Krell, Tino (1996) Studies on the chemistry of enzyme active sites. PhD thesis

http://theses.gla.ac.uk/6141/

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Glasgow Theses Service
http://theses.gla.ac.uk/
theses@gla.ac.uk

Studies on the chemistry of enzyme active sites

A thesis submitted to the University of Glasgow

for the degree of
DOCTOR OF PHILOSOPHY

IN
BIOCHEMISTRY

UNIVERSITY
of
GLASGOW

BY
TINO KRELL

Division of Biochemistry and Molecular Biology
Institute of Biomedical and Life Sciences
University of Glasgow
Glasgow, G12 8QQ
SCOTLAND

OCTOBER, 1996

Dedicated to my parents Monika and Werner

and to my wife Belén

Abstract

A major part of this thesis involved using electrospray mass spectrometry to monitor the chemical modification of amino acid side chains in enzymes. The technique was used specially to locate active site residues in shikimate pathway enzymes and to monitor the dephosphorylation of a phospho enzyme intermediate.

Firstly, site-specific chemical modification in combination with mass spectrometry was used to identify Arg-23 in *Streptomyces coelicolor* type II dehydroquinase (DHQ) as a residue essential for enzyme function. This residue was replaced by lysine, glutamine and alanine residues using site-directed mutagenesis. All the mutants were shown to have much lower turn over numbers as well as lower Km values in comparison to the native enzyme. This makes a role for Arg-23 in substrate binding unlikely. A catalytic role for this residue in stabilising a negatively charged enolate transition state is proposed since the mutant R23A was found to be 10 times less active than R23K and R23Q. Furthermore, Tyr-28 of *S. coelicolor* DHQ and Arg-213 of *Escherichia coli* type I DHQ have been shown to be in or near the active site.

Secondly, mass spectrometry was used to monitor the dephosphorylation of phosphorylated forms of phosphoglycerate mutases (PGAM). The phosphorylated PGAM from *Saccharomyces cerevisiae* was shown to be at least 35 times more stable than the enzyme from *Schizosaccharomyces pombe* which does not contain a C-terminal segment of 14 amino acids which is probably responsible for the differences in stability. The phosphorylated mutant H163Q mutant of *S. cerevisiae* PGAM appeared to be at least 400 times more stable than the native enzyme.

Thirdly, chemical modification and mass spectrometry were used to identify active site residues in *E. coli* shikimate dehydrogenase (SDH). Two lysine residues were shown to be in or near the active site; the inactivation of SDH by treatment with the lysine specific reagent trinitrobenzenesulfonic acid was shown to be due to the modification of only one residue, Lys-65, which is proposed to be involved in binding the carboxylate group of shikimic acid. Arg-154 has also been identified as an active site residue and assigned a role in binding of the 2' phosphate group of NADP.

Shikimate kinase from *Erwinia chrysanthemi* was overexpressed in *E. coli* to an amount of up to 30% of the total cellular protein. The enzyme was purified, characterised and crystallised in complex with shikimic acid and ADP. It was predicted that arginine residues would be present in the binding sites of ADP and shikimic acid. This was confirmed by solving the three dimensional structure of the enzyme.

DECLARATION

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

† This thesis consists of 6 manuscripts either published or to be published in the near future. The following collaborators have contributed to this work.

Chapter 3:

Dr. Alan Cooper (University of Glasgow) - carried out isothermal titration calorimetry experiments

Dr. Sharon Kelly (University of Stirling) - recorded the c.d. spectra

Chapter 4 and 5:

The mass -spectrometric analysis is my own work, the preparation of phospho enzymes was carried out by Dr. Jaqueline Nairn and the chemical modification experiments were carried out by Dr. Sureka Chackrewarthy.

Chapter 6:

Dr. Malcolm Horsburgh (University of Sheffield) - constructed the overexpressing strain for *Erwinia chrysanthemi* shikimate kinase.

