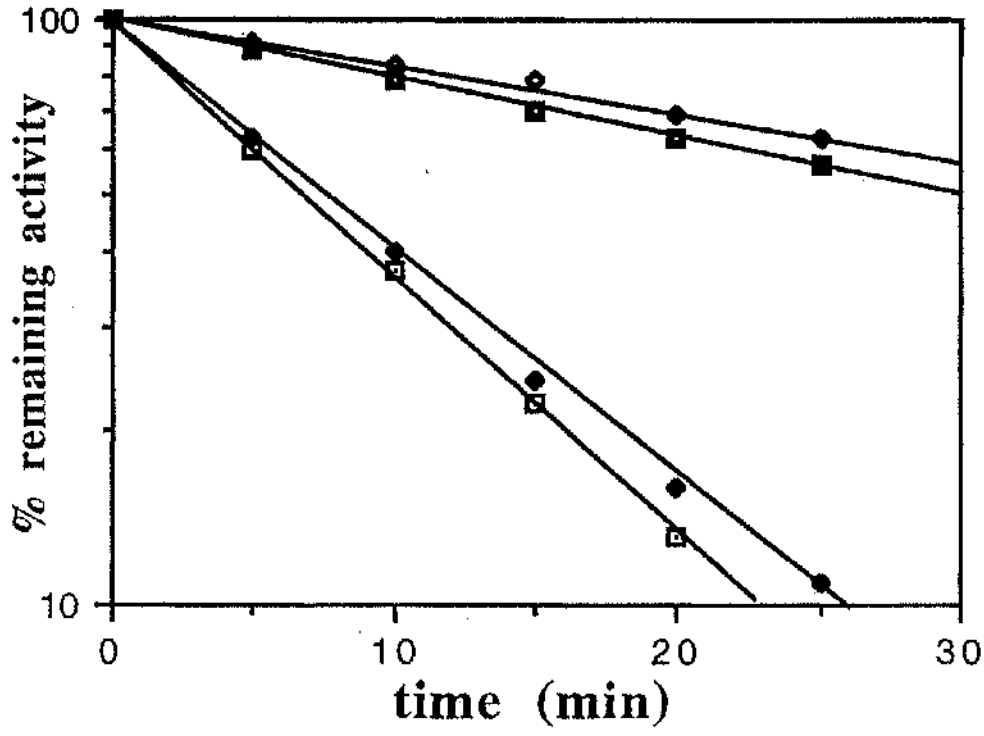


Fig. 5.3 Substrate and coenzyme protection against PGO inactivation

SKDH was incubated with 2.2 mM PGO alone(□), or in the presence of 0.4 mM shikimate(◆), or 0.2 mM $NADP^+$(■) or a mixture of 0.4 mM shikimate and 0.2 mM $NADP^+$(◈). This figure indicates the protection afforded by $NADP^+$ against PGO inactivation.

5.2.4 Stoichiometry of incorporation of PGO into SKDH

The incorporation of PGO into SKDH was determined using samples inactivated to different extents, prepared as described in section 3.7.3. The samples were analysed by ESMS in the positive ion mode as described in section 3.6.5 and quantitative analysis of raw data was done using the MaxEnt deconvolution procedure with 1.0 Da peak width and 1.0 Da channel resolution (Ferridge *et al.*, 1992).

The extent of incorporation proceeded in parallel with the loss of activity. In the ESMS profile of the 90% active sample (Fig. 5.4 A) the major peak was of unmodified enzyme with a molecular mass of 29,414 Da and the only significant modified species had a mass of 29,530 Da which indicated a mass difference of +116. Incorporation of one PGO per arginine with the loss of water would result in a mass difference of +116 (Krell *et al.*, 1995), therefore this peak corresponds to enzyme species with one arginine residue modified.

In the sample that was 60% active the two major peaks related to the unmodified enzyme and to the +116 species. These data suggested that the initial inactivation of the enzyme was due to the modification of one arginine residue, which forms a simple 1:1 adduct with PGO, followed by dehydration. Two other minor modified species were also present with molecular masses of 29,664 Da and 29,646 Da. These relate to species with mass difference of +250 corresponding to the Takahashi adduct with two PGO's per arginine and +232 corresponding to two different arginines each modified with a single PGO molecule. The electrospray profile of the 60% active sample is shown in Fig. 5.4 B.

The 35% active sample was extensively modified and a small percentage of unmodified enzyme remained (Fig. 5.4 C). The more significant modified species had molecular masses of 29530, 29646, 29664, 29780, 29896, 29914, 30012 and 30030 Da corresponding to mass differences of +116, +232, +250, +366 (250 plus 116), +482 (250 plus 116x2), +500 (250x2), +598(250 plus 116x3) and +616(250x2 plus 116) respectively. These data indicate that in extensively modified enzyme the modified arginines could have both 1:1 and 2:1 stoichiometry, and this is in good agreement with the reaction

Fig. 5.4 Electrospray mass spectrometry of PGO modified SKDH

Fig. 5.4 A. ESMS profile of the 90% active SKDH showed a major peak with a molecular mass of 29,414 Da relating to unmodified enzyme (E) and another peak with a +116 mass difference (29,530 Da) indicative of enzyme species with one modified arginine with a single PGO molecule with loss of water.

Fig. 5.4 B The 60% active sample showed two main peaks due to unmodified enzyme (E) (29,414 Da) and +116 species (29,530 Da). Minor modified species with mass differences of +232 (29,646 Da) and +250 (29,664 Da) corresponding to two modified arginines with a single PGO each and the Takahashi adduct of one modified arginine with two PGO's were also present.

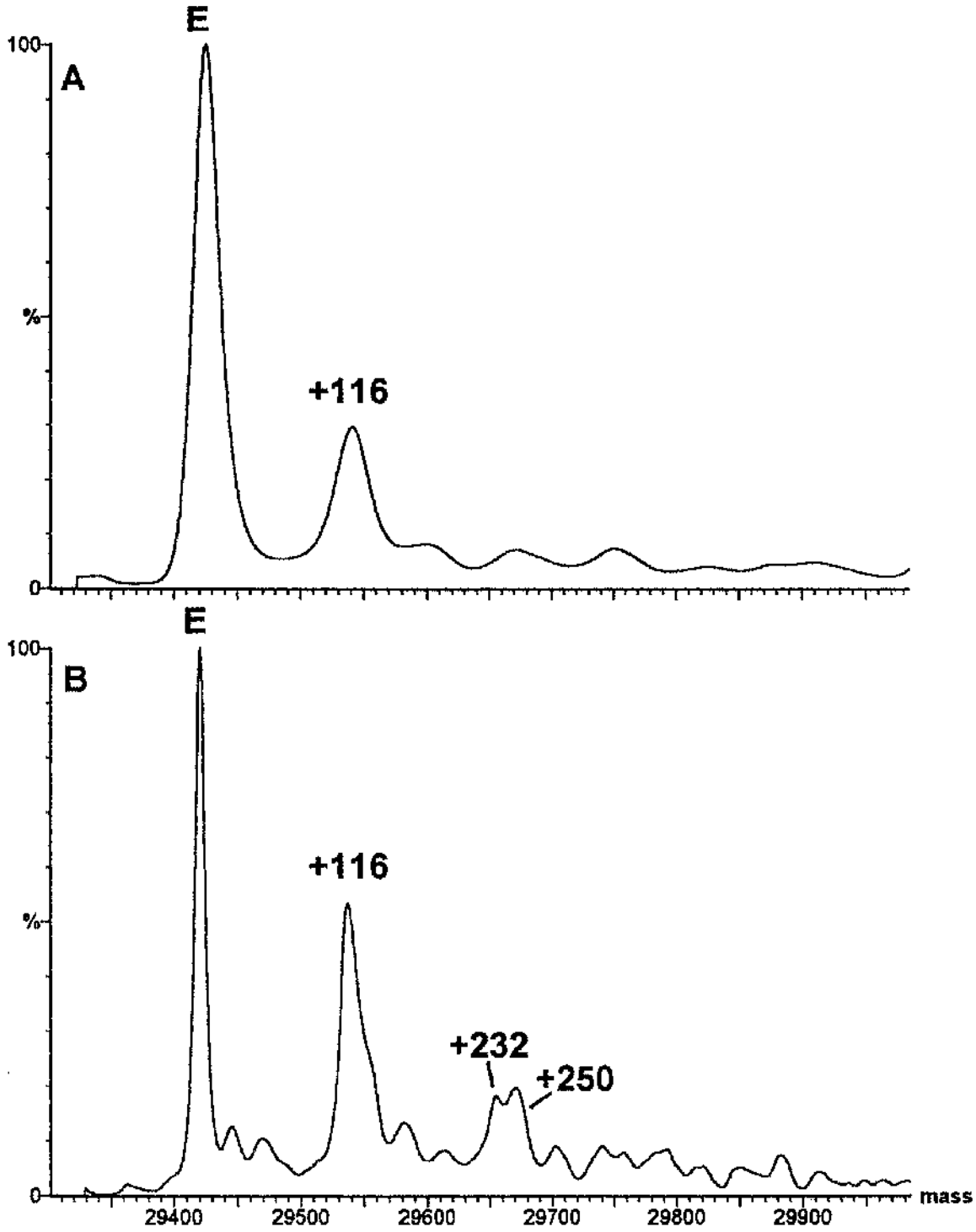
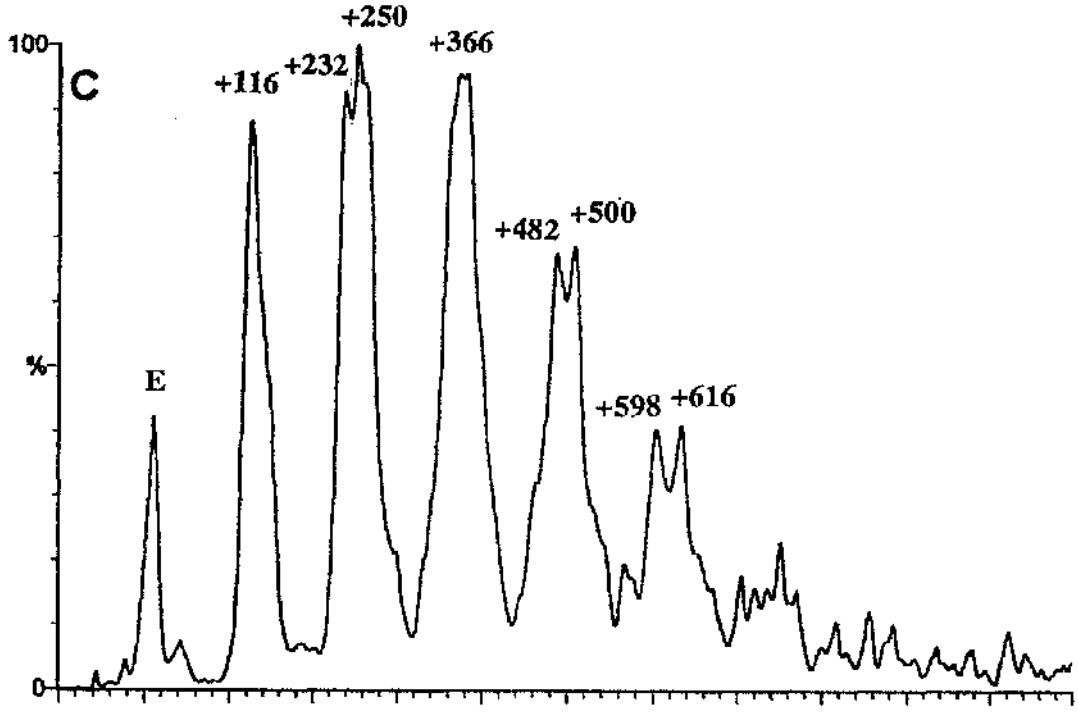
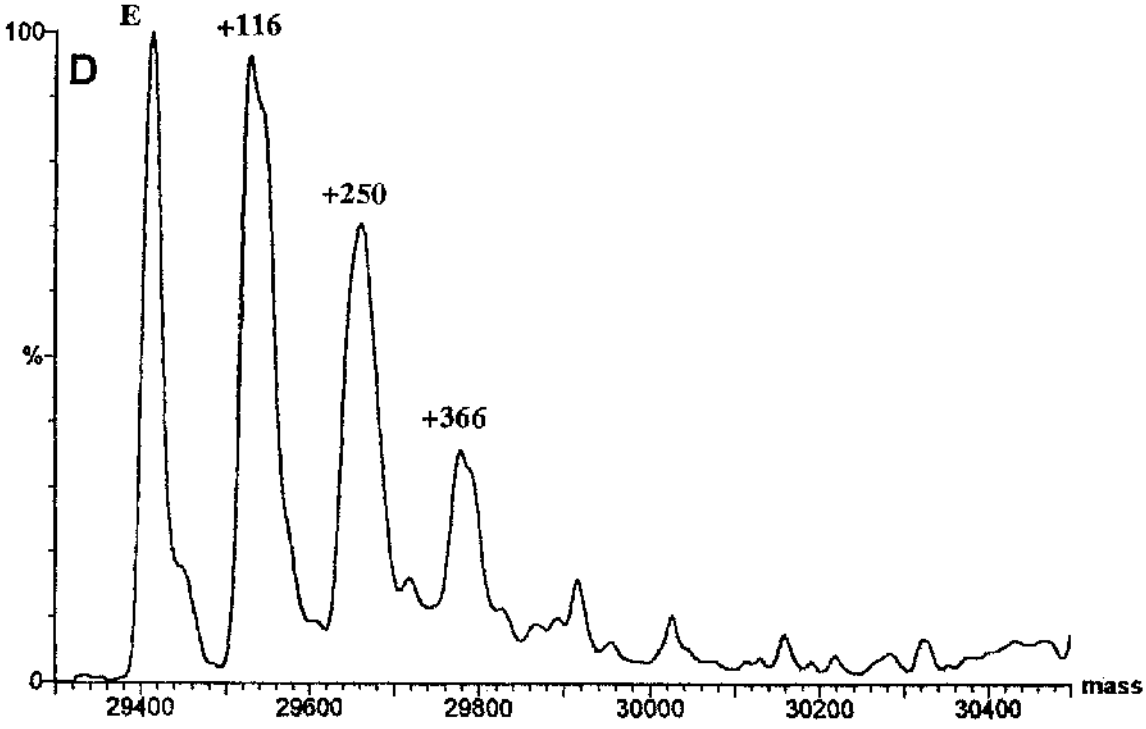

E
100
A
%
+116
0
E
100
B
%
+116
+232
+250
0
29400
29500
29600
29700
29800
29900
mass

Fig. 5.4 C ESMS profile of the 35% active sample (treated with 3.0 mM PGO in the absence of $NADP^+$) showed a number of modified species with mass differences of +116, +232, +250, +366 (116 plus 250), +482 (116x2 plus 250), +500 (250x2), +598 (116x3 plus 250) and +616 (116 plus 250x2).

Fig. 5.4 D ESMS profile of the $NADP^+$ protected sample (treated with 3.0 mM PGO in the presence of 2.0 mM $NADP^+$) showed significant protection indicated by the decrease in the number of modified species. The main peaks were of unmodified enzyme (E) (29,414 Da) and modified species with mass differences of +116, +250 and +366 (116 plus 250).


C
100
%
0
E
+116
+232
+250
+366
+482
+500
+598
+616


D
100
%
0
E
+116
+250
+366
29400
29600
29800
30000
30200
30400
mass

scheme reported by Krell *et al.* (1995). However, due to the variable stoichiometry and the relatively low stability of the PGO adduct determination of the number of arginines modified in this sample proved difficult.

In the presence of $NADP^+$ the sample was 93% active although the same amount of PGO was used as in the 35% active sample. The ESMS profile of this sample (Fig. 5.4 D) had a major peak corresponding to the unmodified enzyme and three other peaks corresponding to modified species. These peaks had molecular masses of 29530, 29664 and 29780 Da, corresponding to mass differences of +116, +250 and +366 (250 plus 116) respectively. Accordingly, in the presence of $NADP^+$ the number of modified species was greatly reduced and binding of $NADP^+$ to the enzyme had protected one or more arginines from PGO modification. The protected arginine(s) could be functionally important in the binding of the coenzyme.

5.3 Differential peptide mapping of PGO modified SKDH by RP-HPLC

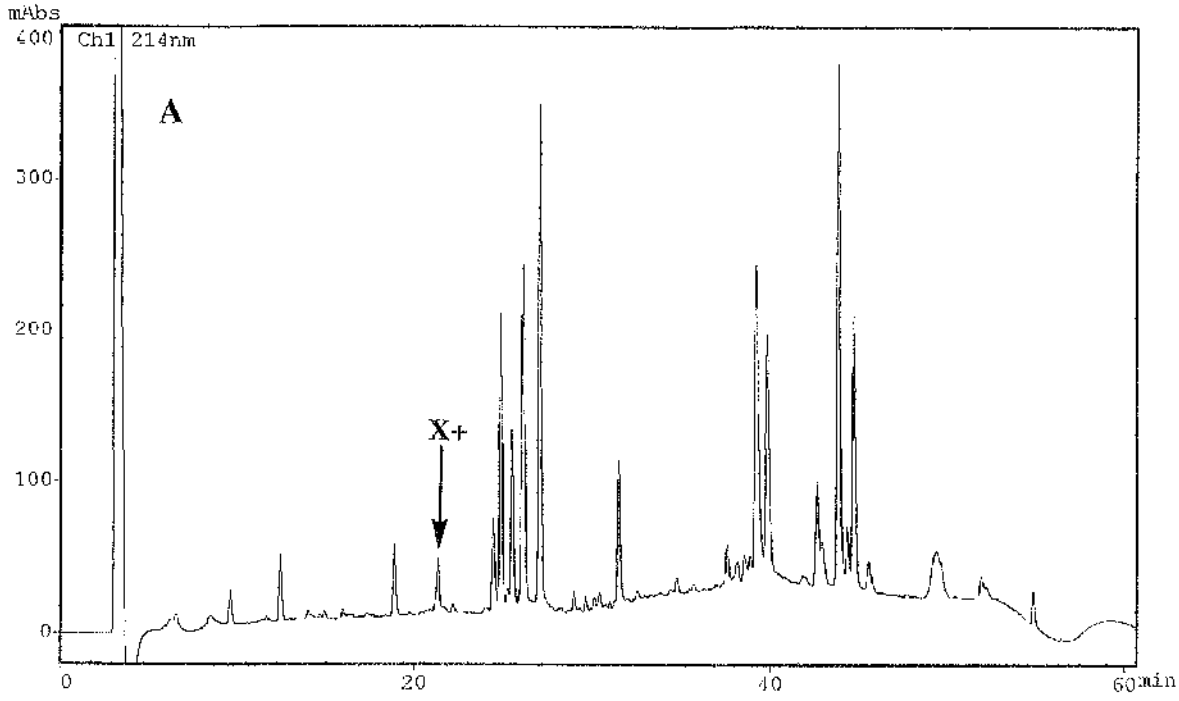
In order to identify peptide(s) containing arginine(s) protected by $NADP^+$, differential peptide mapping of PGO modified SKDH in the absence and presence of $NADP^+$ was attempted. Detailed comparison of the peptide maps at 214 nm would allow the identification of regions protected by $NADP^+$ from PGO modification.