Joe Coyle (University of Leeds) - obtained first crystals of shikimate kinase

Date: 26. 10. 1996 Signature: [redacted]

ACKNOWLEDGEMENTS

The work would not have been possible without the help of many people.
Many thanks to Prof. J.R. Coggins for his enthusiastic support throughout the years. I would like to thank all members of the molecular enzymology group, especially Dr. Graeme Walker for teaching me about protein purification, Dr. Malcolm Horsburgh for his help with the DNA-work, Andrew Elwell for showing me incredible short-cuts on the computer and John Greene for technical assistance and many good jokes.

Special thanks to all members of Glasgow Protein Crystallography group, especially to Dr. Adrian Lapthorn, for teaching me about crystals and computer. Sincere thanks to Dr. Andrew Pitt from Strathclyde University for his help and support with mass spectrometry.

Many thanks to our collaborators in Stirling, especially Prof. N.C. Price, Dr. Jaqueline Nairn and Dr. Sharon Kelly for many helpful discussions, recording spectra and their efficiency. I would also like to thank Dr. Chris Abell and Dr. Conchi González-Bello from the University of Cambridge for interesting discussions, substrate and their hospitality.

Last but not least many thanks to all those people who made my stay in Glasgow really enjoyable: Marta Terrazo, Mairi Munro, Susan Cummings, Ina and Paul Watson, Hazel and Johann De Bono, Colin Johnson, Saiqa Khan, Roy Parkin, Mohammed Abderrahim, Silvia Arribas, Susanne Ridder, Amal Chatterjee, Monica Landoni, Fabio Crestani and Daniel Loh Kuan Yew.

Contents

List of Figures

List of Tables

Abbreviations

A	absorbance
amp	ampicillin
ATP	adenosine triphosphate
BPG	2,3-bisphosphoglycerate
BSA	bovine serum albumin
c.d.	circular dichroism
Da	Dalton
DAHP	3-deoxy-D-arabino-heptulosonate-7-phosphate
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxynucleoside triphosphate
DEAE	diethylaminoethyl
DEPC	diethylpyrocarbonate
DHQ	3-dehydroquinase
DMSO	dimethylsulphoxide
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetate
EM	energy minimisation
EPSP	5-enolpyruvylshikimate 3-phosphate
ES-MS	electrospray mass spectrometry
FPLC	fast protein liquid chromatography
GdnHCl	guanidinium chloride
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
HPLC	high pressure liquid chromatography
IPTG	isopropyl-β-D-thiogalactoside
ITC	isothermal titration calorimetry
kb	kilo base pairs
l.m.p.	low melting point
LB	Luria Bertani medium
LC/ES-MS	liquid chromatography/electrospray mass spectrometry

MaxEnt	Maximum Entropy
MD	molecular dynamics
M_r	relative molecular mass
NAD^+	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide-reduced
$NADP^+$	nicotinamide adenine dinucleotide phosphate
NADPH	nicotinamide adenine dinucleotide phosphate-reduced
PAGE	polyacrylamide gel-electrophoresis
PCR	polymerase chain reaction
PEP	phosphoenol pyruvate
PGAM	phosphoglycerate mutase
PGO	phenylglyoxal
RNAse	ribonuclease
SDH	shikimate dehydrogenase
SDS	sodium dodecyl sulphate
SK	shikimate kinase
TE	Tris-acetate/EDTA buffer
TEMED	N, N, N', N'-tetramethyl-ethylenediamine
TFA	trifluoroacetic acid
TNBS	trinitrobenzene sulfonic acid
TNM	tetranitromethane
TPCK	N-p-toluensulphonyl-L- phenylalanine chloromethylketone
Tris	Tris (hydroxymethyl) aminomethane
U	units of enzyme activity
u.v.	ultra violet
2PG	2-phosphoglycollate
3PGA	3-phosphoglycerate