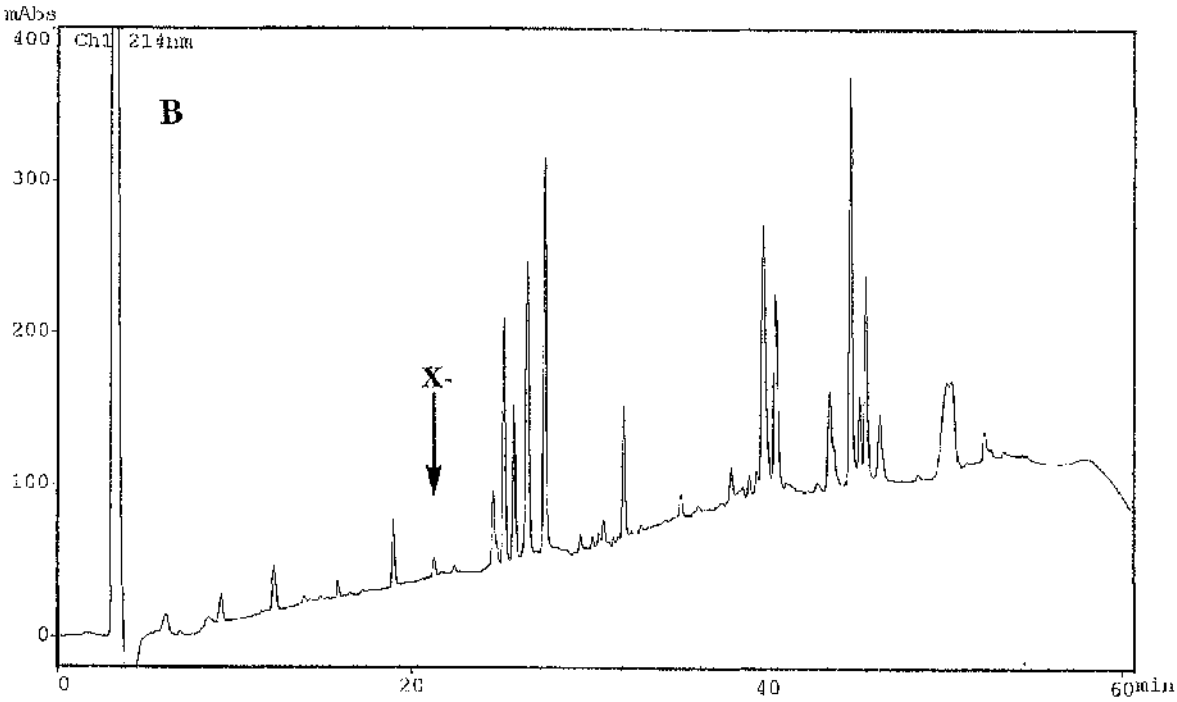
Differential peptide mapping was performed using the tryptic digests of SKDH modified with 3.0 mM PGO in the presence and absence of 2.0 mM $NADP^+$. Preparation and proteolysis of the samples are described in sections 3.7.3 and 3.7.5 respectively. Samples were fractionated using a Vydac 214TP C4 column and peptides were eluted with a split acetonitrile gradient as described in section 3.7.6. Detailed inspection of the peptide maps at 214 nm of PGO modified SKDH in the absence of $NADP^+$ (Fig. 5.5.A) and in the presence of $NADP^+$ (Fig. 5.5.B) showed that the peptide maps were identical except for one protected peptide. This peptide clearly eluting after approximately 21.2 minutes, was present in the sample modified in the absence of $NADP^+$ and was absent in the protected sample. It therefore seemed likely that this was an active site peptide containing a functional arginine. This difference in profiles observed at 21.2

Fig. 5.5 Differential peptide mapping by RP-HPLC

HPLC profiles of tryptic digests of PGO modified SKDH in the absence (A) and presence (B) of $NADP^+$

Experimental details are given in section 3.7.6. Identical quantities of protein were injected (80 μg) and the peptide maps were monitored at 214 nm, the common wave length for the detection of peptides. The peptide maps were identical except for one peptide which had a retention time of 21.2 min designated as X(+) in profile A and X(-) in profile B. mAbs represents milli absorbance units.


mAbs
400
Ch1 214nm
A
300
200
X+
100
0
0
20
40
60min


mAbs
400
Ch1 214nm
B
300
200
X-
100
0
0
20
40
60min

minutes was highly reproducible between separate experiments.

5.4 Identification of $NADP^+$ protected peptide by LCMS

To identify the peptide protected by $NADP^+$, and hence the functional arginine the preferred technique was LCMS as described by Krell *et al.* (1995) for simple and rapid identification of PGO modified peptides in the type II dehydroquinases.

Tryptic digests of SKDH modified with PGO in the absence and presence of $NADP^+$ (different aliquots of the samples used in section 5.3) were analysed using LCMS as described in section 3.7.7 and the peptides were scanned for an added mass of 116 and 250. In the unprotected sample two peptides with +116 mass units were identified. These peptides are referred to as T1 and T2 (Table 5.2). However, at this stage detection of peptides with a mass of +250 proved difficult possibly due to the instability of the +250 adduct (Takahashi, 1968) during the prolonged incubation at alkaline pH required for proteolysis. The HPLC trace and the mass scans of the doubly charged ions of peptides T1+116 and T2+116 are shown in Fig. 5.6.1 A-C respectively.

Table 5.2

Peptide	Sequence	M+2H/2 Da
T1	^{151}TVSRA EELAK^{160}	610.63
T2	^{79}AALAG AVNTL MRLED GR^{95}	938.02

(M= Mass of peptide+116, H=1.00 Da)

Trypsin is a highly specific protease which cleaves at the carboxy group of Lys and Arg (Beynon and Bond, 1989) and the modification of either of these residues should lead to the loss of trypsin sensitive sites. Similar observations of loss of trypsin sensitive sites following modification of lysine have been reported by Xia *et al.* (1993).

Fig. 5.6.1 LCMS of PGO modified SKDH in the absence of $NADP^+$

Experimental details are given in section 3.7.7. HPLC trace (at 214 nm) of the tryptic digest of PGO modified SKDH in the absence of $NADP^+$ **(A)**, mass scans of the doubly charged ions of peptides T1+116 **(B)** and T2+116 **(C)** which resulted in the identification of Arg-154 and Arg-90 as the modified arginines in peptides T1 and T2 respectively.

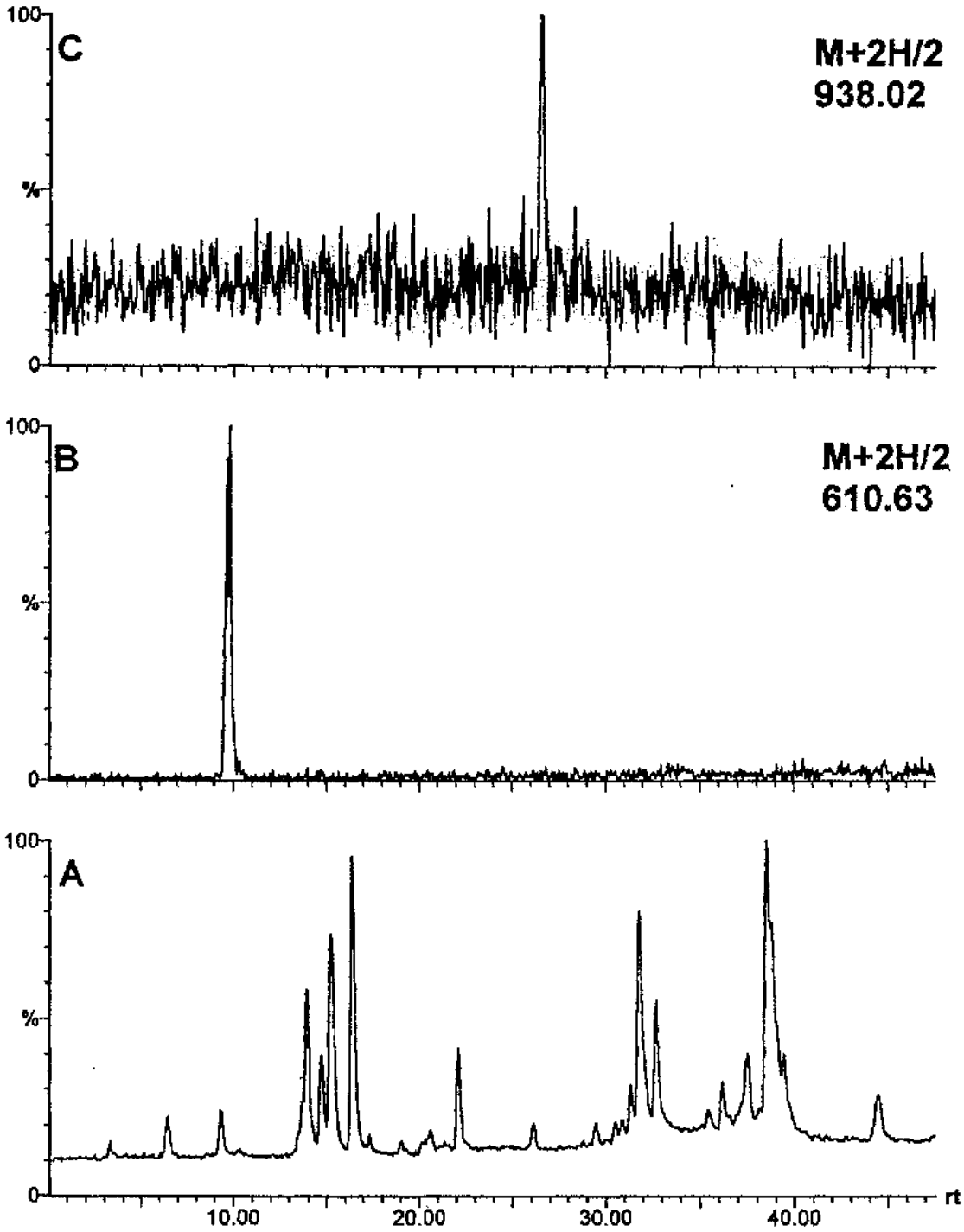
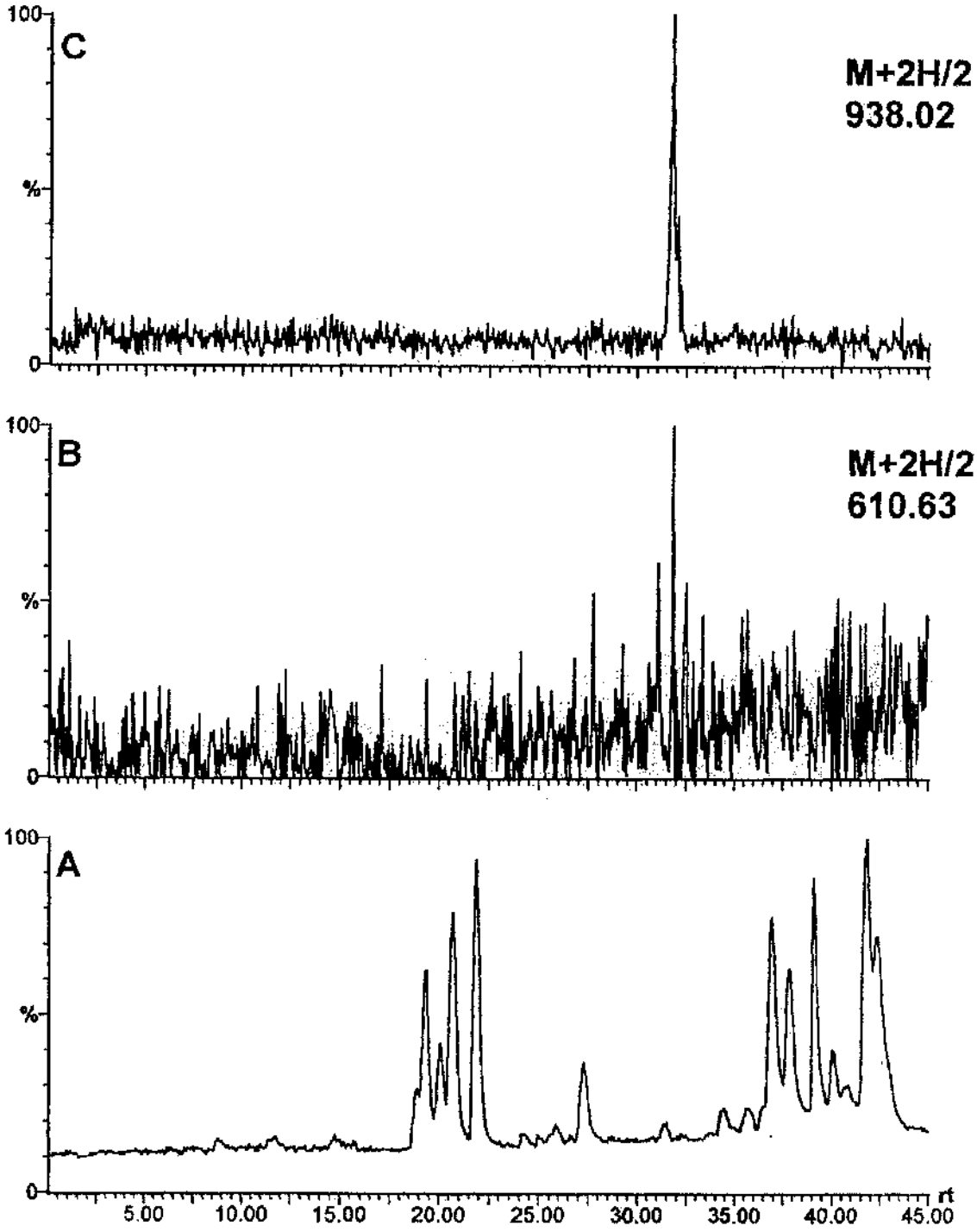

100
C
%
0
M+2H/2
938.02
100
B
%
0
M+2H/2
610.63
100
A
%
0
10.00
20.00
30.00
40.00
rt

Fig. 5.6.2 LCMS of PGO modified SKDH in the presence of $NADP^+$

HPLC trace (at 214 nm) of the tryptic digest of PGO modified SKDH in the presence of $NADP^+$ **(A)**, mass scans of the doubly charged ions of peptides T1+116 **(B)** and T2+116 **(C)**. Absence of a peak in the mass scan for T1+116 indicate that Arg-154 is completely protected from PGO modification in the presence of $NADP^+$.


100
C
%
0
M+2H/2
938.02
100
B
%
0
M+2H/2
610.63
100
A
%
0
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
rt

Therefore, Arg-154 and Arg-90 can be confirmed as the sites of modification in peptides T1 and T2 respectively.

Mass scans following LCMS of the protected sample showed that in the presence of $NADP^+$, Arg-90 in peptide T2 is still accessible to PGO. However, no peak was obtained for the mass of T1+116 indicating that Arg-154 is completely protected by $NADP^+$ from PGO modification. The HPLC trace of the protected sample and the mass scans of the doubly charged ions of peptides T1+116 and T2+116 are shown in Fig. 5.6.2 A-C. Comparison of LCMS data with differential peptide mapping data in section 5.3, showed that the $NADP^+$ protected peptide with a retention time of 21.2 minutes, was the same peptide as peptide T1. Accordingly, it can be concluded that in the presence of $NADP^+$ Arg-154 is completely protected from PGO modification.

5.5 The role of Arg-154

The structural properties of the guanidino group of arginine residues enables multiple hydrogen bond formation with phosphate groups (Cotton *et al.*, 1973). In NAD^+ and $NADP^+$, the pyrophosphate group which bridges the nicotinamide and the adenosine moieties of the molecule provides an excellent site for such hydrogen bond formation. This is best illustrated in dogfish M4 LDH and LADH, where Arg-101 and Arg-47 respectively, are hydrogen bonded to the pyrophosphate group of NAD^+ (Adams *et al.*, 1974; Eklund *et al.*, 1974). However, in $NADP^+$ there is an additional phosphate group at the 2' position of the adenosine moiety which can also be hydrogen bonded with an arginine residue. In human glutathione reductase Arg-218 and Arg-224 interact with this 2' phosphate group and this has been shown to confer specificity for $NADP^+$ (Scrutton *et al.*, 1990). Also in NADP-dependent cytoplasmic malate dehydrogenases arginine residues facilitate the binding of the 2' phosphate group of the adenine ribose (Nishiyama *et al.*, 1993).

In order to investigate whether Arg-154 of SKDH is interacting with the 2' phosphate of the adenine ribose, protection experiments were repeated with NAD^+. SKDH binds with NAD^+ but activity is ten fold less than with $NADP^+$ perhaps because NAD^+ lacks the 2' phosphate group. However, in the presence of NAD^+ no protection was

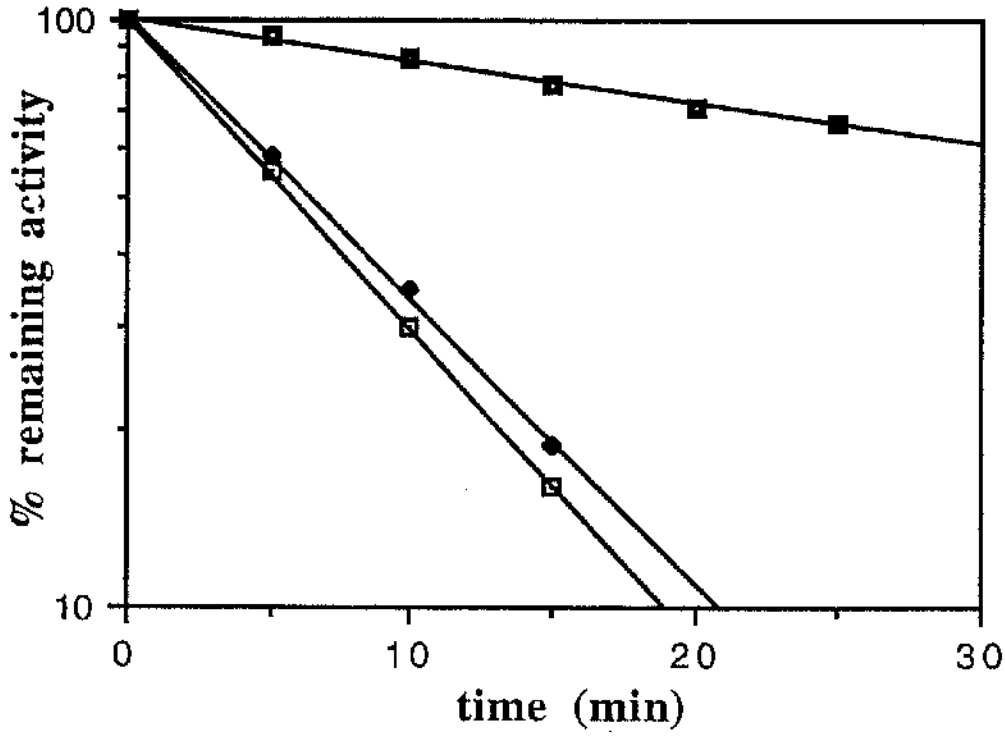


Fig 5.7 Protection against PGO inactivation with NAD^+ and $NADP^+$. SKDH was incubated with 3.0 mM PGO alone (□), or in the presence of 0.2 mM NAD^+ (◆), or 0.2 mM $NADP^+$ (■)

obtained against PGO inactivation whereas, the same concentration of $NADP^+$ afforded strong protection (Fig. 5.7). This suggest that Arg-154 is not interacting with the pyrophosphate bridge of the ADP moiety; instead it appears likely that Arg-154 of *E. coli* SKDH interacts with the phosphate group at the 2' position of the adenosine moiety of $NADP^+$.

5.6 Active site homology

The present investigation led to the identification of Arg-154 as a component of the coenzyme binding site of *E. coli* SKDH. Comparison of amino acid sequence around Arg-154 with matching SKDH sequences from other organisms showed that Arg-154 is conserved in *N.tabacum*, substituted by Lys in four other sequences (*P. aeruginosa*, *A. nidulans*, *S. cerevisiae* and *P.carinii*) and by Asn in *B. aphidicola* sequence (Fig. 5.8). A common feature of Arg, Lys and Asn is they all contain $-NH_2$ groups which are capable of hydrogen bond formation. Although the involvement of arginine and lysine residues in hydrogen bond formation are well known, many cases have been reported where Asn residues are also involved in hydrogen bond formation, such as Asn-240 of *Clostridium symbiosum* glutamate dehydrogenase and Asn-160 of human aldose reductase (Baker *et al.*, 1992; Wilson *et al.*, 1992). Therefore it is reasonable to infer from these data that the residue equivalent to Arg-154 in the *E. coli* sequence is a functionally conserved residue involved in a hydrogen bonding interaction with the 2' phosphate of the adenosine ribose of $NADP^+$.

5.7 Discussion

Arginines play an important role in biological systems in interacting with phosphorylated metabolites. This is best illustrated in the glycolytic pathway, where all of the intermediates are esters of phosphoric acid and all but one of the glycolytic enzymes contain essential arginine residues at their active sites (Riordan *et al.*, 1977). A similar observation is found in pyridine nucleotide-dependent enzymes, where functional arginines are found at coenzyme binding sites. The present work was based on this observation.