1. Introduction

1.1 The shikimate pathway

In plants, as in micro-organisms, the biosynthesis of all aromatic compounds involved in primary metabolism proceeds by way of the shikimate pathway (Haslam, 1993). Seven enzymes (Fig. 1-1) convert erythrose-4-phosphate (a product of the Calvin cycle and/or pentose phosphate pathway) and phosphoenolpyruvate (a glycolytic product) to chorismic acid. This is a common precursor for the synthesis of the aromatic amino acids and other aromatic compounds such as folic acid and vitamin K (Haslam, 1993). The shikimate pathway provides the precursors for many plant and microbial secondary metabolites like lignin and flavonoids and it has been estimated that products of the shikimate pathway account for up to 35% of the dry weight of higher plants (Boudet *et al.*, 1985). This pathway is essential to plants and microorganisms, but it is absent in animals. Inhibitors of the pathway enzymes are therefore potentially antimicrobial agents or herbicides with no mammalian toxicity (Davies *et al.*, 1994; Jude *et al.*, 1996). The commercially successful and widely used herbicide *glyphosate* is an inhibitor of EPSP synthase, the sixth enzyme of the shikimate pathway (Steinruecken and Amrhein, 1980)

The complete pathway in micro-organisms was elucidated in the fifties by Davies and Sprinson who used radiolabelled carbon compounds as well as auxotrophic mutants to isolate intermediates (Davies, 1955; Sprinson, 1961). The situation in plants is more complex: all shikimate pathway enzymes have been detected in higher plants (Jensen, 1986) and there is evidence that the complete pathway occurs in chloroplasts (Schulze-Siebert, 1984; Mousdale and Coggins, 1985). However there is further evidence for a second, either truncated or incomplete shikimate pathway in the cytosol (Jensen, 1986).

The initial study of the pathway enzymes was difficult. This was mainly due to the enzymes being present at a very low level and, therefore, their purification was difficult. Those problems were overcome by the cloning and overexpression of all pathway enzymes in the eighties, especially by Coggins and co-workers (Coggins, 1989).

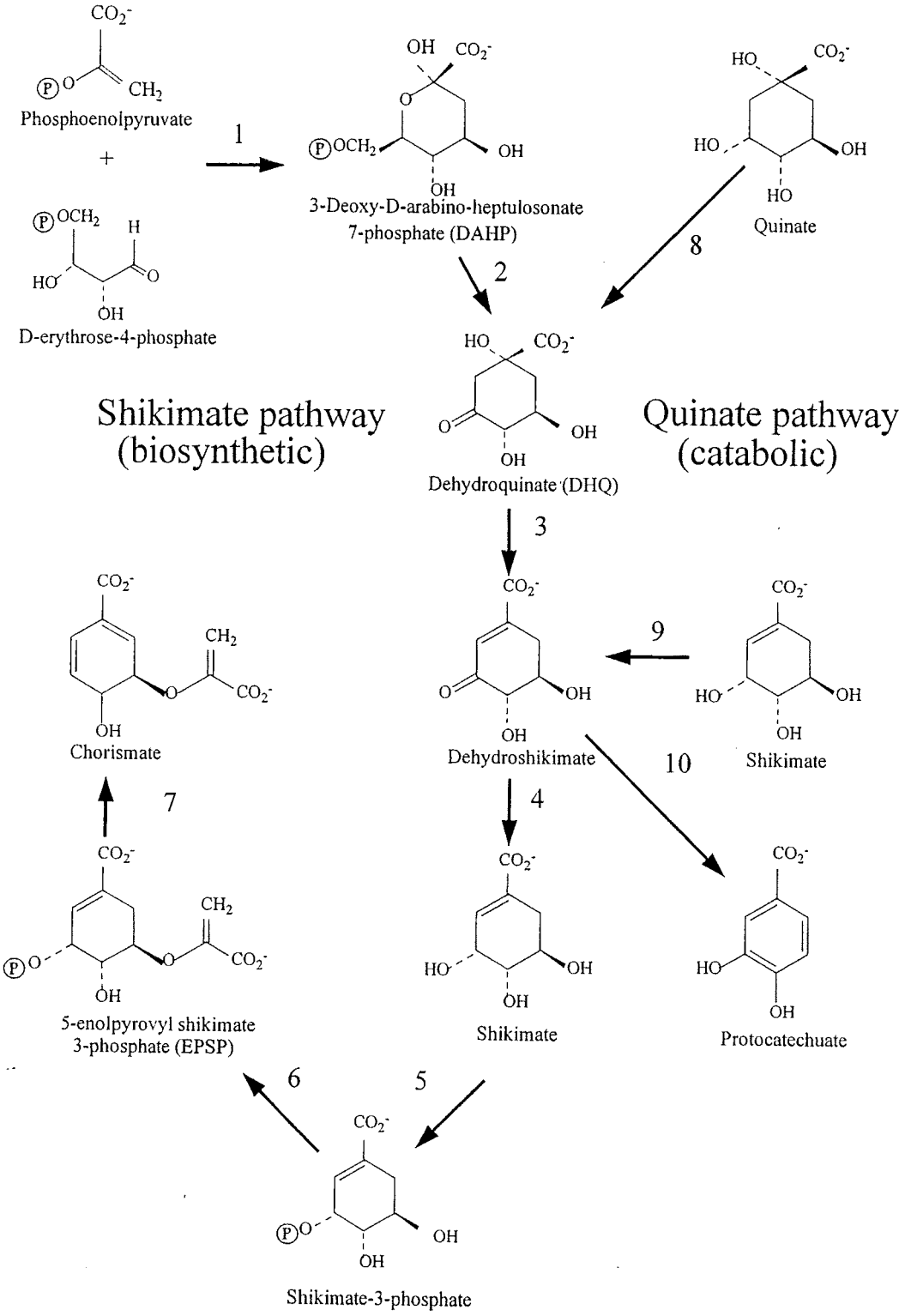


Figure 1-1 The shikimate and quinate pathway.

1	DAHP synthase	**6**	EPSP synthase
2	DHQ synthase	**7**	Chorismate synthase
3	Dehydroquinase	**8**	Quinate dehydrogenase
4	Shikimate dehydrogenase	**9**	Shikimate dehydrogenase
5	Shikimate kinase	**10**	Dehydroshikimate dehydratase

1.1.1 Organisation of shikimate pathway enzymes

The seven enzymes of the shikimate pathway are structurally and mechanistically similar in all species capable of aromatic amino acid biosynthesis, but there are considerable differences in the supramolecular organisation of the enzymes. In *Escherichia coli* (and other prokaryotes like *Salmonella typhi*) the enzymes are monofunctional polypeptides and their genes are widely scattered throughout the chromosome (Table 1-1) (Bachmann, 1983; Sanderson and Roth, 1983).

Table 1-1 The shikimate pathway enzymes of *E. coli* (taken from Haslam, 1993).

Enzyme	Gene	Map pos. (min)	Amino acids	M_r (calc.)	Quaternary structure
DAHP synthase					
DAHP synthase (tyr)	*aroF*	57	356	38 804	Dimer
DAHP synthase (phe)	*aroG*	17	350	37 997	Tetramer
DAHP synthase (trp)	*aroH*	37	347	39 000	Dimer
3-Dehydroquinase synthase	*aroB*	75	362	38 880	Monomer
3-Dehydroquinase dehydratase	*aroD*	37	240	26 377	Dimer
Shikimate Dehydrogenase	*aroE*	72	272	29 380	Monomer
Shikimate Kinase					
Shikimate Kinase I	*aroK*	74	173	19 526	Monomer
Shikimate Kinase II	*aroL*	9	173	18 937	Monomer
EPSP synthase	*aroA*	20	427	46 112	Monomer
Chorismate synthase	*aroC*	51	357	38 183	Tetramer

In the case of *E. coli* the enzymes do not tend to aggregate (Coggins and Boocock, 1986) and are separable (Chaudhuri and Coggins, 1985). In fungi, the enzymes catalysing the five consecutive steps involved in the conversion of DAHP to EPSP (Fig. 1-1) form a single pentafunctional polypeptide, known as the AROM

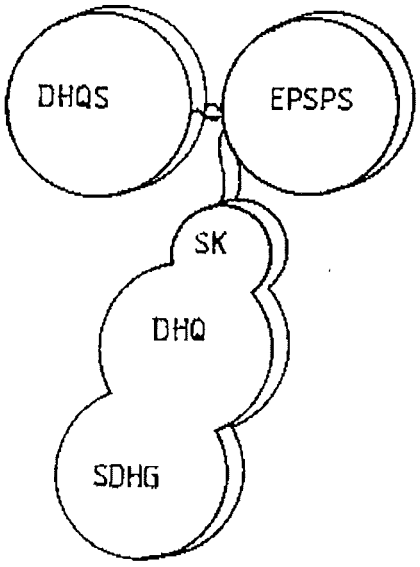


Figure 1-2 Model of the AROM protein as suggested by Hawkins and Lamb (1995).