E. coli	^{149}N	R	T	V	S	**R**	A	E	E	L	A	K^{160}
P. aeruginosa	N	R	T	A	R	**K**	A	V	D	L	A	E
B. aphidicola	N	R	T	I	L	**N**	A	K	I	L	V	K
A. nidulans	G	R	T	P	S	**K**	L	E	N	M	V	S
P. carinii	N	R	S	K	D	**K**	L	N	K	L	Y	H
S. cerevisiae	N	R	T	T	S	**K**	L	K	P	L	I	E
N. tabacum	N	R	T	Y	E	**R**	A	R	E	L	A	D

Fig. 5.8 Sequence alignments of the active site peptide identified by LCMS with corresponding regions of SKDH sequences from other organisms

The numbering is according to the *E. coli* enzyme. This alignment is taken from the GCG PileUp program (Devereux *et al.*, 1987) mounted on the University of Glasgow UNIX System.

PGO is a dicarbonyl reagent commonly used in selective modification of arginines, but it can also react with α-amino groups (Takahashi, 1968; Lundblad and Noyes, 1984). It had been reported that functional arginines have lower pK_a values than other arginines due to the strong positive electric potential of the anion binding sites (Patthy and Thesz, 1980). Thus functional arginines are hyper-reactive and can be modified selectively thereby facilitating structure-function correlation. Incubation of SKDH with PGO resulted in the inactivation of the enzyme in a pseudo first-order fashion. Analysis of kinetic data revealed that the reaction is first order with respect to enzyme and second-order with respect to PGO. This suggested that the inactivation of the enzyme could result either from modification of one essential arginine with two PGO's or from modification of two essential arginines with a single PGO each. However, a possible explanation for this could be obtained through ESMS data. The 90% active sample revealed only one modified species with a mass difference of +116 relating to one arginine residue modified with a single PGO with loss of water. In the 60% active sample this modified species was more significant indicating that initial inactivation of the enzyme was due to the modification of one arginine. Therefore it seems likely that there is only one essential arginine which forms a 1:1 intermediate initially with PGO which rapidly dehydrates to form a stable product. This is in good agreement with the reaction scheme reported by Krell *et al.* (1995). Also, according to this scheme the initial 1:1 intermediate could react with a second PGO molecule to form a 2:1 adduct. Therefore a modified arginine could have both 1:1 and 2:1 stoichiometry. This is clearly indicated in the ESMS profile of the 35% active sample where both adducts of modified arginines are revealed. However, the bulky nature of the PGO molecule could cause the addition of a second PGO to the initial intermediate to be rate limiting, leading to an inactivation which is second-order in PGO (Peters *et al.*, 1981). The appearance of a minor peak for +250 mass units (2:1 adduct) in the 60% active sample may strengthen the above argument.

PGO modification in the presence of $NADP^+$ resulted in the retention of most of the enzyme activity. ESMS data of the $NADP^+$ protected sample showed a significant reduction of the number of modified species providing evidence for the

protection afforded by $NADP^+$. These data suggested the involvement of at least one functional arginine in $NADP^+$ binding. Having inferred this, their identification using differential peptide mapping and LCMS was attempted.

In a control experiment, a tryptic digest of unmodified (native) SKDH was analysed using LCMS. There are thirteen arginines in the SKDH sequence, and they should be located on twelve peptides in a tryptic digest since trypsin is insensitive to sites between Arg-Pro. Six of these arginine containing peptides were identified in the native digest but detection of the other smaller peptides proved difficult possibly due to their weak binding. LCMS of tryptic digests of PGO modified SKDH in the absence and presence of $NADP^+$ resulted in the identification of Arg-154 as a $NADP^+$ protected arginine. Furthermore, LCMS of a tryptic digest of the 90% active sample (in which only one arginine is modified) led to the identification of Arg-154 as the modified arginine, confirming that it is the primary site of modification. Thus Arg-154 is clearly hyper-reactive and has a functional role as has been shown for other hyper-reactive arginines (Patthy and Thesz, 1980). Based on the results of the protection experiment with NAD^+, it appears that this arginine is involved in binding of the 2' phosphate group of the adenine ribose.

CHAPTER 6

Involvement of a Histidine residue as a general acid/base in the catalytic mechanism of *Escherichia coli* shikimate dehydrogenase

6.1 Introduction

The role of histidine residues as general acid/base catalysts in enzyme mechanisms depends upon the special features of the imidazole group. It is a flat heterocyclic molecule structurally related to both pyrrole and pyridine and has a pKa of~7.0 (Schneider, 1978). Donor and acceptor properties are thus combined in imidazole at neutral pH making it an optimal catalyst and a buffer in biological systems (Eigen, 1964).

A large number of enzymes contain essential histidine residues at their active sites. For example, in the hydrolysis of RNA by bovine pancreatic ribonuclease, His-12 and His-119 are involved in donation and abstraction of protons (Richards and Wyckoff, 1971; Fersht, 1985) and the conversion of pyruvate to lactate is facilitated by the donation of a proton from His-195 in dogfish M4 lactate dehydrogenase (Adams *et al.*, 1970). The histidine in the catalytic triad of serine proteases abstracts a proton from a serine residue and enhances the nucleophilicity of the oxyanion to attack the carbonyl carbon of the peptide bond (Kraut, 1977; Warshel *et al.*, 1989). In the active site of the zinc protease carboxypeptidase A the imidazole groups of His-196 and His-69 serve as ligands for the binding of zinc atom which is essential for activity (Lipscomb, 1973). Histidine residues also participate in binding of substrates via hydrogen bonding and/or electrostatic interactions as in the case of carbamoyl phosphate synthetase in which His-312 is involved in the binding of glutamine (Miran *et al.*, 1991).

The role of His-195 in proton transfer in lactate dehydrogenase is shown in Fig. 6.1. The first direct evidence for this role for histidine was obtained using chemical modification with the histidine specific reagent diethylpyrocarbonate (DEPC) (Holbrook and Ingram, 1973). The involvement of a histidine residue in proton transfer is thought to be a general feature of pyridine nucleotide dependent dehydrogenases (Price and Stevens, 1989). Thus, it was expected that a histidine side chain would be involved as a general acid/base in the catalytic mechanism of *E. coli* SKDH. In preliminary experiments it was observed that *E. coli* SKDH could be rapidly inactivated by DEPC (R. Syme and J.R. Coggins, unpublished data). The present chapter confirms and extends this work and describes attempts to identify the particular histidine involved.

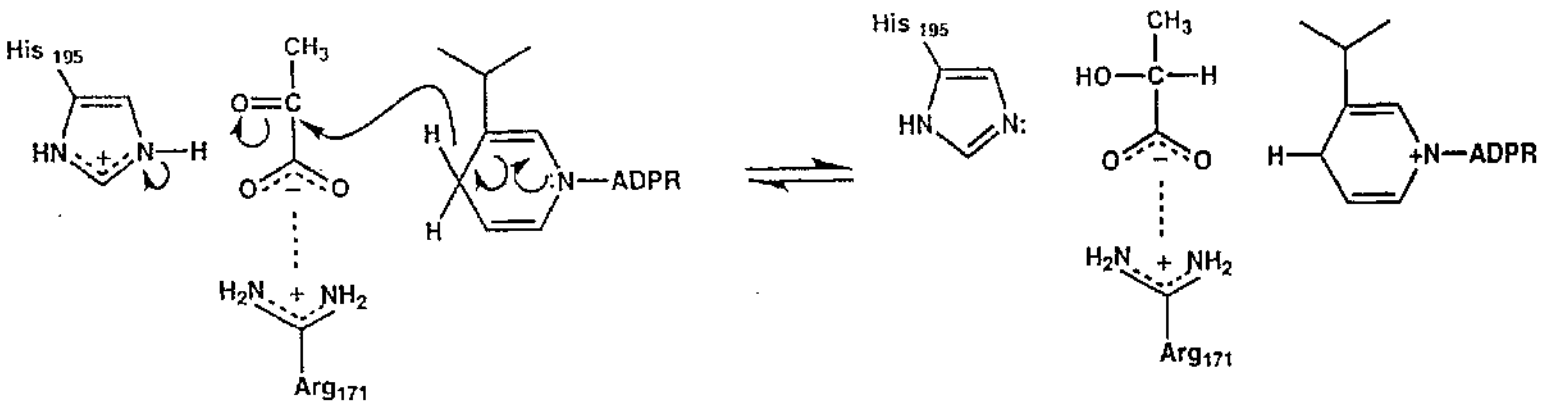


Fig. 6.1 His-195 functions as a proton-donor in the conversion of pyruvate to lactate by lactate dehydrogenase.

6.2 Chemical modification with diethylpyrocarbonate (DEPC)

6.2.1 Background

DEPC has proven to be the most commonly used reagent for selective modification of histidine residues in proteins (Lundblad and Noyes, 1984). Until the development of this reagent photo-oxidation techniques were used for histidine modification but problems persisted with specificity; methionine, tryptophan, tyrosine, serine and threonine were also sensitive to photo-oxidation in somewhat different extents (Ray and Koshland, 1962). DEPC is specific for reaction with imidazole groups of histidine residues within the pH range 5.5 to 7.5 (Lundblad and Noyes, 1984). The reaction yields an *N* -carbethoxyhistidine derivative, due to the substitution at one of the nitrogen positions of the imidazole ring (Fig. 6.2 A). This derivative has an absorption maximum between 230-250 nm, and the number of modified histidine residues can be calculated from the molar absorption difference ($\Delta\varepsilon_{240}$= 3200 M^{-1} cm^{-1})(Miles, 1977). The reaction can be reversed by addition of hydroxylamine which removes the carbethoxy group and regenerates histidine (Fig. 6.2 B)(Miles, 1977). In the presence of excess DEPC disubstitution of the imidazole ring may result in a derivative which has an absorption maximum between 220-240 nm, and the treatment of this derivative with hydroxylamine does not regenerate histidine but results in the scission of the imidazole ring (Miles, 1977).

DEPC can also react with other nucleophiles such as cysteine, tyrosine and primary amino groups (Muhlrad *et al.*, 1967) in proteins. Modification of sulfhydryl residues, although not well documented with protein bound cysteines can be determined by a decrease in free sulfhydryl groups (Lundblad and Noyes, 1984) and O-carbethoxylation of tyrosine residues can be detected by a decrease in absorbance at 278 nm (Burstein *et al.*, 1974). O-carbethoxylation can be slowly reversed by treatment with hydroxylamine (Melchior and Fahrney, 1970) but the reaction with primary amino groups yields a derivative which is stable to hydroxylamine (Lundblad and Noyes, 1984). Therefore in the modification of histidine residues with DEPC the possibility of these side reactions should not be ruled out without investigation.

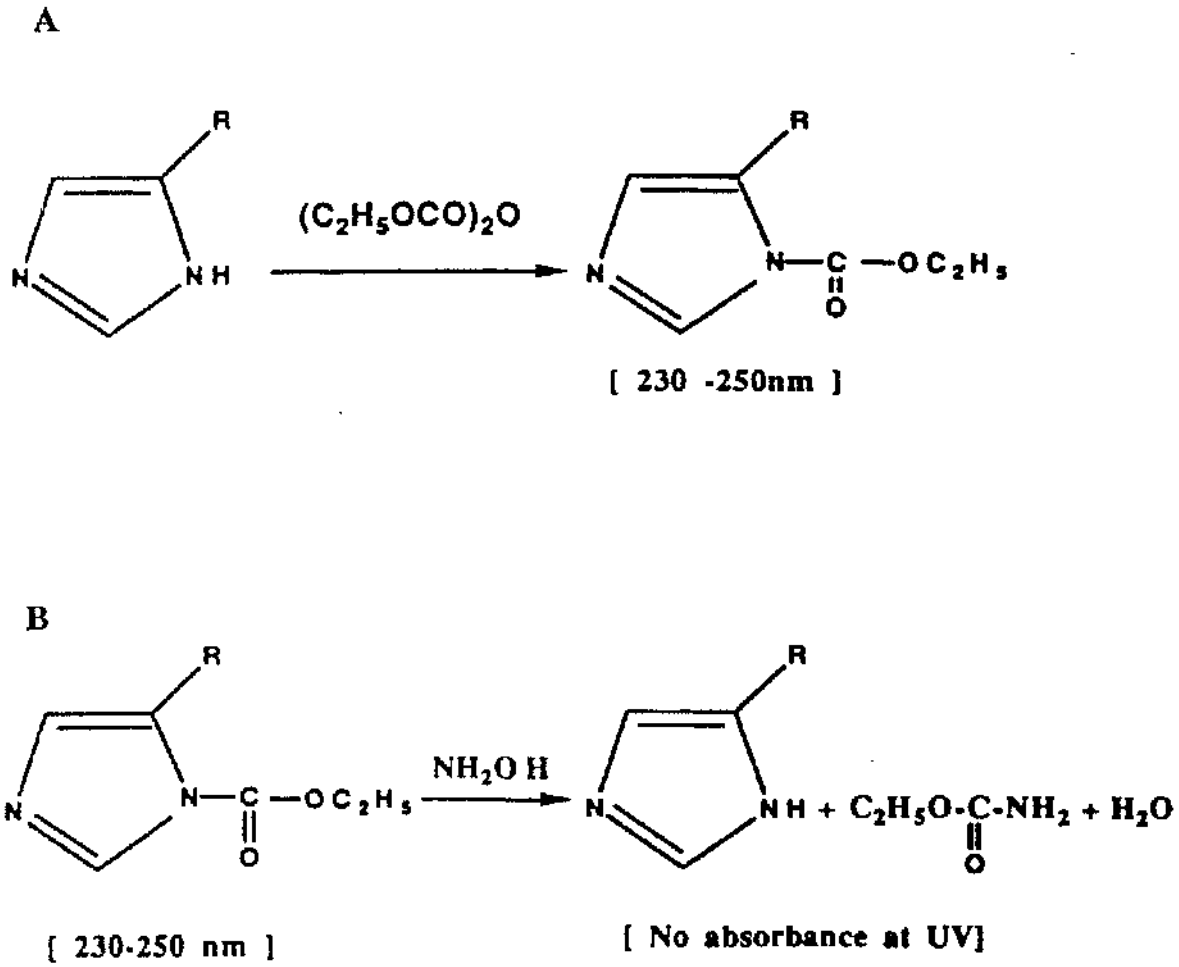


Fig. 6.2 The reaction of diethylpyrocarbonate with the imidazole side chain of histidine (A) and the reaction of hydroxylamine with carbethoxyhistidine (B)

6.2.2 Kinetics of DEPC inactivation of SKDH

Incubation of SKDH with DEPC in 50 mM sodium phosphate buffer, pH 7.0 at 25°C resulted in a rapid loss of enzyme activity. In the control experiment, in the absence of DEPC, no loss of activity was detected. The fraction of activity remaining was calculated as a percentage and plots of logarithm of percent activity remaining against time at various DEPC concentrations were linear (Fig. 6.3 A) indicating that the inactivation followed pseudo first-order kinetics. It was evident that the rate of inactivation was dependent on reagent concentration. Treatment of kinetic data and calculation of pseudo-first order rate constants were done as described in section 4.2.2. The half life values and the pseudo first-order rate constants for the inactivation process are summarised in Table 6.1.

Table 6.1

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

[DEPC] mM	$t_{1/2}$ (min)	k_{obs}/min
0.3	25.1	0.028
0.5	8.5	0.082
0.8	6.5	0.107
1.0	5.3	0.130
1.2	3.9	0.178
1.5	3.1	0.226

The second-order rate constant calculated from a plot of pseudo first-order rate constants against DEPC concentration (Fig.6.3 B) was 154 $M^{-1}min^{-1}$. The order of the reaction with respect to DEPC calculated from a plot of log k_{obs} against log [DEPC] (using equation 7 in

Fig. 6.3 Kinetics of inactivation of SKDH with DEPC

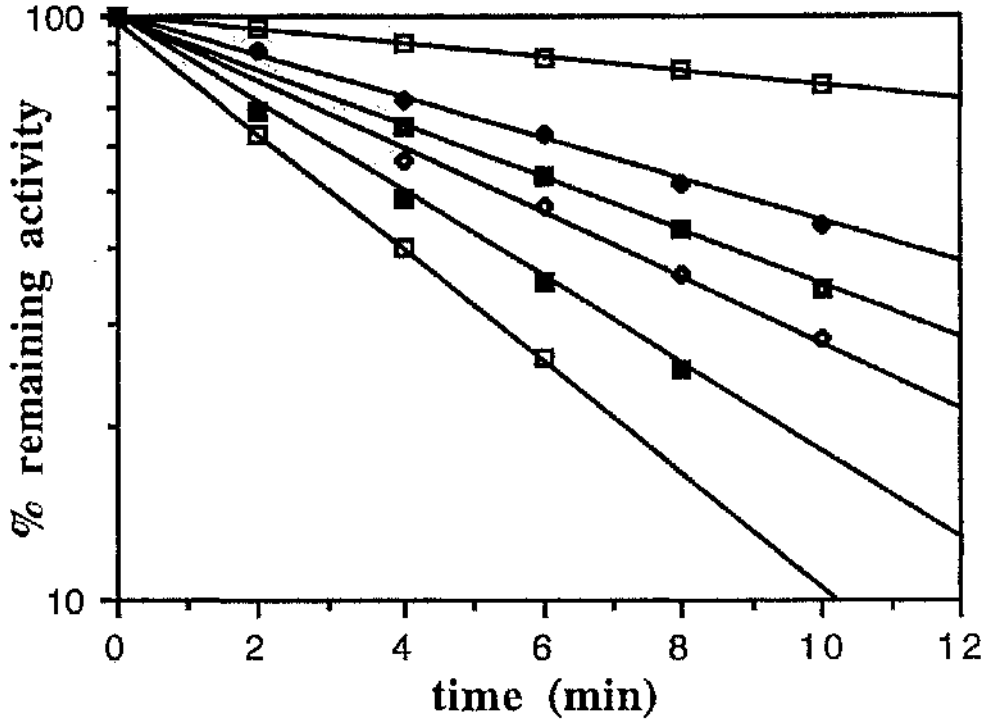
Fig. 6.3 A Pseudo first-order plots for inactivation

SKDH (2-3 μM) was incubated with increasing concentrations of DEPC in 50 mM sodium phosphate buffer, pH 7.0 at 25°C. Aliquots were removed from the reaction mixture at time intervals, quenched in imidazole and assayed for residual activity. Experimental details are given in section 3.8.2. The concentrations of DEPC used were 0.3 mM(▫), 0.5 mM(♦), 0.8mM(▪), 1.0mM(◆), 1.2mM(■) and 1.5mM(▫).

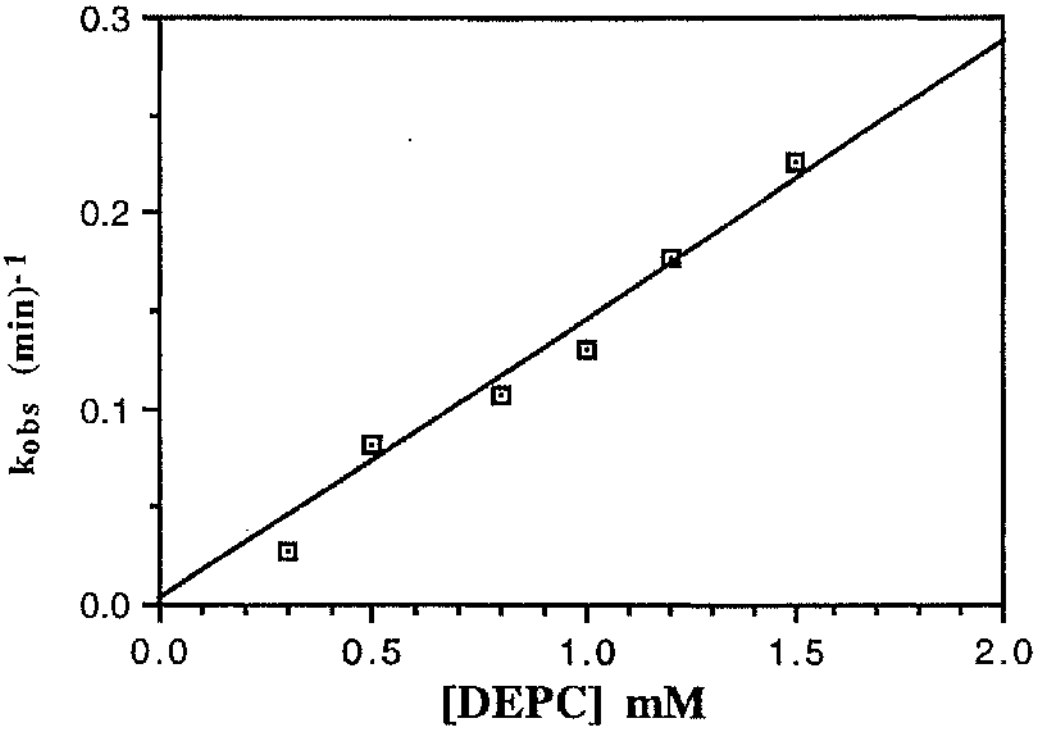
Fig. 6.3 B Determination of the second-order rate constant of inactivation

Pseudo first-order rate constants calculated in part A were replotted against DEPC concentration, and the second-order rate constant was calculated to be 154 $M^{-1}min^{-1}$ from the slope.

A


100
10
% remaining activity
0
2
4
6
8
10
12
time (min)

B


0.3
0.2
0.1
0.0
k_{obs} (min)$^{-1}$
0.0
0.5
1.0
1.5
2.0
[DEPC] mM

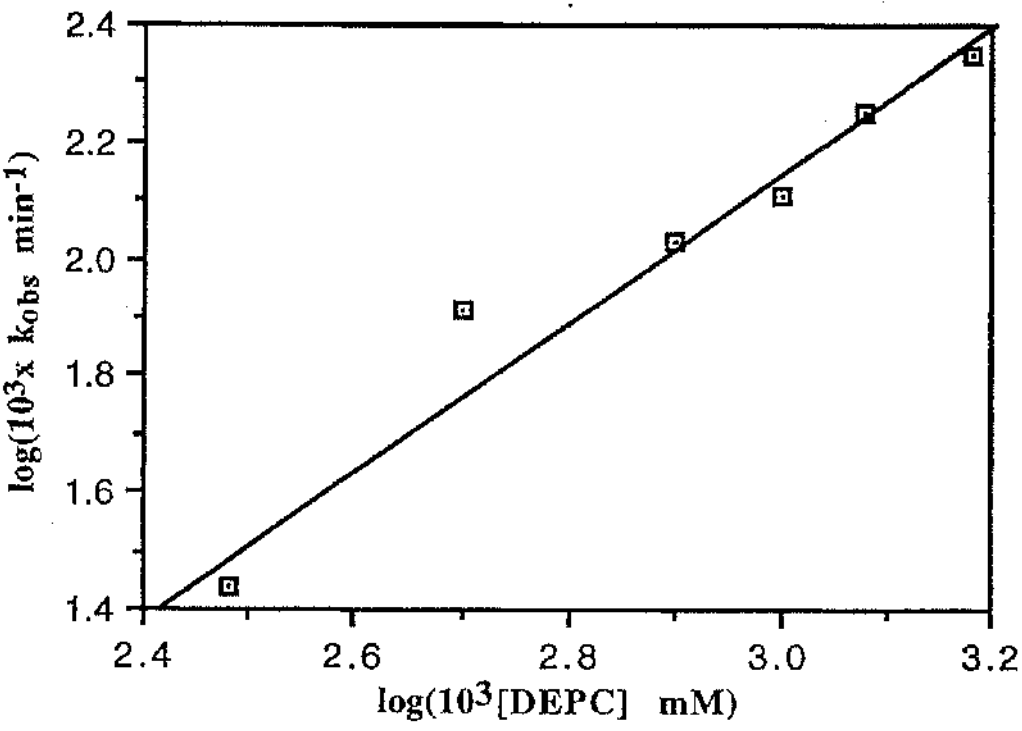


Fig 6.3 C Determination of the order of the reaction

The order of reaction with respect to DEPC was calculated from a plot of log (pseudo first-order rate constant) against log (DEPC concentration). The slope of this plot was calculated to be 1.14 indicating that the modification of a single histidine residue inactivated the enzyme.

section 4.2.2) was 1.14 indicating that modification of one essential histidine by DEPC results in the inactivation of the enzyme (Fig. 6.3 C).

6.2.3 Substrate protection against DEPC inactivation

Incubation of SKDH with DEPC in the presence of shikimate and $NADP^+$ resulted in the retention of enzyme activity (Fig. 6.4) suggesting that the inactivation was active site directed. The percentages of protection afforded were calculated using the equation in section 4.2.3. Shikimate alone afforded 54% protection whereas $NADP^+$ alone afforded 58% protection against DEPC inactivation. A maximum protection of 75% was obtained in the presence of both $NADP^+$ and shikimate. The protection afforded by shikimate was enhanced in the presence of $NADP^+$ and a similar observation was made in protection experiments against TNBS (section 4.2.3). According to these data, it appears that the proper binding of shikimate to the enzyme occurs only in the presence of the coenzyme. Thus shikimate alone is not able to afford good protection for residues involved in its binding or catalysis.

The protection afforded by shikimate and $NADP^+$ together against inactivation indicate that the essential histidine is protected from acylation when the enzyme is complexed with shikimate and $NADP^+$. In many dehydrogenases substrate reduction is facilitated by the transfer of a proton from an imidazole group of a histidine residue and a hydride ion from NAD(P)H (Fig. 6.1). In view of this and the nature of the protection afforded, it can be proposed that the essential histidine is involved in the proton transfer step of the oxidation of shikimate.

6.2.4 Characterisation of DEPC modified SKDH

DEPC reacts with cysteines, tyrosines and lysines in addition to histidines (Muhlard *et al.*, 1967). In order to investigate this possibility, characterisation of the reaction of DEPC with SKDH was attempted.

The difference absorption spectrum of DEPC modified SKDH recorded against unmodified SKDH revealed an absorption maximum at 240 nm (Fig. 6.5) which is

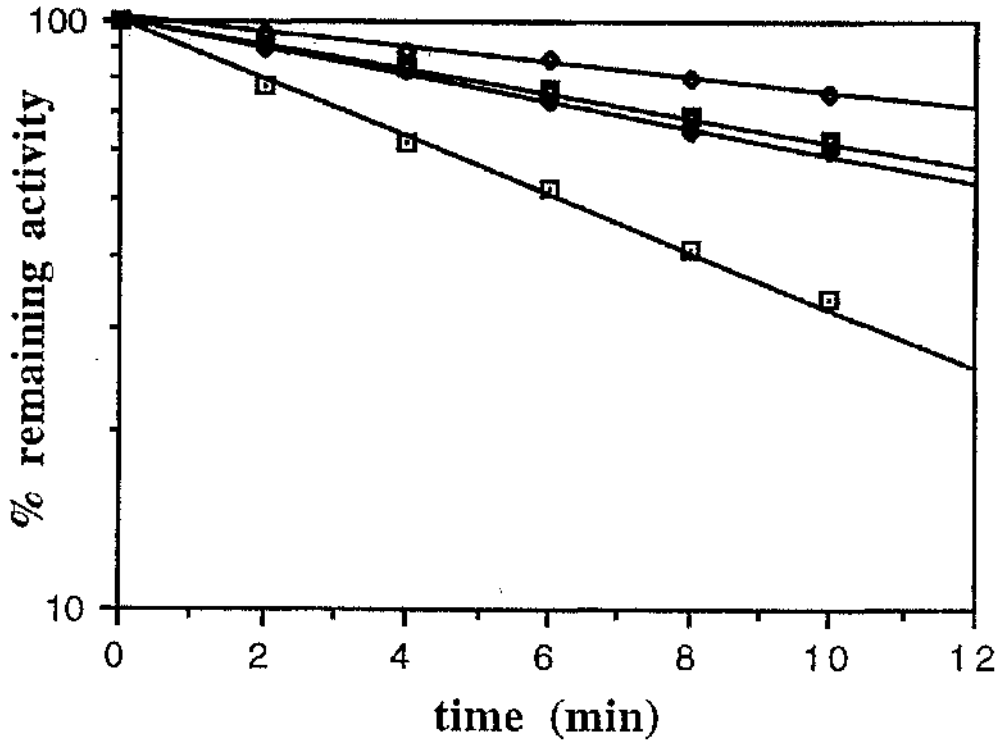


Fig 6.4 Substrate and coenzyme protection against DEPC inactivation

SKDH (2-3 μM) was incubated with 1.0 mM DEPC in 50 mM sodium phosphate buffer, pH 7.0 at 25°C alone (□), or in the presence of 4.0 mM shikimate (◆), or 4.0 mm $NADP^+$ (■) or a mixture of both in 4.0 mM concentration (◆). This figure indicates that maximum protection is obtained when shikimate and $NADP^+$ are both present in mixture.

characteristic of the carbethoxylation of histidine residues (Ovadi *et al.*, 1967; Miles, 1977). The magnitude of the maximum increased with incubation time indicating time dependent carbethoxylation. Upon addition of neutral hydroxylamine to the reaction mixture a decrease in the maximum was detected indicating decarbethoxylation. This provided evidence for the reaction of DEPC with histidine. However, O-carbethoxylation of tyrosine residues can also be reversed by treatment with hydroxylamine but the complete absence of spectral change at 278 nm ruled out the possible formation of O-carbethoxytyrosine, which would be characterised by a decrease in the difference spectrum at 278 nm ($\Delta\varepsilon_{278}$=1310 $M^{-1}cm^{-1}$)(Burstein *et al.*, 1974).

6.2.5 Correlation of activity with carbethoxylation and decarbethoxylation reaction

The extent of carbethoxylation was related to inactivation of SKDH by monitoring the activity during the reaction. With the increase in absorbance at 240 nm a concomitant decrease in activity was detected (Fig 6.6 A). This indicated that the inactivation of the enzyme was due to the carbethoxylation of histidine residues. The number of histidine residues modified were calculated from the extinction coefficient for carbethoxyhistidine ($\Delta\varepsilon_{240}$=3200 M^{-1} cm^{-1}) and complete inactivation correlated with the modification of approximately seven histidine residues (Fig. 6.6 B). These data suggested that both essential and nonessential histidine residues were being modified by DEPC.

Hydroxylamine treatment resulted in a decrease in absorbance at 240 nm and restoration of enzyme activity (Fig. 6.6 A). This indicated the regeneration of histidine residues following decarbethoxylation. However, complete restoration of activity was not possible, even after prolonged incubation. According to Miles (1977), the failure of hydroxylamine to fully reactivate a modified enzyme implies either, modification of other residues, for example lysine residues which are important for the activity or conformation of the enzyme, or the reaction of two equivalents of DEPC with one equivalent of histidine followed by a Bamberger reaction to open the imidazole ring. The latter can be ruled out as there is no marked increase in absorbance in the 230-250 nm region during hydroxylamine

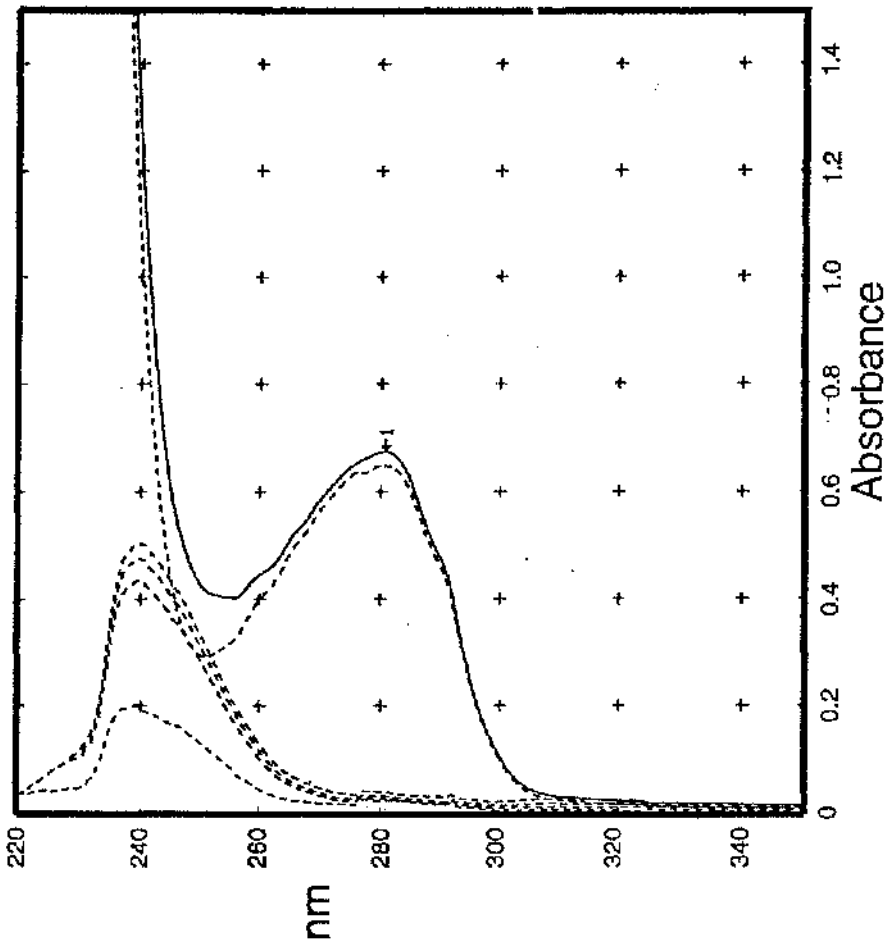


Fig. 6.5 Difference absorption spectra of DEPC modified SKDH

Difference absorption spectra of DEPC modified SKDH were recorded against unmodified SKDH. The absorption maximum at 240 nm is indicative of N-carbethoxyhistidine ($\Delta\varepsilon_{240}$=3200 M^{-1} cm^{-1}; Miles, 1977). The increase in the maximum at 240 nm relates to spectra recorded at 2min, 5 min, 10 min and 15 min after the addition of DEPC. A_{280} absorbance of unmodified SKDH(.....) and modified SKDH (____) are also shown.

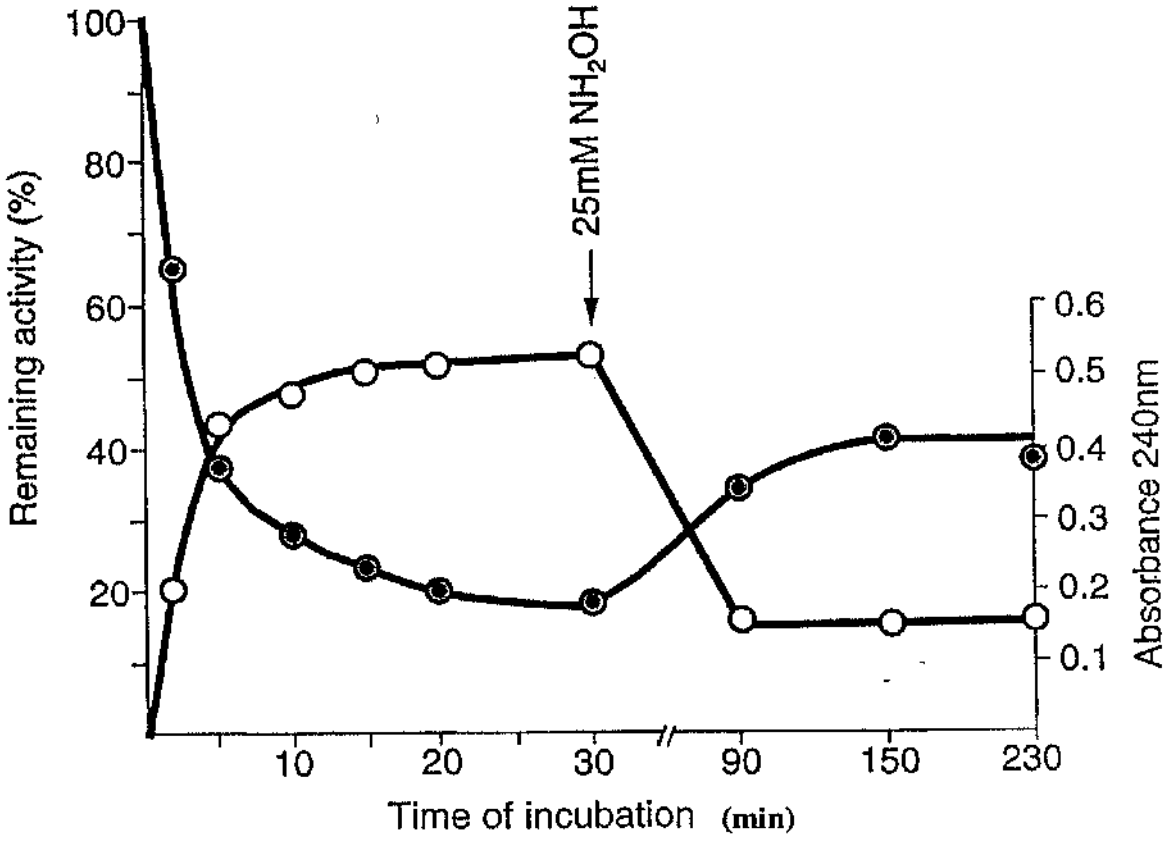
Fig. 6.6 A Correlation of activity with carbethoxylation and decarbethoxylation

With the increase in absorbance at 240 nm (o) a concomitant decrease in SKDH activity (•) was detected indicating that inactivation of the enzyme was due to carbethoxylation of histidine residues. Addition of 25 mM neutral hydroxylamine to the reaction mixture resulted in partial restoration of enzyme activity and the recovery of most of the initial absorbance at 240 nm.

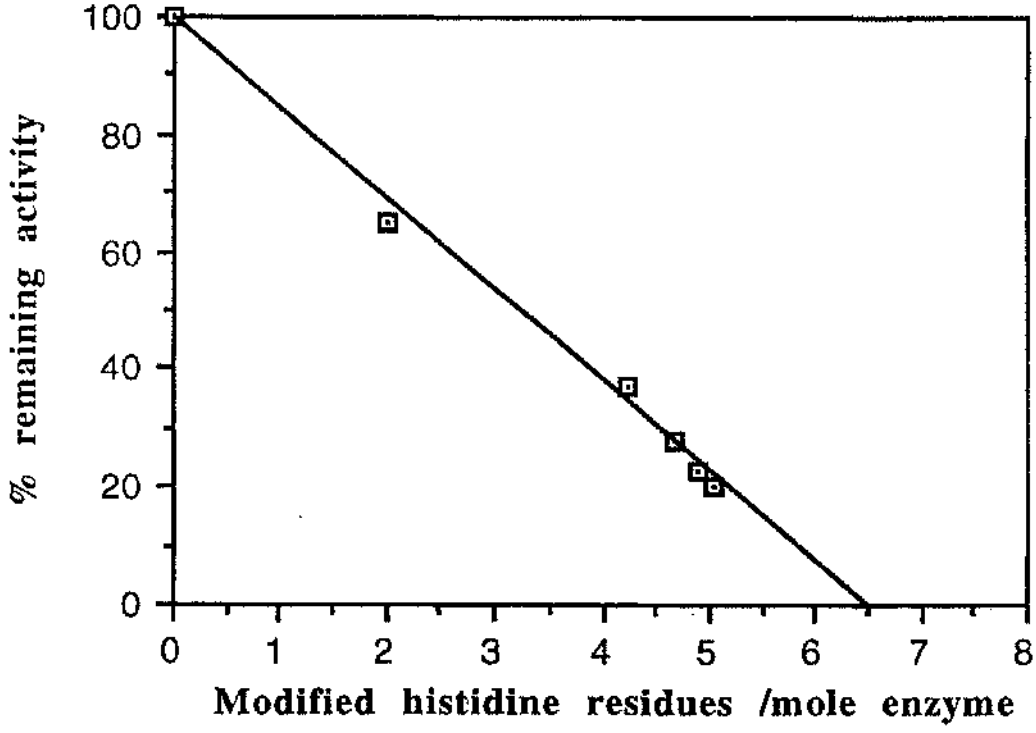
Fig.6.6 B Relationship between remaining activity and the number of histidine residues modified

The number of histidine residues modified were calculated from A_{240} ($\Delta\varepsilon_{240}$=3200 M^{-1} cm^{-1}; Miles, 1977) and were plotted against remaining activity. Extrapolation of the plot to zero activity correlated to the modification of approximately seven residues. Data were taken from **A**.

A


100
80
60
40
20
Remaining activity (%)
25mM NH_2OH
0.6
0.5
0.4
0.3
0.2
0.1
Absorbance 240nm
10
20
30
90
150
230
Time of incubation (min)

B


100
80
60
40
20
0
% remaining activity
0
1
2
3
4
5
6
7
8
Modified histidine residues /mole enzyme

treatment characteristic of the Bamberger reaction (Loosemore and Pratt, 1976). Further evidence to rule out disubstitution of the imidazole group during inactivation was obtained from the ESMS data presented in section 6.2.7. SKDH has an essential lysine at the active site (identified as Lys-65 in chapter 4) but carbethoxylation of this amino group is unlikely since e-amino groups are less effective nucleophiles at neutral pH than imidazole groups and their reaction with DEPC is reported to be very slow (Wells, 1973; Church *et al.*, 1985). Alternatively the partial recovery of activity may be due to irreversible denaturation of the enzyme by hydroxylamine, and this notion is supported further by the recovery of most of the initial absorbance at 240 nm (Fig. 6.6 A). Similar observations have been reported by Badet and Badet (1992) with *E. coli* glucosamine-6-phosphate synthase .