DHQS	Dehydroquinate synthase
EPSPS	5-Enolpyrovyl shikimate 3-phophate (EPSP) synthase
SK	Shikimate kinase
DHQ	Dehydroquinase
SDHG	Shikimate dehydrogenase

complex (Giles *et al.*, 1967). The AROM protein was found to be the product of the *arom* gene cluster which consists of a single open reading frame (Duncan *et al.*, 1987; Charles *et al.*, 1986). The existence of the pentafunctional AROM protein has been demonstrated directly for *Neurospora crassa* (Lumsden and Coggins, 1977), *Aspergillus nidulans* (Gillies, 1994), *Saccharomyces cerevisiae* (Graham *et al.*, 1993), other fungal and yeast species and in *Euglena gracilis* (Patel and Giles, 1979). The *arom* genes have been cloned and characterised from *Aspergillus nidulans* (Charles *et al.*, 1986), *Saccharomyces cerevisiae* (Duncan *et al.*, 1987) and *Pneumocystis carinii* (Banerji *et al.*, 1993). A high degree of conservation can be observed in a multiple sequence alignment of the AROM proteins with the five monofunctional *E. coli* enzymes, supporting the hypothesis that AROM proteins have evolved by fusion of ancestral monofunctional genes (Hawkins, 1987).

The model of the AROM protein as shown in Figure 1-2 can be used to summarise our knowledge about that multidomain protein.

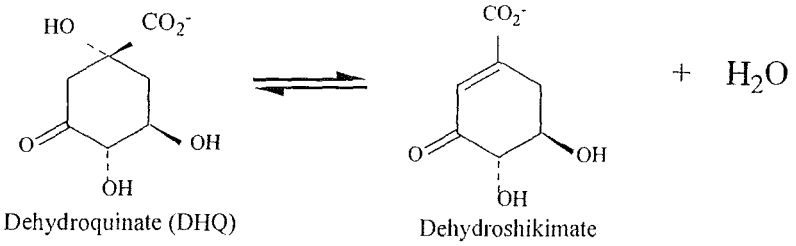
- The dehydroquinate synthase domain of the AROM protein (*A. nidulans*) can be overexpressed in *E. coli* and was found to retain efficient catalytic activity when compared with the intact pentafunctional AROM protein. (Moore *et al.*, 1994).
- The three C-terminal domains (SDHG, DHQ, SK) function most efficiently as a tri-domain unit (Hawkins and Smith, 1991).
- EPSP synthase is only active when covalently linked to DHQ synthase (Smith and Coggins, 1983; Moore and Hawkins, 1993).
- The AROM protein is a homo-dimer, DHQ synthase and shikimate dehydrogenase form part of the dimerisation interface (Case and Giles, 1971).
- The isolated DHQ domain is a very poor enzyme when compared to the native *Neurospora crassa* AROM protein (Hawkins *et al.*, 1993).
- Shikimate dehydrogenase and shikimate kinase domains are not active as monofunctional domains, but are active in combination with DHQ (Lamb *et al.* 1996).

The following advantages are usually put forward for the occurrence of multidomain enzymes: a) the facility to produce enzymatic activities in a fixed stoichiometric ratio, b) the potential for interactions between domains thereby giving the potential for allosteric interactions and c) the potential for metabolite channelling and the protection of unstable intermediates (Hawkins and Lamb, 1995). The latter seems to be unlikely in the case of the AROM protein since the intermediates are reasonably stable. The five enzyme activities of the AROM protein are separately measurable (Coggins *et al.*, 1987) which indicates an unhindered movement of substrates and products.