6.2.6 pH dependence of DEPC inactivation

The pH dependence of inactivation can provide useful information about the pK_a value of the modified residues (Takeuchi *et al.*,1986). DEPC is known to react only with the unprotonated form of the imidazole group of histidine residues (Holbrook and Ingram, 1973). The apparent pK_a of the essential histidine residue at the active site of SKDH was determined by following the pH dependence of the pseudo first-order rate constant for DEPC inactivation in phosphate buffer. The relationship between the pseudo first-order rate constant and pH can be expressed as,

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

which in linear form becomes,

$$k_{obs}(H^+) = K_a(k_{obs})_{max} - K_a k_{obs}$$

where K_a is the 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 $k_{obs}(H^+)$ against k_{obs} yielded a straight line (Fig 6.7); the $(k_{obs})_{max}$ calculated from

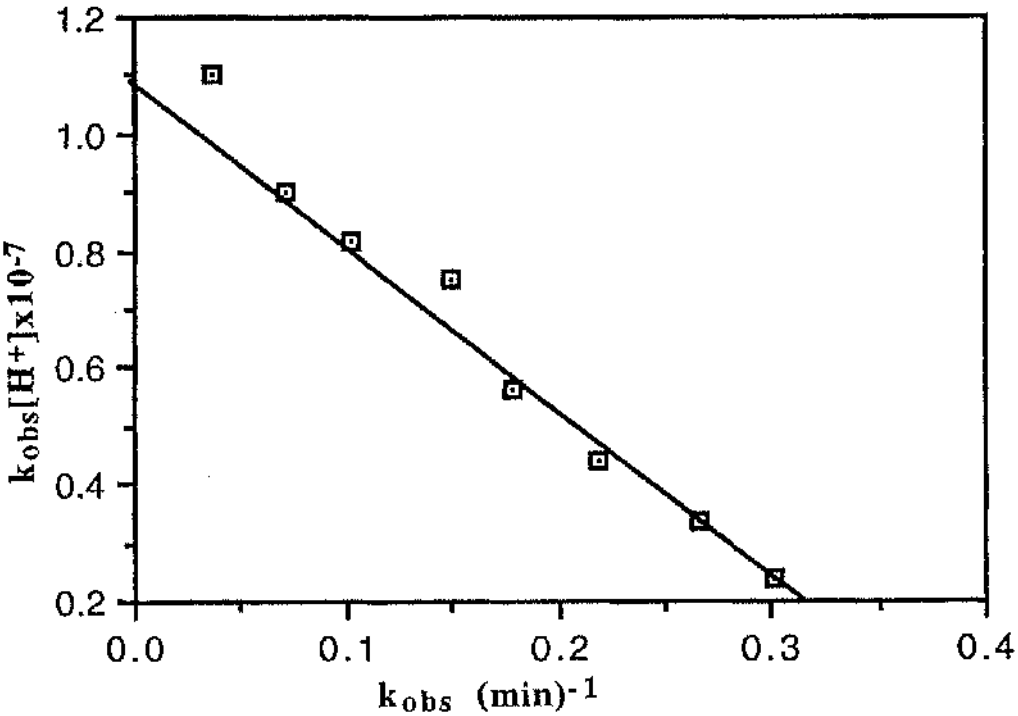


Fig. 6.7 pH dependence of DEPC inactivation

SKDH was inactivated with DEPC over the pH range 5.5-7.1. The k_{obs} values were calculated from pseudo first-order plots at various pH values and were plotted against $k_{obs}[H^+]$. The pK_a of the reacting group calculated from the slope is 6.6.

the ordinate intercept was 0.392 min^{-1} and the pK_a of the reacting group was calculated to be 6.6 from the slope. This is a typical pK_a for a histidine residue (Cosineau and Meighen, 1976; Holbrook and Ingram, 1973) and provides further evidence that DEPC inactivation of SKDH was due to the modification of a histidine residue.

6.2.7 Determination of the stoichiometry of incorporation of DEPC

The stoichiometry of incorporation of DEPC into SKDH in the absence and presence of protecting ligands was determined by ESMS in the positive ion mode as described in section 3.6.5. Preparation of unprotected and protected samples is described in section 3.8.6.

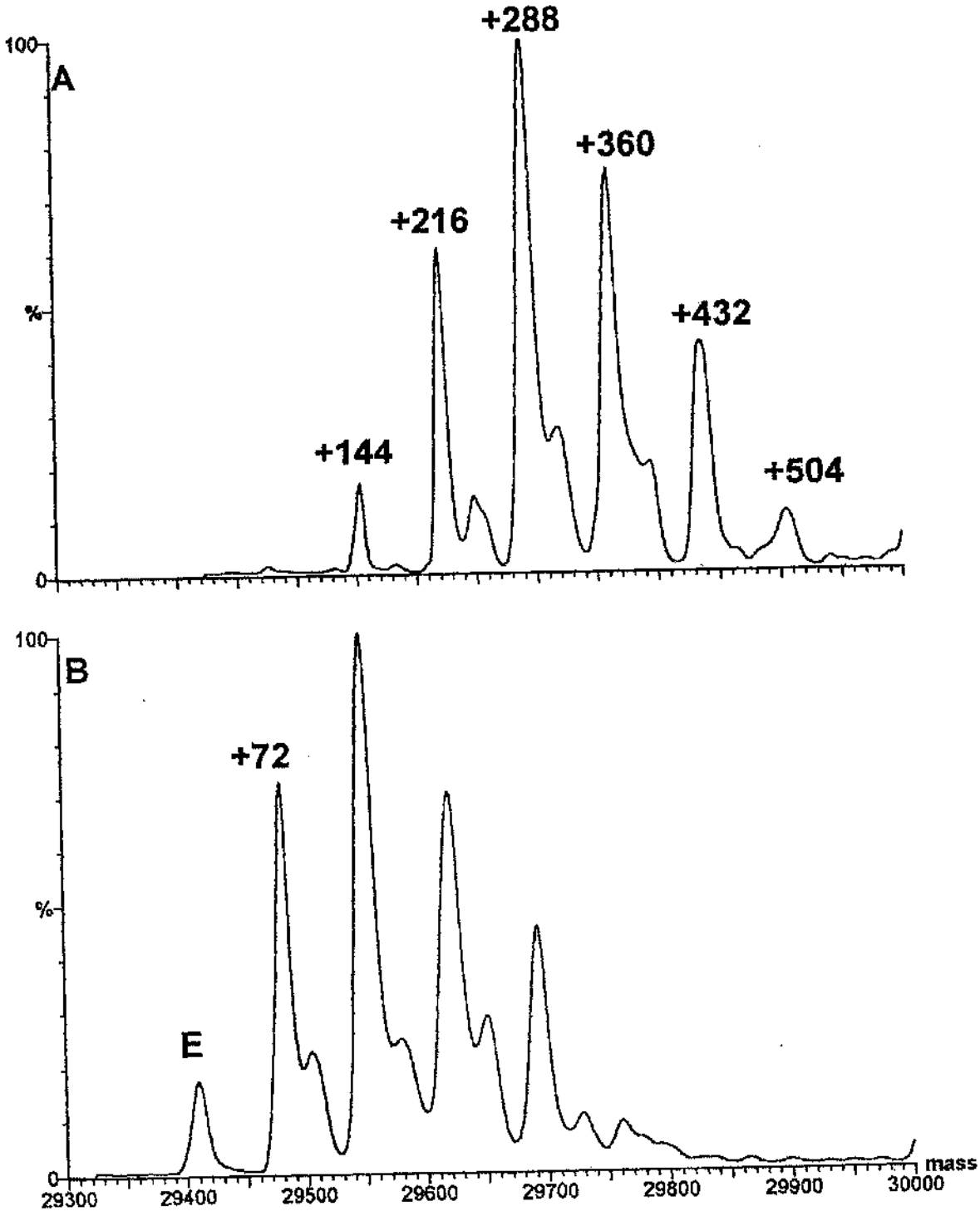
There are eight histidine residues in the primary sequence of SKDH and in the unprotected sample which was >70% inactivated, seven of these residues were modified (Fig. 6.8 A). The absence of a peak corresponding to unmodified enzyme, although the sample was not completely inactivated, indicated the modification of both essential and nonessential residues. The reaction of one equivalent of DEPC with one equivalent of imidazole would result in a mass difference of +72 Da according to Fig. 6.2 A. The significant modified species had mass differences of +144 (M_R=29558); +216 (M_R=29630); +288 (M_R=29702); +360 (M_R=29774); +432 (M_R=29846) and +504 (M_R=29918) Da relating to enzyme species with two to seven modified histidine residues respectively. These data also show that extensive di-substitution of the imidazole group had not taken place.

Although the protected sample was 65% active only a small amount of the unmodified enzyme remained (29414 Da)(Fig. 6.8 B). Significant modified species had mass differences of +72 (M_R=29486); +144 (M_R=29558); +216 (M_R=29630); +288 (M_R=29702) and +360 (M_R=29774) Da relating to the modification of one to five histidine residues respectively. Accordingly, the occupancy of the active site by shikimate and NADP+ has protected two histidine residues from DEPC modification. Thus the possibility exists, that these two histidine residues are located at the active site and their identification would contribute towards the characterisation of the active site of *E. coli* SKDH.

Fig. 6.8 Stoichiometry of incorporation of DEPC into SKDH by ESMS

A. ESMS profile of DEPC modified SKDH (>70% inactivated) in the absence of shikimate and $NADP^+$ showed a number of modified species with mass differences of +144; +216; +288; +360; +432 and +504 Da indicative of enzyme species with two to seven modified histidine residues respectively.

B. ESMS profile of DEPC modified SKDH (65% active) in the presence of shikimate and $NADP^+$ showed a minor peak corresponding to unmodified enzyme (29414 Da) and five modified species with mass differences of +72; +144; +216; +288 and +360 Da relating to enzyme species with one to five modified histidine residues. Accordingly, two residues are protected by shikimate and $NADP^+$ from DEPC modification.


100
A
%
0
+288
+360
+216
+432
+144
+504
100
B
%
0
+72
E
29300
29400
29500
29600
29700
29800
29900
30000
mass

6.3 Differential peptide mapping of DEPC modified SKDH by RP-HPLC

In order to identify the two protected histidine residues the technique of differential peptide mapping was employed, also successfully used by Vangrysperre *et al.* (1989) and Deka *et al.* (1992) to study the DEPC mediated inactivation of *Streptomyces violaceoruber* D-xylose isomerase and *E. coli* type I dehydroquinase respectively. This method relies on the fact that *N*-carbethoxyhistidine is sufficiently stable around neutral pH (Melchior and Fahrney, 1970) to permit non radioactive DEPC to be used, and labelled peptides are simply detected by their absorbance at 240 nm.

DEPC modified SKDH in the absence and presence of shikimate and $NADP^+$ (aliquots from the samples prepared in section 3.8.6) were digested using subtilisin following denaturation in 4M GdnHCl. After 2 hours of digestion complete digests were obtained which were then fractionated by reverse phase HPLC as described in section 3.8.9 using a solvent system maintained at pH 6.4. Comparison of peptide maps at 240 nm (Fig. 6.9 A-B) resulted in the identification of three peptides which were protected by shikimate and $NADP^+$. All three peptides were located in the hydrophobic region of the peptide map and had retention times of 39.3 min, 43.9 min and 48.0 min. Since DEPC is a hydrophobic reagent carbethoxylated peptides are more hydrophobic than their unmodified counterparts and this hydrophobicity shift results in an increase the retention time of the modified peptides (Biscoglio *et al.*,1986). These differences in the peptide maps were highly reproducible between separate experiments unlike the other regions where differences were less pronounced and irreproducible.

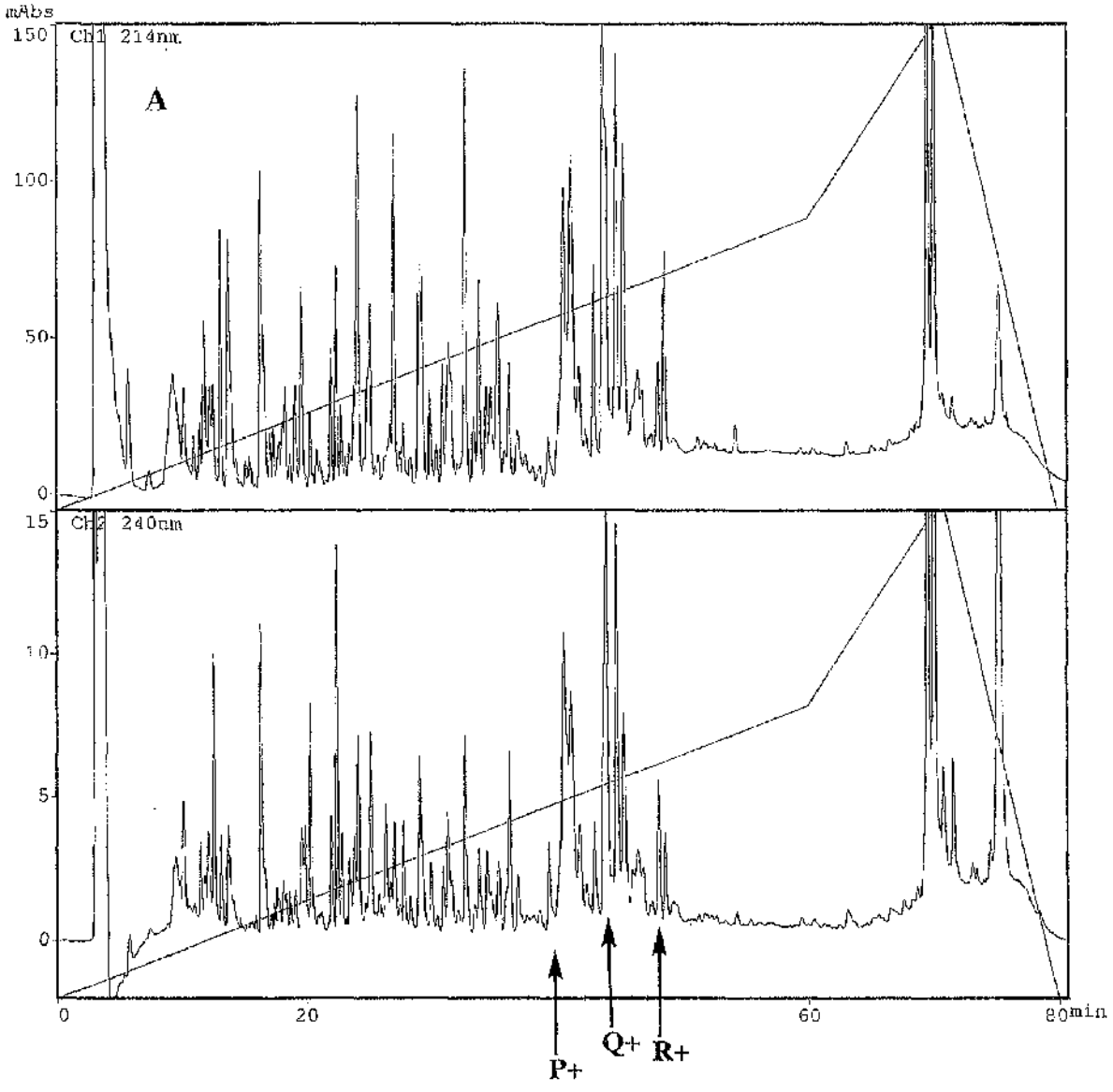
6.4 Amino-acid sequencing of the isolated peptides

The peptides eluting at 39.3 min, 43.9 min and 48.0 min (referred to as P,Q and R respectively) were collected and subjected to Edman degradation. However, no sequence was obtained for peptide Q since it comprised of a mixture of peptides, and unambiguous sequence determination was not possible. The amino-acid sequences of peptide P and R are summarised in Tables 6.2 and 6.3 respectively.

Fig. 6.9 Differential peptide mapping by RP-HPLC

HPLC profiles of subtilitic digests of DEPC modified SKDH in the absence (A) and presence (B) of shikimate and $NADP^+$

Peptide maps were monitored at 214 nm, the common wave length for peptide detection and at 240 nm for the detection of peptides with *N*-carbethoxyhistidines. In the presence of shikimate and $NADP^+$ three peptides are protected from carbethoxylation. They are designated as P+, Q+ and R+ in profile (**A**) and as P-, Q- and R- in profile (**B**). mAbs represents milli absorbance units.


mAbs
150
Ch1 214nm
A
100
50
0
15
Ch2 240nm
10
5
0
0
20
60
80min
P+
Q+
R+

B

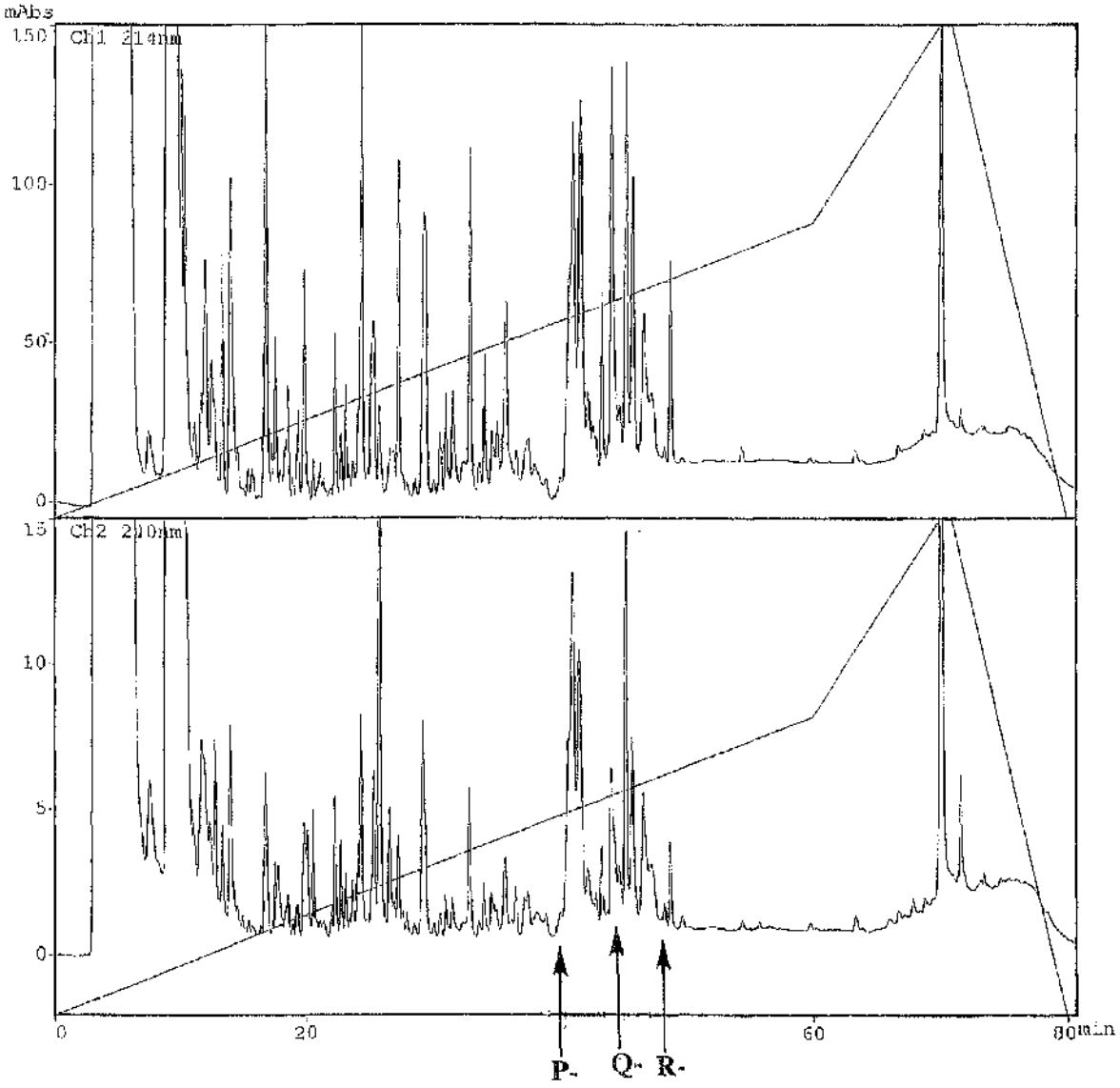


Table 6.2

Amino acid sequence of peptide P

Cycle number	Residue	Yield (pmoles)
1	Ala	50
2	Val	45
3	Phe	33
4	Gly	13
5	Asn	4
6	Pro	8
7	Ile	6
8	Ala	3
9	?	
10	?	

This sequence is incomplete and represents only a part of the full length of peptide P. The isolated peptide had to be further purified by a second HPLC run prior to sequencing. As a result very little material was available for sequencing, and because of the relatively poor sequencing yield at step 4 only eight cycles in the sequencer was possible.

The above sequence does not contain any histidine residues or any other residue that is likely to react with DEPC. However, the residue following the above sequence is a histidine and of the eight histidine residues in the *E. coli* sequence, this histidine is the most conserved. The above sequence is followed by a highly conserved stretch of residues shown in *boldfaces* below.

^{5}A V F G N P I A **H S K S P**^{17}

It is evident from these data that His-13 is the site of DEPC modification in peptide P. Thus, His-13 appears to be the essential histidine responsible for the DEPC mediated inactivation of SKDH.

Table 6.3

Amino-acid sequence of peptide R

Cycle Number	Residue	Yield (pmoles)
1	Leu	27
2	Leu	19
3	Trp	10
4	His	13
5	Gly	16
6	Val	11
7	Leu	10
8	Pro	6
9	Asp	3
10	Val	1.7
11	Glu	4
12	Pro	3
13	Val	4
14	Ile	3
15	Lys	1
16	Gln	1
17	---	---
18	---	---
19	---	---

The complete peptide R contained 16 amino acid residues and corresponds to residues 250-265 in the deduced amino acid sequence of *E. coli* . The only histidine residue within this peptide is His-253 and this must be the site of DEPC modification in peptide R. His-253 is not a conserved residue and is therefore unlikely to be a catalytically important active site residue. It's protection by substrate and coenzyme from DEPC modification could be simply due to it's presence in the active site region.

6.5 Active site homology

The above work led to the identification of the essential histidine residue as His-13. Alignment of the amino-acid sequence surrounding this residue with matching regions of other known SKDH enzymes show that it is a conserved region and His-13 is conserved in all but one sequence. In *A. nidulans* this position is occupied by a Gln residue. However the *A. nidulans* sequence has been subjected to many corrections since it's initial sequencing and a definitive confirmation of this region in the sequence is needed. It shall be noted that a single base change (A/G to C/U) would change the Gln in the *A. nidulans* sequence to a His. The sequence alignments are shown in Fig. 6.10.

6.6 The role of His-13

As described in section 6.1 histidine residues function in the proton transfer step in many pyridine nucleotide dependent dehydrogenases. On the basis of the work presented above His-13 is proposed to have this role in SKDH.

Besides chemical modification the study of the pH-dependence of enzyme-catalysed reactions can provide information about the nature of the amino-acid side chains that participate in catalysis. Because of the proposed role for His-13 in SKDH it will be expected that the activity of the enzyme would vary with the ionisation state of His-13. Histidines typically have pK_a values in the range of 6.0-7.5 and so the pH dependence of the oxidation of shikimate by SKDH was investigated over the pH range 5.5-10.8.

In this investigation, the V_{max} value at each pH studied was determined by varying the concentration of shikimate while keeping the concentration of $NADP^+$

E. coli	^{5}A V F G N P I A **H** S K S P^{17}
P.aeruginosa	C V F G N P I G **H** S K S P
B. aphidicola	A L F G N P I D **H** S Q S P
N. tabacum	G **I** I G K P V S **H** S K S P
P. sativum	G I I G K P V S **H** S K S P
S. cerevisiae	F V V G K P I G **H** S R S P
P. carinii	F L F G K P I K **H** S Q S P
A. nidulans	A I F G S P I S Q S R S P

Fig. 6.10 Sequence alignments surrounding His-13 of the *E. coli* sequence with matching regions of other known SKDH sequences. His-13 is shown in *boldfaces* and is conserved in all the sequences except the *A. nidulans* sequence. These alignments were taken from the GCG PileUp program (Devereux *et al.*, 1987) mounted on the Glasgow University Unix system.

constant. A plot of pH against log V_{max} is shown in Fig. 6.11. Extrapolation of the linear portions of the plot yielded a pK_a value of 7.3 at the point of intersection which is constant with a histidine side chain being involved in the catalytic mechanism. The pK_a of the group reacting with DEPC was determined to be 6.6 (section 6.2.6), and although these values are slightly different they are within range of pK_a values for imidazole side groups of histidine residues (Lundblad and Noyes, 1984).

These data together with the chemical modification data identifying of His-13 as the essential histidine residue, provide strong evidence to confirm the initial proposal that a histidine residue functions as a general acid/base in the catalytic mechanism of SKDH.

6.7 Discussion

Chemical modification of SKDH by DEPC at pH 7.0 resulted in time dependent inactivation of the enzyme and analysis of kinetic data revealed that the inactivation was due to the modification of one essential histidine residue. The rate constant for the inactivation process was calculated to be 154 M^{-1} min^{-1}. This value is higher than values frequently found for essential histidines in proteins (Lundblad and Noyes, 1984; Church *et al.*, 1985) and suggested that the essential histidine reacting with DEPC was highly reactive. In addition to histidines DEPC can also react with tyrosines, cysteines and lysines (Miles, 1977; Lundblad and Noyes, 1984). To investigate whether there were any side reactions of this kind difference spectra of DEPC modified SKDH were recorded. These showed an absorption maximum at 240 nm typical of *N*-carbethoxyhistidine (Miles, 1977). Carbethoxylation paralleled inactivation and activity was partially recovered following hydroxylamine treatment. The possibility of tyrosines reacting with DEPC was ruled out since there was no decrease in difference spectra at 278 nm (Burstein *et al.*, 1974; Miles, 1977). There are three cysteine residues in the primary sequence of SKDH and reaction with 5,5'-dithiobis(2-nitrobenzoate) (Ellman, 1959) in the presence and absence of 0.1% sodium dodecyl sulphate showed that two of these residues are in a buried conformation and only one sulfhydryl group is exposed in the native protein. However,

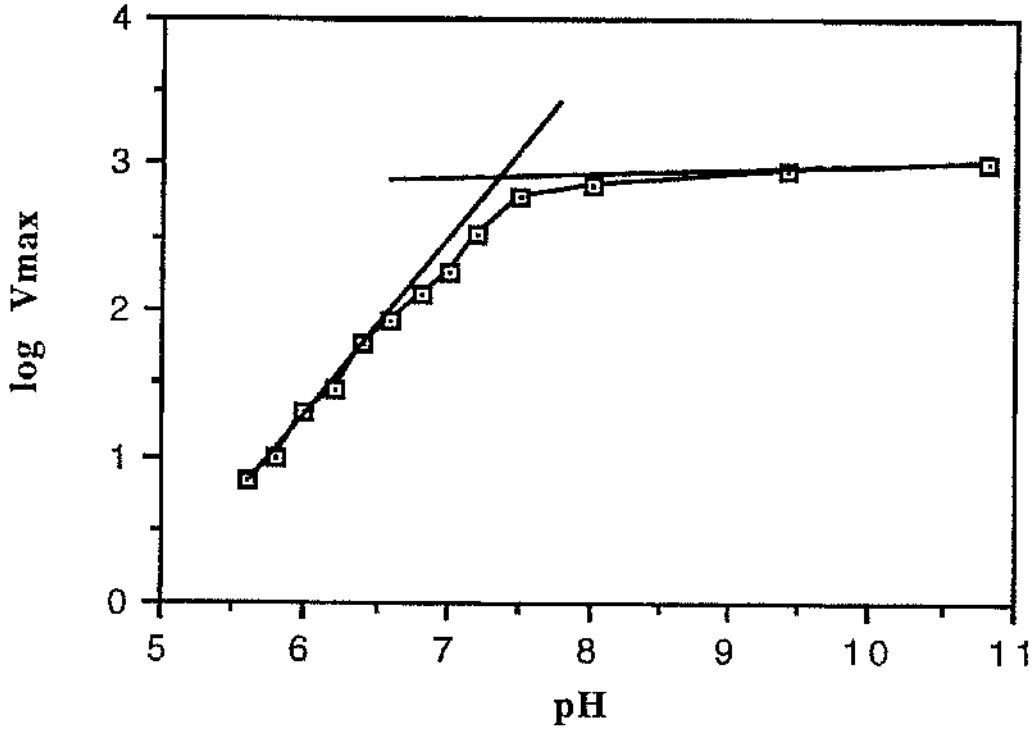


Fig. 6.11 pH dependence of the SKDH catalysed reaction

Extrapolation of the linear portions of the plot of pH against log Vmax intersect at pH 7.3 indicating that a group with a pK_a of 7.3 is involved in the catalytic reaction.

incubation of SKDH with cysteine specific reagent *N*-ethylmaleimide (Lundblad and Noyes, 1984) did not inactivate the enzyme even at high reagent concentrations, indicating the absence of any active site sulfhydryl groups. Therefore any contribution towards DEPC inactivation of SKDH by a modified cysteine was very unlikely. The reaction of reactive lysines with DEPC is reported to have very low rates (Wells, 1973; Church *et al.*, 1985) and a second-order rate constant of 154 M^{-1} min^{-1} is much too high for a typical lysine residue. The pH dependence of inactivation demonstrated that the inactivation was due to the modification of a group with a pK_a of 6.6 which was also consistent with the essential residue modified by DEPC being a histidine residue.

Complete inactivation of the enzyme correlated with the modification of approximately seven histidine residues, calculated from the molar absorption difference for *N*-carbethoxyhistidine ($\Delta\varepsilon_{240}$=3200 M^{-1} cm^{-1}; Miles, 1977). This number is in good agreement with the ESMS data where the maximum stoichiometry of incorporation in an extensively modified sample (>70% inactivated) was determined to be seven. In the protected sample only five histidines were modified as revealed by ESMS and therefore two histidines were inaccessible to DEPC in the presence of shikimate and $NADP^+$.

Attempts to identify the sites of DEPC modification by LCMS were unsuccessful. The most widely used buffer system for the fractionation of DEPC modified peptides is sodium/potasium phosphate at neutral pH (Vangrysperre *et al., 1989*; Deka *et al.*, 1992). Since LCMS operates on volatile buffers these buffers could not be used, alternatively a phosphoric acid based buffer system at pH 4.5 was employed. LCMS of a chymotryptic digest of DEPC modified SKDH resulted in the identification of two modified histidine residue; His-164/179 and His-253. The limited stability of the carbethoxylated histidine residues of proteins is a main problem encountered in their identification (Holbrook and Ingram, 1973; Burstein *et al.*, 1974). Therefore, failure to identify other modified histidine residues by LCMS can be attributed to the instability of the carbethoxyhistidine residues at pH 4.5. However, cases have been reported where *N*-carbethoxyhistidines had been sufficiently stable around neutral pH to permit their isolation (Vangrysperre *et al.*, 1989; Deka *et al.*, 1992). Hence, the alternative approach of

differential peptide mapping using sodium phosphate buffer, pH 6.4 was employed to identify the peptides containing the two protected histidine residues.

Sequencing of the substrate protected peptides identified His-13 and His-253 as the two protected histidine residues. His-253 is not a conserved residue and it's protection by substrate may simply reflect it's presence in the active site region. His-13 is a highly conserved residue and appears to be the essential histidine responsible for the DEPC mediated inactivation of SKDH. The pH-dependence of the enzyme catalysed reaction implicated the involvement of a histidine residue in the catalytic mechanism. Based on a general mechanism for the bond forming events of pyridine nucleotide dependent dehydrogenases (Bloxham *et al.*, 1975), a reaction scheme involving His-13 for the conversion of 3-dehydroshikimate to shikimate by SKDH can be proposed as shown in Fig. 6.12. Full confirmation of this role for His-13 must await site-directed mutagenesis and eventually three dimensional structure determination.

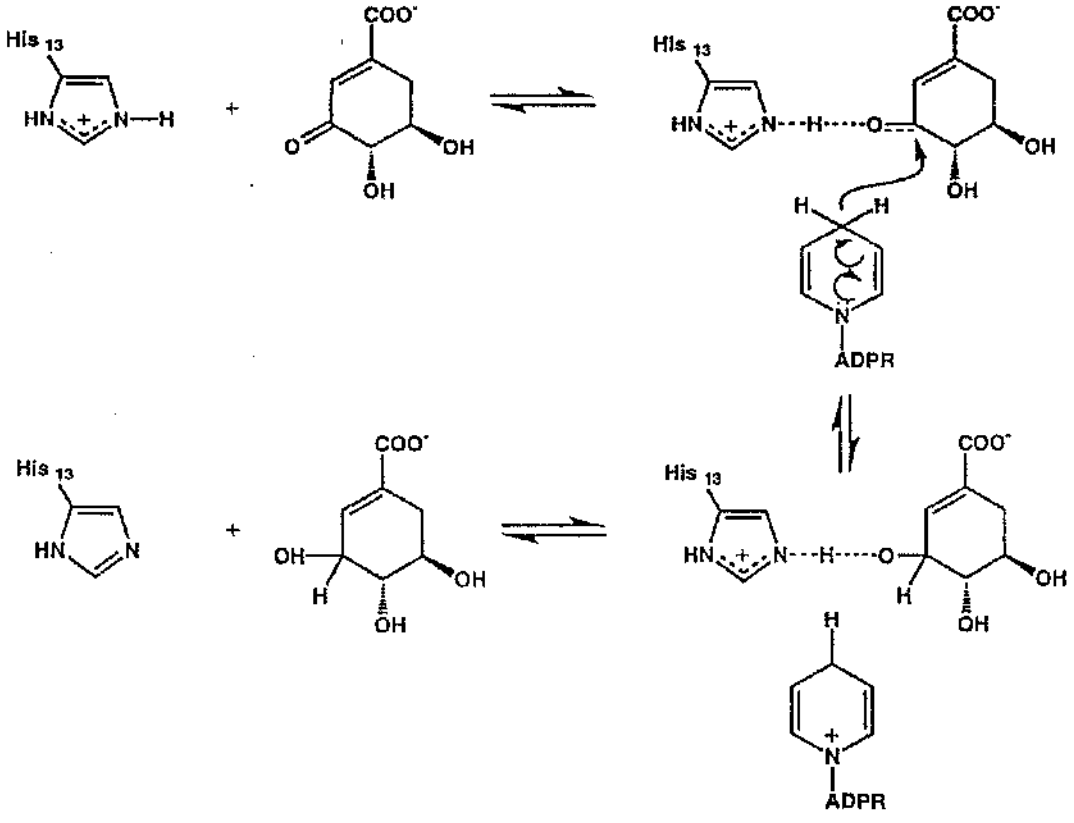


Fig. 6. 12 The proposed reaction scheme for SKDH involving His-13 as a general acid/base

The substrate (3-dehydroshikimate) is activated by a hydrogen bond formation between substrate carbonyl and protonated form of His-13 (for the forward reaction). This polarises the substrate carbonyl group and facilitates hydride transfer from NADPH after which complete protonation of the carbonyl group occurs. Similarly, in the reverse reaction the hydroxyl hydrogen of shikimate forms a bond with the unprotonated form of His-13 facilitating hydride removal.

CHAPTER 7

General discussion and future prospects

7.1 The active-site of *E. coli* SKDH

The preceding chapters have described in detail the progress made towards the characterisation of the active-site of *E.coli* SKDH. It has been shown that;

(I) Lys-65 mediates the binding of shikimic acid via an interaction between it's ε-amino group and the carboxyl group of shikimate.

(II) Arg-154 is a component of the $NADP^+$ binding site, and interacts with the 2' phosphate group of the adenosine moiety.

(III) His-13 is probably involved in proton transfer as a general acid/base facilitating substrate reduction/oxidation.

Thus, these findings can be summarised in a schematic model of the proposed active-site of *E. coli* SKDH as shown below (Fig. 7.1).

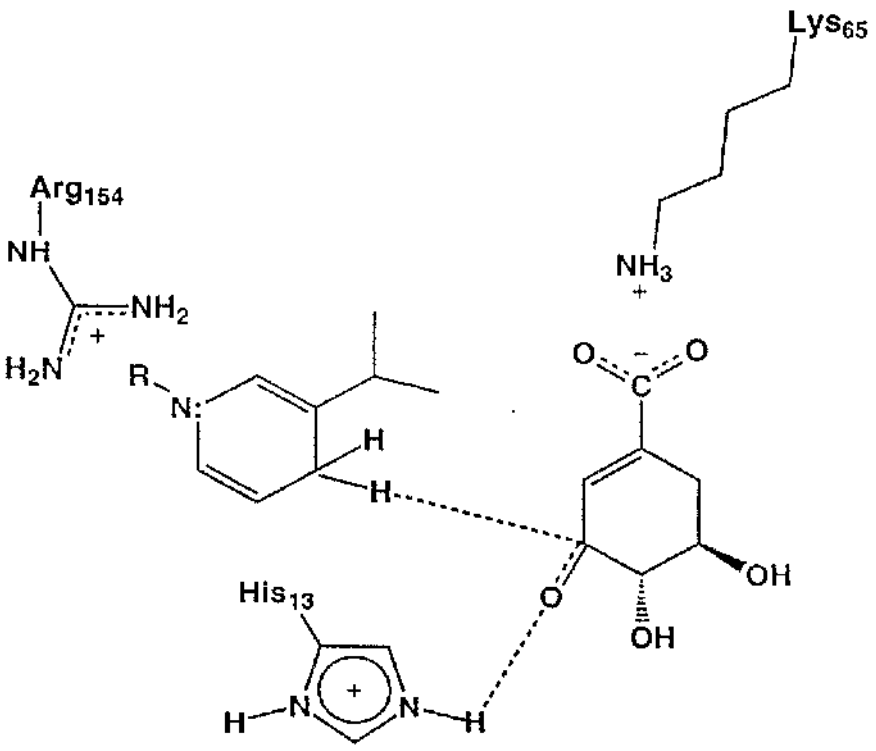


Fig. 7.1 Schematic model of the proposed active-site of *E. coli* SKDH containing the residues identified in this project.

7.2 Structural and mechanistic features of *E. coli* SKDH

In the absence of the crystal structure chemical modification studies together with sequence data can give useful information about the structural and mechanistic features of an enzyme molecule. SKDH is monomeric in structure, and contains 272 amino-acid residues in the sequence. Alignment of all known amino-acid sequences of SKDH from various organisms (see appendix) shows that most of the conserved residues are located in the N-terminal region of the polypeptide chain. The identification of His-13 and Lys-65 as residues involved in catalysis and substrate binding respectively, suggest that this region of the polypeptide chain is involved in substrate binding and catalysis. Furthermore, the presence of Lys-15 in this region and it's proximity to His-13 explains the partial nature of protection afforded by substrate and coenzyme against TNBS modification. In the case of His-253, which from the protection studies must be at or near the active-site, could be present on a loop of the polypeptide chain which in the folded structure comes close to the active-site.

The 'fingerprint' characteristic of the ADP binding βαβ fold lies between residues 121 and 151 (Anton and Coggins, 1988). The α helix dipole of this fold is known to interact with the pyrophosphate group of the coenzyme and facilitate it's binding. In enzymes that bind $NADP^+$ by this fold, the final position of this fingerprint should not be a negatively charged residue in order accommodate the 2' phosphate group of the adenosine moiety (Wierenga *et al.*, 1986), and a positively charged side chain nearby would provide a binding site for the adenosine phosphate (Branden and Tooze, 1991). Thus in SKDH which is $NADP^+$ specific, the final position of the fingerprint is occupied by Thr-151 and two residues further in the sequence is Arg-154, which was identified as the residue interacting with the 2' phosphate group of $NADP^+$. Similar situations are encountered in other $NADP^+$ specific enzymes; for example in the *E. coli* glutathione reductase the final position in the fingerprint is occupied by Val-197 and Arg-198 is involved in binding the adenosine phosphate (Scrutton *et al.*, 1990) .