In plants the enzymes dehydroquinase and shikimate dehydrogenase which catalyse step three and four of the shikimate pathway, form a bifunctional enzyme whilst the other five enzymes are found to occur as monofunctional species (Polley, 1978; Mousdale *et al.*, 1987). The bifunctional enzyme gene has been cloned and characterised from *Pisum sativum* (Deka *et al.*, 1994) and *Nicotinaba tabacum* (Bonner and Jensen, 1994).

The following enzymes are being studied in this work.

1.1.2 3-Dehydroquinase (DHQ)



Dehydroquinate (DHQ) Dehydroshikimate

It is generally believed that enzymes have evolved to catalyse reactions by optimal mechanisms and it is therefore extremely unusual to find two mechanistically different enzymes that catalyse the same reaction. This is the case with the two types

of dehydroquinase (EC 4.2.1.10) which catalyse the dehydration of dehydroquinate to form dehydroshikimate. (Harris *et al.*, 1993).

This reaction is common to two metabolic pathways: the biosynthetic shikimate pathway for the synthesis of aromatic compounds in plants and micro-organisms (Haslam, 1974) and the catabolic quinate pathway (Fig. 1-1) in fungi which enables the organism to use quinate as a sole carbon and energy source (Giles *et al.*, 1985). There are two classes of DHQ which have different biochemical and biophysical properties (Table 1-2). The type I enzymes are only involved in the biosynthetic shikimate pathway (Chaudhuri *et al.*, 1986), whereas type II enzymes have been found to have either a biosynthetic (White *et al.*, 1990; Garbe *et al.*, 1991; Bottomley *et al.*, 1996b) or a catabolic role (Hawkins *et al.*, 1982) and in at least one species a dual role (Euverink *et al.*, 1992)(Table 1-2).

1.1.2.1 The type I enzymes

The type I enzymes are dimers with a M_r of about 56 000 , they are heat labile and catalyse a *cis* elimination of the elements of water (Hansen and Rose, 1963; Turner *et al.*, 1975). The enzyme mechanism involves the formation of a Schiff-base intermediate between the substrate and a lysine residue of the enzyme followed by the abstraction of a proton by a general base (Butler *et al.*, 1974; Chaudhuri *et al.*, 1991; Deka *et al.*, 1992). The resulting carbanion is stabilised by the Schiff-base which functions as a electron sink. In *E. coli* this lysine residue was identified as K-170 by trapping the Schiff-base intermediate using borohydride reduction and is conserved in all type I sequences (Chaudhuri *et al.*, 1991). The formation of a covalent enzyme-substrate intermediate might be the reason for the unusual *cis*-stereochemistry for the elimination of water. The enzymatic elimination of water usually proceeds in *trans* mode (Hupe, 1987). Using site specific chemical modification and site-directed mutagenesis the general base has been shown to be His-143 (Deka *et al.*, 1992; Leech *et al.*, 1995). Furthermore, Met-23 and Met-205 have been identified as likely active site residues but their role in catalysis has not been established (Kleanthous and Coggins, 1990).

1.1.2.2 The type II enzymes

The type II enzymes are heat stable dodecamers with a M_r of about 190 000 (Table 1-2). Unlike type I enzymes type II enzymes catalyse a *trans* elimination of the elements of water (Harris *et al.*, 1993) and the enzyme was found to be resistant to borohydride treatment (Kleanthous *et al.* 1992). This observation together with the fact that there are no conserved lysine residues in a multiple sequence alignment of type II dehydroquinases provide evidence that the enzyme mechanism does not involve the formation of a Schiff-base. There are only a few mechanistic details known about the type II enzymes. An enolate intermediate in the enzyme mechanism has been proposed by Harris *et al.* (1996).

The two types of DHQs can be compared with the two classes of aldolases which catalyse the conversion of fructose 1,6-bisphosphate to dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate (Lai and Horecker, 1972). The mechanisms of both types of aldolase proceed via the formation of a carbanion intermediate. Type I aldolases, like the type I DHQs, stabilise the carbanion intermediate by forming a Schiff-base with the substrate (Grazi *et al.*, 1962) whereas the type II aldolases employ a tightly bound metal ion for carbanion stabilisation (Kadonaga and Knowles, 1983). However, very recent studies have shown that the type II DHQ of *Aspergillus nidulans* does not require metal ions for enzyme function (Bottomley *et al.*, 1996a) and it is therefore still unclear how a negatively charged intermediate is stabilised by the type II enzymes.