Many dehydrogenases follow a compulsory order mechanism, where the enzyme complexes with the coenzyme to form a binary complex which then forms a ternary

complex with the substrate. In protection experiments against TNBS and DEPC inactivation shikimate alone afforded partial protection, which was enhanced in the presence of $NADP^+$. This may imply that in the absence of $NADP^+$, shikimate forms a weak complex with the enzyme. The binding of the coenzyme at the enzyme surface may generate a second binding site for the substrate which did not previously exist (Takenaka and Schwert, 1955). Thus, the possibility exists, that shikimate binds at two sites one of which is generated by the binding of $NADP^+$. Alternatively, the binding of the coenzyme may alter the affinity of the pre-existing substrate binding site. Therefore, it appears that the proper binding of shikimate is facilitated only in the presence of the coenzyme. Moreover, product inhibition and isotope exchange studies on *P. sativum* SKDH have shown that it functions through an ordered mechanism (Balinsky *et al.*, 1971). In the protection experiments mentioned above, $NADP^+$ alone afforded significant protection although Lys-65 and His-13 are residues involved in shikimate binding and catalysis. Therefore, it is possible that binding of $NADP^+$ induces a conformational change in the protein molecule and this may also support the notion that a second binding site for shikimate is generated upon $NADP^+$ binding. Alternatively, the protection afforded by $NADP^+$ against TNBS and DEPC inactivation could be simply due to steric hindrance in the active site pocket.

7.3 Future prospects

As a part of a complete structural and mechanistic study of *E. coli* SKDH, the main objective of the project was to study the active-site. In fulfilling this objective, three active-site residues have been identified and a knowledge of how SKDH converts 3-dehydroshikimate to shikimate has been obtained. This work should be followed by site-directed mutagenesis of the identified residues, to confirm their roles in the mechanism.

On the mechanistic side, the next step would be to determine the binding order of the mechanism and to measure substrate binding constants. The first of these could be done by product inhibition studies while the second would require equilibrium dialysis or microcalorimetry. The main limitation in doing product inhibition studies is the non-availability of the natural substrate, 3-dehydroshikimate. This compound is not stable

enough to be stored over a period of time and has to be synthesised from 3-dehydroquinate immediately before use. Equilibrium dialysis require radio labelled reagents which are also not readily available. Therefore, microcalorimetry would be the preferred method to study the ligand binding in the case of SKDH.

Having obtained a knowledge of the active site structure and the catalytic mechanism, the next step would be to determine the three dimensional structure of SKDH. The first prerequisite for solving the three dimensional structure by X-ray crystallography is a well-ordered crystal that will diffract X-rays strongly. Crystal growth depends on many parameters such as pH, protein concentration, nature of solvent and precipitant and the presence of added ions or ligands. Therefore, in a preliminary experiment all these parameters were screened and the optimal conditions for the crystallisation of SKDH were determined. Microcrystals of SKDH were obtained by the sitting drop method, in 5 mg/ml protein concentration and a precipitant solution containing 18-20% ammonium sulphate in 50 mM bis-tris/HCl buffer pH 6.0-6.2. This experiment was repeated using identical conditions except that 2% PEG-600 was added to the precipitant solution. This resulted in slow growth and improved the size of the crystals, which are shown in Fig. 7.2. Thus, the combination of parameters suitable for the crystallisation of SKDH has been established. The next step, will be to obtain larger crystals which can diffract X-rays, hopefully by refinement of the conditions mentioned above.

Fig 7.2 Crystals of *E. coli* SKDH formed by the sitting-drop method

REFERENCES

Adams, M. J., Ford, G.C., Koekoek, R., Lentz Jr., P.J., McPherson Jr., A., Rossmann, M.G., Smiley, I.E., Schevitz, R.W. and Wonacott, A.J. (1970) *Nature* **227**, 1098-1103

Adams, M.J., Buehner, M., Chandrasekhar, K., Ford, G.C., Hackert, M.L., Liljas, A., Rossman, M.G., Smiley, I.E., Allison, W.S., Everse, J., Kaplan, N.O. and Taylor, S.S. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1968-1972

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

Alberg, D.G. and Bartlett, P.A. (1989) *J. Am. Chem. Soc.* **111**, 2337-2338

Amrhein, N., Schab, J. and Steinrucken, H.C. (1980) *Naturwissenschaften* **67**, 356-357

Anderson, K.S., Sikorski, J.A. and Johnson, K.A. (1988a) *Biochemistry* **27**, 7395-7406

Anderson, K.S., Sikorski, J.A., Benesi, A.J. and Johnson, K.A. (1988b) *J. Am. Chem. Soc.* **110**, 6577-6579

Anton, D.L., Hedstrom, L., Fish, S.M. and Abeles, R.H. (1983) *Biochemistry* **22**, 5903-5908

Anton, I. and Coggins, J.R. (1984) *Biochem. Soc. Trans.* **12**, 275-276

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

Aronson, B.D., Somerville, R.L., Epperly, B.R. and Dekker, E.E. (1989) *J. Biol. Chem.* **264**, 5226-5232

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

Baker, P.J., Britton, K.L., Engel, P.C., Farrants, G.W., Lilley, K.S., Rice, D.W. and Stillman, T.J. (1992) *Proteins: Structure, Function and Genetics* **12**, 75-86

Balasubramanian, S., Coggins, J.R. and Abell, C. (1995) *Biochemistry* **34**, 341-348

Balinsky, D. and Davies, D.D. (1961a) *Biochem. J.* **80**, 292-296

Balinsky, D. and Davies, D.D. (1961b) *Biochem. J.* **80**, 296-300

Balinsky, D., Dennis, A.W. and Cleland, W.W. (1971) *Biochemistry* **10**, 1947-1952

Banerji, S., Wakefield, A.E., Allen, A.G., Maskell, D.J., Peters, S.E. and Hopkin, J.M. (1993) *J. Gen. Microbiol.* **139**, 2901-2914

Barca, J.L. and Giles, N.H. (1978) *Biochim. Biophys. Acta* **524**, 1-14

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

Bender, S. L., Widlanski, T. and Knowles, J.R. (1989) *Biochemistry* **28**, 7560-7572

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

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

Beynon, R. J. and Bond, J.S. (1990) In *Proteolytic enzymes*, IRL Press, Oxford University Press, England

Bhat, K.V., Bhat, S.R. and Chandel, K.P.S. (1992) *J. Horticult. Sci.* **67**, 737-743

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

Bloxham, D.P., Giles, I.G., Wilton, D.C. and Akhtar, M. (1975) *Biochemistry* **14**, 2235-2241

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

Bonner, C.A. and Jensen, R.A. (1994) *Biochem. J.* **302**, 11-14

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

Borders, C.L., Jr. and Riordan, J.F. (1975) *Biochemistry* **14**, 4699-4704

Bordier, P, Condron, R., Traeger, J.C., Mcauliffe, J.C. and Stone, B.A. (1992) *J. Biol. Chem.* **267**, 25059-25066

Boudet, A.M., Graziana, A. and Ranjeva, R. (1985) In *The Biochemistry of Plant Phenolics* (C.F. Van Sumere and P.J. Lea, eds.) *Annual Proceedings of the Phytochemical Society of Europe* 25, Clarendon Press, Oxford

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

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

Branden, C.I., Eklund, H., Nordstrom, B., Boiwe, T., Soderlund, G., Zeppezauer, E., Ohlsson, I. and Akeson, A. (1973) *Proc. Nat. Acad. Sci. U.S.A.* **70**, 2439-2442

Branden, C. and Tooze, J. (1991) In *Introduction to Protein Structure*, Garland Publishing, Inc., New York

Braunstein, A.E. (1970) *The Enzymes* (P.D. Boyer ed.) 3rd ed., Vol.IX, pg 379-480, Academic Press, New York

Bruce, N.C. and Cain, R.B. (1990) *Arch. Microbiol.* **154**, 179-186

Buehner, M., Ford, G.C., Moras, D., Olsen, K.W. and Rossmann, M.G. (1973) *Proc. Nat. Acad. Sci. U.S.A.* **70**, 3052-3054

Buehner, M., Ford, G.C., Moras, D., Olsen, K.W. and Rossmann, M.G. (1974) *J. Mol. Biol.* **90**, 25-49

Bugg, T.D.H., Abell, C. and Coggins, J.R. (1988) *Tetrahedron Lett.* **29**, 6779-6782

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

Byng, G.S. and Turner, J.M. (1977) *Biochem. J.* **164**, 139-145

Cain, R.B. (1972) *Biochem. J.* **127**, 15p

Carr, S.A., Hemling, M.E., Bean, M.F. and Roberts, G.D. (1991) *Anal. Chem.* **63**, 2802-2824

Casati, R., Bealc, J.M. and Floss, H.G. (1987) *J. Am. Chem. Soc.* **109**, 8102- 8104

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

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

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

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

Chaudhuri, S., Anton, I.A. and Coggins, J.R. (1987) *Methods in Enzymology* **142**, 315-320

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

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

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

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

Conn, E.E. (1986) In *The Shikimic Acid Pathway*, Plenum Press, New York

Cotton, F.A., Hazen, E.E., Jr., Day, V.W., Larsen, S., Norman, J.G., Jr., Wong, S.T.K. and Johnson, K.H. (1973) *J. Am. Chem. Soc.* **95**, 2367-2369

Cousineau, J. and Meighen, E. (1976) *Biochemistry* **15**, 4992-4500

Covey, T.R., Huang, E.C. and Henion, J.D. (1991) *Anal. Chem.* **63**, 1193-1200

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

Dalziel, K. (1963) *J. Biol. Chem.* **238**, 2850-2858

Dalziel, K. (1975) In *The Enzymes*, (P.D. Boyer ed.), (3rd ed.), Vol.XI, Academic

Press, New York

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

Dardenne, G.A., Larsen, P.O.and Wieczorkowska, E. (1975) *Biochim. Biophys. Acta* **381**, 416-423

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

Davies, G.M., Barrett-Bee, K.J., Jude, D.A., Lehan, M., Nichols, W.W., Pinder, P.E., Thain, J.L., Watkins, W.J. and Wilson, G. (1994) *Antimicrobial agents and Chemotherapy* **38**, 403-406

DeFeyter, R. and Pittard, J. (1986a) *J.Bacteriol.* **165**, 226-232

DeFeyter, R. and Pittard, J. (1986b) *J.Bacteriol.* **165**, 331-333

DeFeyter, R., Davidson, B.E. and Pittard, J. (1986) *J.Bacteriol.* **165**, 233-239

Degani, C., Beiles, A., Elbatsri, R., Goren, M. and Gazit, S. (1995) *J. Amer. Soc. Hort. Sci* . **120**, 307-312

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

Deka, R.K., Anton, I.A., Dunbar, B. and Coggins, J.R. (1994) *FEBS Lett* **349**, 397-402

DeLeo, A.B. and Sprinson, D.B. (1968) *Biochem. Biophys. Res. Commun.* **32**, 873-877

Dennis, A.W. and Balinsky, D. (1972) *Int. J. Biochem.* **3**, 93-102

Devereux, J., Haeberli, P. and Marquess, P. (1987) *The Program Manual for the Sequence Analysis Software Package*, Genetics Computer Group, Madison , WI

Doy, C.H. and Brown, K.D. (1965) *Biochim. Biophys. Acta* **104**, 377-389

Duncan, K., Lewendon, A. and Coggins, J.R. (1984a) *FEBS Lett.* **170**, 59-63

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

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

Eigen, M. (1964) *Angew. Chem. Int. Ed. Engl.* **3**, 1

Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T. and Branden, C.I. (1974) *FEBS Lett.* **44**, 200-204

Eklund, H., Plapp, B.V., Samama, J.P. and Branden, C.I. (1982) *J.Biol. Chem.* **257**, 14,349-14,358

Ellman, G.L. (1959) *Arch. Biochem. Biophys.* **82**, 70- 77

Elsemore, D.A. and Ornston, L.N. (1994) *J. Bacteriol.* **176**, 7659-7666

Ely, R. and Pittard, J. (1979) *J. Bacteriol.* **138**, 933-943

Erickson, W.R. and Gould, St. J. (1987) *J. Am. Chem. Soc.* **109**, 620-621

Essar, D.W., Eberly, L., Hadero, A. and Crawford, I.P. (1990) *J. Bacteriol.* **172**, 867-883

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

Ewart, C.D.C., Jude, D.A., Thain, J.L. and Nichols, W.W. (1995) *Antimicrobial agents and Chemotherapy* **39**, 87-93

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

Eyzaguirre, J. (1987) In *Chemical modification of enzymes*, Ellis Horwood Ltd., West Sussex, England

Fenn, J.B, Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1989) *Science* **246**, 64-71

Ferridge, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A. and Skilling, J. (1992) *Rapid Commun. Mass Spectrom.* **6**, 707-711

Fersht, A. (1985) *Enzyme Structure and Mechanism* (2nd ed.) Freeman, New York

Fischer, H.O.L. and Dangschat, G. (1932) *Chem. Ber.* **65**, 1009-1031

Fischer, H.O.L. and Dangschat, G. (1934) *Helv. Chim. Acta*, **17**, 1196-1200

Fisher, H.F., Conn, E.E., Vennesland, B. and Westheimer, F.H. (1953) *J. Biol. Chem.* **202**, 687-697

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

Freedman, R.B. and Radda, G.K. (1968) *Biochem. J.* **108**, 383-391

Freudenberg, K., Meisenheimer, H., Lane J.T. and Plankenhorn, E. (1940) *Annalen*, **543**, 162-171

Fromm, H.J. and Nelson, D.R. (1962) *J. Biol. Chem.* **237**, 215-220

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

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

Gaertner, F.H. (1987) *Methods in Enzymology* **142**, 362-366

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

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

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

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

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

Gibson, F. and Pittard, A. J. (1968) *Bact eriological Review* **32**, 465-492

Goldfarb, A.R. (1966) *Biochemistry* **5**, 2570-2574

Gorlach, J, Schmid, J. and Amrhein, N. (1993) *Plant Mol. Biol.* **23**, 707-716

Gourley, D.G., Coggins, J.R., Isaacs, N.W., Moore, J.D., Charles, I.G. and Hawkins, A.R. (1994) *J. Mol. Biol.* **241**, 488-491

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

Graziana, A. and Boudet, A.M. (1980) *Plant Cell Physiol.* **21**, 793-802

Hammond, S.J., Williamson, M.P., Williams, D.H., Boeck, L.D. and Marconi, G.G. (1982) *J. Chem. Soc., Chem. Commun.* 344-346

Hasan, N. and Nester, E.W. (1978) *J. Biol. Chem.* **253**, 4999-5004

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

Haslam, E. (1993) In *Shikimic Acid- Metabolism and Metabolites*, John Wiley and Sons, Chichester

Hawkins, A.R., Reinert, W.R. and Giles, N.H. (1982) *Biochem. J.* **203**, 769-773

Hawkins, A.R. (1987) *Current Genetics* **11**, 491-498

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

Hawkins, A.R., Moore, J.D. and Adeokun, A.M. (1993) *Biochem. J.* **296**, 451-457

Henstrand, J.M., Mccue, K.F., Brink, K., Handa, A.K., Herrmann, K.M. and Conn, E.E. (1992) *Plant Physiol.* **98**, 761-763

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

Hill, E.J., Tsernoglou, D., Webb, L. and Banaszak, L.J. (1972) *J. Mol. Biol.* **72**, 577-591

Hol, W.G.J., Van Duijnen, P.T. and Berendsen, H.J.C. (1978) *Nature* **273**, 443-446

Holbrook, J.J. and Gutfreund, H. (1973) *FEBS Letters* **31**, 157-169

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

Holbrook, J.J., Lodola, A. and Illsley, N.P. (1974) *Biochem. J.* **139**, 797-800

Holbrook, J.J., Liljas, A., Steindel, S.J. and Rossmann, M.G. (1975) In *The Enzymes* (P.D. Boyer. ed.) 3rd ed., Vol. XI , Academic Press, New York

Hollenberg, P.F., Flashner, M. and Coon, M.J. (1971) *J. Biol. Chem.* **246**, 946-953

Horecker, B.L. and Kornberg, A. (1948) *J. Biol. Chem.* **175**, 385-389

Horecker, B.L., Tsolas, O. and Lai, C.Y. (1970) *The Enzymes* (P.D. Boyer ed.) 3rd ed., Vol.VII, pg 213-253, Academic Press, New York

Huang, L., Montoya, A.L. and Nester, E.W. (1975) *J. Biol. Chem.* **250**, 7675-7681

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

Hungerer, C., Troup, B. and Jahn, D.- EMBL/GenBank Accession No. X85015

Hunter, M.J. and Ludwig, M.L. (1962) *J. Am. Chem. Soc.* **84**, 3491-3504

Hunter, M.J. and Ludwig, M.L. (1972) *Methods in Enzymology* Vol XXV, pp 585-596

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

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

Huynh, Q.K., Bauer, S.C., Bild, G.S., Kishore, G.M.and Borgmeyer, J.R. (1988b) *J. Biol. Chem.* **263**, 11636-11639

Huynh, Q.K. (1993) *Biochem. J.* **290**, 525-530

Jaquinod, M., Potier, N., Klarskov, K., Reymann, J.M., Sorokine, O., Kieffer, S., Barth, P., Andriantomanga, V., Biellmann, J.F. and Van Dorsselaer, A. (1993) *Eur. J. Biochem.* **218**, 893-903

Jornvall, H., Bahr-Lindstrom, H.V., Jany, K., Ulmer, W. and Froschle, M. (1984) *FEBS Lett.* **165**, 190-196

Karas, M. and Hillenkamp, F. (1988) *Anal. Chem.* **60**, 2299-2301

Karrer, P. and Link, K.P. (1927) *Helv. Chim. Acta* **10**, 794-799

Kim C.-G., Kirschning, A., Bergon, P., Ahn, Y., Wang, J.J., Shibuya, M. and Floss, H.G. (1992) *J. Am. Chem. Soc.* **114**, 4941-4943

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

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

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

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

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) *Biochem. J.* **282**, 687-695

Knowles, W.S., Anderson, K.S., Andrew, S.S., Phillion, D.P., Ream, J.E., Johnson, K.A. and Sikorski, J.A. (1993) *Bioorganic and Medicinal Chem. Lett.* **3**, 2863-2868

Koshiba, T. (1978) *Biochim. Biophys. Acta* **522**, 10-18

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

Kotaki, A., Harada, M. and Yagi, K. (1964) *J.Biochem. Tokoyo* **55**, 553-561

Kraut, J. (1977) *Ann. Rev. in Biochem.* **46**, 331-358

Krell, T., Pitt, A. R. and Coggins, J.R. (1995) *FEBS Letters* **360**, 93-96

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

Lamb, H.K., Moore, J.D., Lakey, J.H., Levett, L., Wheeler, K.A., Lago, H., Coggins, J.R. and Hawkins, A.R. (1995) - manuscript submitted for publication

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

Lange, L.G., Riordan, J.F. and Vallee, B.L. (1974) *Biochemistry* **13**, 4361-4370

Lee, W.T. and Levy, H.R. (1992) *Protein Science* **1**, 329-334

Leo, G.C., Sikorski, J.A. and Sammons, P.D. (1990) *J. Am. Chem. Soc.* **112**, 1653-1654

Levin, J.G. and Sprinson, D.B. (1960) *Biochem. Biophys. Res. Commun.* **3**, 157-163

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

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

Linhart, Y.B., Davis, M.L. and Mitton, J.B. (1981) *Biochem. Genet.* **19**, 641-646

Lipscomb, W.N. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 3797-3801

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

Loewus, F. A., Westheimer, F.H. and Vennesland, B. (1953) *J. Am. Chem. Soc.* **75**, 5018-5023

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

Lourenco, E.J. and Neves, V.A. (1984) *Phytochemistry* **23**, 497-499

Lourenco, E.J., Silva, G.M.L. and Neves, V.A (1991) *J. Agric. Food Chem.* **39**, 458-462

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

Lundblad, R.L. and Noyes, C.M. (1984) In *Chemical Reagents for Protein Modification*,

Vol.I, CRC Press, Boca Raton, Florida

Mann, M. and Wilm, M. (1995) *TIBS* **20**, 219-224

Mann, M., Meng, C.K. and Feng, J.B. (1989) *Anal. Chem.* **61**, 1702-1708

McKee, J.S. and Nimmo, H.G. (1989) *Biochem. J.* **261**, 301-304

Means, G.E. and Feeney, R.E. (1971) In *Chemical Modification of Proteins*, Holden-Day Inc., San Francisco, California

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

Meyer, H.E., Eisermann, B., Heber, M., Hoffmann-Posorske.E., Korte, H., Weigt, C., Wegner, A., Hutton, T., Donella-Deana, A. and Perich, J.W. (1993) *FASEB J.* **7**, 776-782

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

Milhausen, M. and Levy, H.R. (1975) *Eur.J. Biochem.* **50**, 453-461

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

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

Miller, M.J., Anderson, K.S., Braccolino, D.S., Cleary, D.G.,Gruys, K.J., Han, C.Y., Lin, K.C, Pansegrau, P.D., Ream, J.E., Sammons, R.D. and Sikorski, J.A. (1993) *Bioorg.and Med.Chem. Lett.* **3**, 1435-1440

Miller, M.J., Braccolino, D.S., Cleary, D.G., Ream, J.E., Walker, M.C. and Sikorski, J.A. (1994) *Bioorg. and Med.Chem. Lett.* **4**, 2605-2608

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

Moore, J.D. and Hawkins, A.R. (1993) *Mol. Gen. Genet.* **240**, 92-102

Moore, B.S. and Floss, H.G. (1994) *J. Nat. Prod.* **57**, 382-386

Moore, J.D., Coggins, J.R., Virden, R.and Hawkins, A.R. (1994) *Biochem. J.* **301**, 297-304

Moras, D., Olsen, K.W., Sabesan, M.N., Buehner, M., Ford, G.C. and Rossmann, M.G. (1975) *J.Biol. Chem.* **250**, 9137-9162

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

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

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

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

Muhlrad, A., Hegyi, G. and Toth, G. (1967) *Biochim. Biophys. Acta* **2**, 19-26

Nagano, H. and Zalkin, (1970) *Arch. Biochem. Biophys.* **138**, 58-65

Nakatsukasa, W.M. and Nester, E.W. (1972) *J. Biol. Chem.* **247**, 5972-5979

Neale, A.D., Scopes, R.K., Kelly, J.M. and Wettenhall, R.E.H. (1986) *Eur. J. Biochem.* **154**, 119-124

Nimmo, H.G. (1986) *Biochem. J.* **234**, 317-323

Nishiyama, M., Birktoft, J.J. and Beppu, T. (1993) *J. Biol. Chem.* **268**, 4656-4660

Ogino, T., Garner, C., Markley, J.L. and Herrmann, K.M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5828-5832

Ohlsson. I., Nordstrom, B. and Branden, C.I. (1974) *J. Mol. Biol.* **89**, 339-354

Okuyama, T. and Satake, K. (1960) *J. Biochem. Tokyo* **47**, 454-466

Olive, C., Geroch, M.E. and Levy, H.R. (1971) *J.Biol. Chem.* **246**, 2047-2057

Ovadi, J., Libor, S. and Elodi, P. (1967) *Biochim. Biophys. Acta* **2**, 455-458

Padgette, S.R., Smith, C.E., Huynh, Q.K. and Kishore, G.M. (1988a) *Arch. Biochem. Biophys.* **266**, 254-262

Padgette, S.R., Huynh, Q.K., Aykent, S., Sammons, R.D., Sikorski, J.A. and Kishore, G.M. (1988b) *J. Biol. Chem.* **263**, 1798-1802

Patel, V.B. and Giles, N.H. (1979) *Biochim. Biophys. Acta* **567**, 24-34

Patthy, L. and Thesz, J. (1980) *Eur. J. Biochem.* **105**, 387-393

Peters, R.G., Jones, W.C. and Cromartie, T.H. (1981) *Biochemistry* **20**, 2564-2571

Pinto, J.E.B.P., Suzich, J.A. and Herrmann, K.M. (1986) *Plant Physiol.* **82**, 1040-1044

Pittard, A.J. and Wallace, B.J. (1966) *J. Bacteriol.* **92**, 1494-1508

Pittard, J. (1987) In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Vol.I, Ed., Neidhardt, F.C., American Society for Microbiology, Washington D.C.

Pixley, F.J., Wakefield, A.E., Banerji, S. and Hopkin, J.M. (1991) *Mol. Microbiol.* **5**, 1347-1351

Polley, L.D. (1978) *Biochim. Biophys. Acta* **526**, 259-266

Pompliano, D.L., Reimer, L.M., Myrvoid, S. and Frost, J.W. (1989) *J. Am. Chem. Soc.* **111**, 1866-1871

Price, N.C. and Stevens, L. (1989) In *Fundamentals of Enzymology* (II Ed.), Oxford University Press, New York

Ramjee, M.N., Balasubramanian, S., Abell, C., Coggins, J.R., Davies, G.M., Hawkes, T.R., Lowe, D.J. and Thorneley, R.N.F. (1992) *J. Am. Chem. Soc.* **114**, 3151-3153

Ramjee, M.N., Coggins, J.R., Hawkes, T.R., Lowe, D.J. and Thorneley, R.N.F. (1991) *J. Am. Chem. Soc.* **113**, 8566-8567

Ray, W.J., Jr. and Koshland, D.E. (1962) *J. Biol. Chem.* **237**, 2493- 2505

Ray, J.M. and Bauerle, R. (1991) *J. Bacteriol.* **173**, 1894-1901

Richards, F.M. and Wyckoff, H.W. (1971) *The Enzymes* (P.D. Boyer ed.) 3rd ed.,vol. **IV**, pg. 647-784, Academic Press, New York

Richardson, J.S. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2619-2623

Richardson, J.S. (1977) *Nature* **268**, 495-500

Riordan, J.F., McElvany, K.D. and Borders, C.L. Jr. (1977) *Science* **195**, 884-886

Riordan, J.F. (1979) *Mol. Cell. Biochem.* **26**, 71-92

Robinson, T. (1981) In *The Biochemistry of Alkaloids* (II Ed.) Springer-Verlag, New York

Rossmann, M.G., Moras, D. and Olsen, K.W. (1974) *Nature* **250**, 194-199

Rossmann, M.J., Liljas, A., Branden, C.I. and Banaszak, L.J. (1975)*The Enzymes*, (P.D. Boyer ed.) (3rd ed.), Vol XI, Academic Press, New York

Rouhbakhsh, D.- EMBL/Genbank Accession No. U09230

Rubin, J.L., Gaines, C.G. and Jensen, R.A. (1982) *Plant Physiol.* **70**, 833-839

Rubin, J.L., Gaines, C.G. and Jensen, R.A. (1984) *Plant Physiol.* **75**, 839-845

Rubin, J.L. and Jensen, R.A. (1985) *Plant Physiol.* **79**, 711-718

Satake, K., Okuyama, T., Ohashi, M. and Shinoda, T. (1960) *J. Biochem. Tokyo* **47**, 654-660

Schaller, A., Windhofer, V. and Amrhein, N. (1990) *Arch. Biochem. Biophys.* **282**, 437-442

Schaller, A., Afferden, M.V., Windhofer, V., Bulow, S., Abel, G., Schmid, J. and Amrhein, N. (1991) *Plant Physiol.* **97**, 1271-1279

Schmidt, D.E.Jr. and Westheimer, F.H. (1971) *Biochemistry* **10**, 1249-1253

Schmid, J., Schaller, A., Leibinger, U., Boll, W. and Amrhein, N. (1992) *The Plant J.* **2**, 375-383

Schneider, F. (1978) *Angew. Chem. Int. Ed. Engl.* **17**, 583-592

Schwert, G.W., Miller, B. R. and Peanasky, R.J. (1967) *J. Biol. Chem.* **242**, 3245-3252

Scrutton, N.S., Berry, A. and Perham, R.N. (1990) *Nature* **343**, 38-43

Servos, S., Chatfield, S., Hone, D., Levine, M., Dimitriadis, G., Pickard, D., Dougan, G., Fairweather, N. and Charles, I.A. (1991) *J. Gen. Microbiol.* **137**, 147-152

Sharma, R., Bhatnagar, R.K. and Sarin, N.B. (1993) *Physiologia Plantarum* **88**, 281-286

Shneier, A., Harris, J., Kleanthous, C., Coggins, J.R., Hawkins, A.R. and Abell, C. (1993) *Bioorg. and Med. Chem. Lett.* **3**, 1399-1402

Siuzdak, G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11290-11297

Smith, D.D.S. and Coggins, J.R. (1983) *Biochem. J.* **213**, 405-415

Soon, C.Y., Shepherd, M.G. and Sullivan, P.A. (1978) *Biochem. J.* **173**, 255-262

Srinivasan, P.R., Rothschild, J. and Sprinson, D.B. (1963) *J. Biol. Chem.* **238**, 3176-3182

Stalker, D.M., Hiatt, W.R. and Comai, L. (1985) *J. Biol. Chem.* **260**, 4724-4728

Staub, M. and Denes, G. (1969) *Biochim. Biophys. Acta.* **178**, 599-608

Steinrucken, H.C. and Amrhein, N. (1980) *Biochem. Biophys. Res. Comm.* **94**, 1207-1212

Steinrucken, H.C. and Amrhein, N, (1984) *Eur. J. Biochem.* **143**, 351-357

Strauss, A. (1979) *Mol. Gen. Genet.* **172**, 233-241

Sugimoto, S. and Shiio, I. (1980) *J. Biochem.* **87**, 881-890

Suzich, J.A., Dean, F.J.D. and Herrmann, K.M. (1985) *Plant Physiol.* **79**, 765-770

Takahashi, K. (1968) *J. Biol. Chem.* **243**, 6171-6179

Takenaka, Y. and Schwert, G.W. (1956) *J. Biol. Chem.* **223**, 157-170

Takeuchi, M., Asano, M., Kameda, Y. and Matsui, K. (1986) *J. Biochem.* **99**, 1571-1577

Tresguerres, M.E.F., Torrontegui, G.de, Ingledew, W.M. and Canovas J.L. (1970) *Eur. J. Biochem.* **14**, 445-450

Turner, M.J., Smith, B.W. and Haslam, E. (1975) *J. Chem. Soc.- Perkin* **1**, 52-55

Turner, J.M. and Messenger, A. (1986) *Adv. Microbiol. Physiol.* **27**, 211-275

Vangrysperre, W., Ampe, C., Kersters-Hilderson, H. and Tempst, P. (1989) *Biochem. J.* **263**, 195-199

Walsh, C. (1979) In *Enzymatic Reaction Mechanisms*, Freeman, San Francisco, California

Wang, Y., Jones, J.D., Weller, S.C. and Goldsbrough, P.B. (1991) *Plant Mol. Biol.* **17**, 1127-1138

Warshel, A., Naray-Szabo, G., Sussman, F. and Hwang, J.K. (1989) *Biochemistry* **28**, 3629-3638

Webb, L.E., Hill, E.J. and Banaszak, L.J. (1973) *Biochemistry* **12**, 5101-5109

Weiss, U. and Edwards, J.M. (1980) In *The Biosynthesis of Aromatic Compounds*, Wiley-Interscience, New York

Welch, G.R., Cole, K.W. and Gaertner, F.H. (1974) *Arch. Biochem. Biophys.* **165**, 505-518

Wells, M.A. (1973) *Biochemistry* **12**, 1086-1093

Westheimer, F.H., Fisher, H.F., Conn, E.E.and Vennesland, B. (1951) *J. Am. Chem. Soc.* **73**, 2403

White, P.J., Millar, G. and Coggins, J.R. (1988) *Biochem. J.* **251**, 313-322

White, P.J., Young, J., Hunter, I.S., Nimmo, H.G. and Coggins, J.R. (1990) *Biochem. J.* **265**, 735-738

Widersten, M., Kolm, R.H., Bjornestedt, R. and Mannervik, B. (1992) *Biochem. J.* **285**, 377-381

Widlanski, T., Bender, S.L.and Knowles, J.R. (1989a) *Biochemistry* **28**, 7572-7582

Widlanski, T., Bender, S.L.and Knowles, J.R. (1989b) *J. Am. Chem. Soc.* **111**, 2299-2300

Wierenga, R.K., De Maeyer, M. C.H. and Hol. W.G.J. (1985) *Biochemistry* **24**, 1346-1357

Wierenga, R.K., Terpstra, P. and Hol, W.G.J. (1986) *J. Mol. Biol.* **187**, 101-107

Williamson, S.C., Yu, H. and Davis, T.M. (1995) *J. Heredity* **86**, 74-76

Wilson, D.K., Bohren, K.M., Gabbay, K.H. and Quiocho, F.A. (1992) *Science* **257**, 81-84

Winer, A.D. and Schwert, G.W. (1958) *J. Biol. Chem.* **231**, 1065-1083

Wratten, C.C. and Cleland, W.W. (1963) *Biochemistry* **2**, 935-941

Wratten, C.C. and Cleland, W.W. (1965) *Biochemistry* **4**, 2442-2451

Xia, C., Meyer, D.J., Chen, H., Reinemer, P., Huber, R. and Ketterer, B. (1993) *Biochem. J.* **293**, 357-362

Yamamoto, E. (1980) *Phytochemistry* **19**, 779-781

Yaniv, H. and Gilvarg, C. (1955) *J. Biol. Chem.* **213**, 787-795

Yu, H. and Kiang, Y.T. (1993) *J. Heredity* **84**, 489-492

Zewe, V. and Fromm, H.J. (1965) *Biochemistry* **4**, 782-792

Appendix

```
 eco ...MET..YA VFGNPIAHSK SPFIHQQFAQ QLNIEHPYGR VLAPINDFIN45
 pae ...MDR..YC VFGNPIGHSK SPLIHRLFAE QTGEALVYDA QLAPLDDFPG
 Bac MCKLEKFNYA LFGNPIDHSQ SPKIHNFFAT QTGILHIYKA INIPLDQFSS
 psa QLGPETKVYG IIGKPVSHSK SPILFNEAFK TVGFNGVFVF LLVDDLANFL
 nta QLGPDTRIFG IIGKPVSHSK SPLLYNEAFR SVGFNGVYMP LLVDDVANFF
 sce MGGIEPKELF VVGKPIGHSR SPILHNTGYE ILGLPHKFDK FETESAQLVK
 ani .....PKKFA IFGSPISQSR SPALHNTLFA QVGLPHNYTR LETTNAQDVQ
 pca MGLLPEKKYF LFGKPIKHSQ SPNIHNLGFE ILGLPYKYQL FETDSISELK
Cons .....e...Y ..G.Pi.HS. SP..hN.... ..g....y.. ..........

 eco TLNAFFSAGG KGANVTVPFK EEAFARADEL TERAALAGAV NTLMRLEDGR195
 pae FARRFF.EQG KGANVTVPFK EEAYRLVDEL SERATRAGAV NTLIRLADGR
 Bac VVSDFFKKNI KGANVTAPFK KEAYFFSDKL TERAKIAQSV NTLKKISDKC
 psa RTY..SSTDF VGFSVTIPHK ESALKCCDEV DPVAKSIGAV N.........
 nta RTY..SSLDF AGSAVTIPHK EAIVDCCDEL NPTAKVIGAV NCVVSRLDGK
 sce EKLLDGNKNF GGAAVTIPLK LDIMQYMDEL TDAAKVIGAV NTVIPLGNKK
 ani E..FIRSPDF GGGSVTINLK LDIMPLLDEV AAFAFIIGAV NTIIPVSTGK
 pca EILHLE..EF GGASVTIPLK TNISILLDEI SDHAALIGSV NTITR..TYN
Cons .......... .G..Vtip.K .......De. ...A...gaV Nt........

 eco .....LLGDN TDGVGLLSDL .ERLSFIRPG LR........ ILLIGAGGAS131
 pae .....LRGDN TDGAGLLRDL TANAGVDVRG KR........ VLLLGAGGAV
 Bac .....ILGDN TDGIGLLSDL V.RLNFIKKN FS........ ILILGAGGAV
 psa .......... .......... .......... .......... ..........
 nta .....LFGCN TDYVGAISAI EEALQGSQPS MSGSPLAGKL FVVIGAGGAG
 sce .....FKGDN TDWLGIRNAL INNGV...PE YVGHTAG... .LVIGAGGTS
 ani NTPSRLVGRN TDWQGMILSL RKAGVYGPKR KDQEQSA... .LVVGGGGTA
 pca NGQYILKGEN TDWQGIIKAI KNFNKFE..K SFENFSG... .FIIGAGGAS
Cons .....l.G.N TD..G..... .......... .......... .l..GaGG..

 eco RGVLLPLLSL DCA.VTITNR TVSRAEELAK LFAHTGSIQA LSMDE.....175
 pae RGVLEPFLGE CPAELLIANR TARKAVDLAE RFADLGAVHG CGFAE.....
 Bac KGVLLPLLSL GCS.VYILNR TILNAKILVK QFNKYGKIFV FDRQN.....
 psa .......... .......... .......... .......... ..........
 nta KALAYGAKEK GARVVI.... .......... ......ANRT YERARELADV
 sce RAALYALHSL GCKKIFIINR TTSKLKPLIE SLPSEFNIIG IESTKSIEEI
 ani RAAIYALHNM GYSPIYIVGR TPSKLENMVS TFPSSYNIRI VESPSSFESV
 pca RAAIYALLSL GISPIYLINR SKDKLNKLYH FFNTN.HIIP ITEYHELNNI
Cons ......l... .........R .......l.. .f.....i.. ..........

 eco LEGHEFDLII NATSSGISGD IPAIPSSLIH PG.I...... ........YC210
 pae VEGP.FDLIV NGTSASLAGD VPPLAQSVIE PGRT...... ........VC
 Bac FKQQNFDLVI NAMSRNTEKK NFTL...ILI TSKR...... ........FF
 psa .......... .......... .......... .......... ..........
 nta VGGQALSLDE LSNFHPENDM ILANTTSIGM QPKVDDTPIF KEALRYYSLV
 sce KEHVGVAVSC VPADKPLDDE LLSKLERFLV KGAHAA.... .....FVPTL
 ani PH...VAIGT IPADQPIDPT MRETLCHMFE RAQEADAEAV KAIEHAPRIL
 pca NFDIRIGIST IPTDNPIDPS VLEIAKIFFN LKRKSSEG.. ........IF
Cons .......... .......... .......... .......... ..........
```

```
 eco YDMFYQKGKT PFLAWCEQRG SKRNADGLGM LVAQAAHAFL LWHGVLPDVE260
 pae YDMMYAKEPT AFNRWAAERG AARTLDGLGM LVEQAAEAFF LWRGVRPASA
 Bac YDMNYSTRNT PFINWCSKAG GSFISNGIGM LVFQAAYSFL EWHNVLPEIN
 psa .......... .......... .......... .......... ..........
 nta FDAVYTPKIT RLLREAHES. GVKIVTGVEM FIGQAYEQYE RFTGLASSKG
 sce LEAAYKPSVT PVMTISQDKY QWHVVPGSQM LVHQGVAQFE KWTGFKGPFK
 ani LEMAYKPQVT ALMRLASDS. GWKTIPGLEV LVGQGWYQFK YWTGISPLYE
 pca LDMAYGSNTT D.LTIIAKAC NWKIIHGLEI LLEQGSEQFL LWTETYIPYN
Cons .D..Y....T .......... ......G... lv.Q....f. .w........

 eco PVIKQLQEEL SA272 ..... .......... .......... ..........
 pae PVLETLRRQL ATV*..... .......... .......... ..........
 Bac YIINILNIK* .......... .......... .......... ..........
 psa .......... .......... .......... .......... ..........
 nta TFQENYGWIL RARSLSLFNA ALLVTFPPKS LHSCVIAMVL DSSALPFVLR
 sce AIFDAVTKE. .......... .......... .......... ..........
 ani SARACSSPLI .......... .......... .......... ..........
 pca QVKYAILGPN K......... .......... .......... ..........
Cons .......... .......... .......... .......... ..........

 eco ..
 pae ..
 Bac ..
 psa ..
 nta RN
 sce ..
 ani ..
 pca ..
Cons ..
```

Alignment of SKDH sequences taken from the GCG PileUp program (Devereux *et al.*, 1987) mounted on the Glasgow University UNIX system. References; *E. coli* (Anton and Coggins, 1988), *Pseudomonas aeruginosa* (accession no. X85015), *Buchnera aphidicola* (accession no. U09230), *Pisum sativum* (Deka *et al.*, 1994), *Nicotiana tabacum* (Bonner and Jensen, 1994), *Saccharomyces cerevisiae* (Duncan *et al.*, 1987), *Aspergillus nidulans* (*Charles et al.*, 1986; J. Moore and A.R. Hawkins, personal communication) and *Pneumocystis carinii* (Banerji *et al.*, 1993).