It has been suggested, that the two classes of DHQs have evolved as a result of convergent evolution (Hawkins, 1987). However, since both enzyme types are apparently unrelated at the sequence level (Kleanthous *et al.*, 1992) the only point of convergence seems to be the common overall reaction.

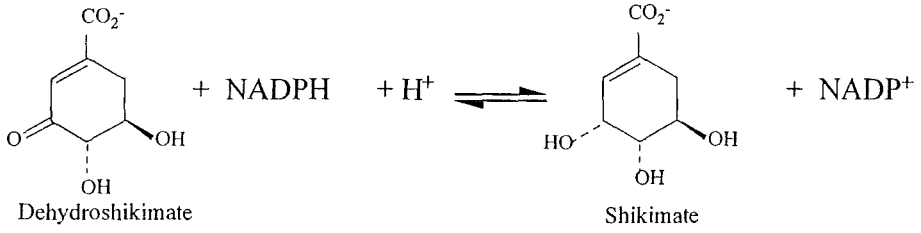
Type I and type II enzymes have been crystallised (Boys *et al.*, 1992; Gourley *et al.*, 1994) and the structure determinations are in progress.

Table 1-2 Biochemical and biophysical properties of dehydroquinases. (updated version of table from Kleanthous *et al.*, 1992)

S=shikimate; Q=quinate; Heat stability is indicated by survival of >90% of enzyme activity when heated to 70°C for 10 min. Unless otherwise stated, the K_m was measured at pH 7.0 (25°C); * estimates from the DHQ domain of the AROM protein; ‡ determined at pH 8.0 (37°C).

Organism	Path-way	Subunit	Heat Stability	sub-units	K_m	k_{cat}	k_{cat}/K_m	References
type I		M_r		No.	(μM)	(s^{-1})	($M^{-1}s^{-1}$)	
E. coli	S	27 466	No	2	16	135	8.4×10^6	(Chaudhuri *et al.*, 1986; Chaudhuri *et al.*, 1991)
S. typhi	S	27 706	-	2	18	200	1.1×10^7	(Servos *et al.*, 1991; Moore *et al.*, 1993)
N. crassa	S	-	No	2	5	-	-	(Lumsden and Coggins, 1977; Gaertner and Cole, 1977)
A .nidulans	S	25 600*	No	-	-	-	-	(Hawkins, 1987)
S. cerevisiae	S	26 377*	No	-	-	-	-	(Duncan *et al.*, 1987)
Pisum sativum	S	28 000	-	-	37	-	-	(Deka *et al.*, 1994)
type II								
A. nidulans	Q	16 500	Yes	12	150	1300	8.7×10^6	(Hawkins *et al.*, 1982; Da Silva *et al.* 1986, Bottomley *et al.*, 1996a)
N. crassa	Q	18 500	Yes	12	70	-	-	(Hautala *et al.* 1975; Geever *et al.*, 1989)
M tuberculosis	S	14 000	Yes	-	-	-	-	(Garbe *et al.*, 1991)
A. methanolica	Q/S	12 000	Yes	12	121	-	-	(Euverink *et al.*, 1992)
S. coelicolor	S	16 000	Yes	12	650‡	50	7.7×10^4	(White *et al.*, 1990)
H. pylori	S	18 481	Yes	-	590	2	3389	(Bottomley, 1996b)

1.1.3 Shikimate Dehydrogenase (SDH)



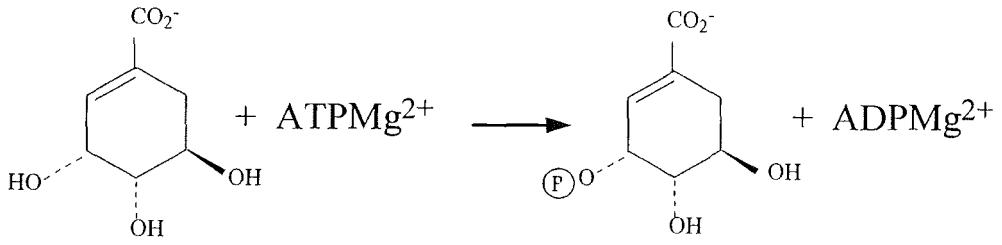
The fourth step of the shikimate pathway, the reduction of dehydroshikimate to shikimate is catalysed by shikimate dehydrogenase (EC 1.1.1.25). The enzyme was found to be strongly selective for NADPH over NADH (Balinsky *et al.*, 1971). Early studies have shown that the enzyme can be inhibited by the cysteine specific reagent p-hydroxymercuribenzoate (Balinsky and Davies, 1961) and by a number of simple aromatic compounds like p-hydroxybenzoate or protocatechuic acid (Balinsky and Dennis, 1970).

In general there are two groups of pyridine nucleotide dependent dehydrogenases; enzymes with or without catalytically active metal ions like zinc or iron (Aronson *et al.*, 1989; Neale *et al.*, 1986). Many metal independent dehydrogenases like lactate dehydrogenase (LDH) or malate dehydrogenase (MDH) (Adams, 1987) have been shown to possess a essential histidine residue which forms a hydrogen bridge with the carbonyl group of the substrate. This leads to a polarisation of the carbonyl group which initiates the hydride ion transfer. In metal dependant dehydrogenases, like alcohol dehydrogenase (Adams, 1987), the initial polarisation is achieved by the metal ion. SDH does not appear to need metal for its function since EDTA has no inhibitory effect on enzyme activity and divalent ions have no activating effect (Balinsky and Dennis, 1970). This is consistent with the observation that SDH can be inactivated by treatment with the histidine specific reagent diethylpyrocarbonate (Chackrewarthy, 1995).

The *E. coli* SDH has been purified and the gene cloned in Glasgow. The enzyme was found to be a monomer with a calculated M_r of 29 414 which is unusual for a dehydrogenase (Chaudhuri and Coggins, 1985; Anton and Coggins, 1988). The availability of larger amounts of recombinant enzyme has facilitated studies of enzyme-substrate interactions. Substrate analogues which lack the C-4 and C-5 hydroxyl groups and the carbon-carbon double bond have been synthesised and enzyme kinetic parameters with these analogues determined (Bugg *et al.*, 1988). Interestingly, the C5-deoxy analogue was found to be a reasonable substrate

with a k_{cat} of 75 s^{-1} (compared with k_{cat}=100 s^{-1} for dehydroshikimate) which implies that the C-5 hydroxyl group has little effect on the specificity. In contrast the C4-deoxy substrate was a very poor substrate (k_{cat}=0.06 s^{-1}) indicating a crucial role for the C4-group in molecular recognition. This is in agreement with the occurrence of the so called VDL motif which is conserved in several shikimate pathway enzymes (Bugg *et al.*, 1991). Since the C-4 hydroxyl group is present in all of the shikimate pathway intermediates it was suggested that this particular group forms a hydrogen bond with the aspartate residue of this motif (Bugg *et al.*, 1991).

1.1.4 Shikimate kinase (SK)



Shikimate Shikimate-3-phosphate

Shikimate kinase (EC 2.7.1.71), the fourth enzyme of the shikimate pathway, converts shikimic acid into shikimate-3-phosphate using ATP as a co-substrate (Weiss and Mingioli, 1956). In *E. coli* this reaction is catalysed by two different isoforms, SKI and SKII (Ely and Pittard, 1979; Berlyn and Giles, 1969). This is unusual for an enzyme in the middle of a metabolic pathway and it has been suggested that shikimate may be a branch point for another yet unknown pathway (Weiss and Edwards, 1980).

1.1.4.1 The two isoenzymes of Shikimate kinase

In *E. coli*, the *aro* genes which encode the enzymes of the shikimate pathway are not clustered into a single operon but are scattered around the chromosome (Table 1-1) (Pittard, 1987). The gene for SKII has been designated as *aroL*, which is the first gene of the two gene operon *aroLM* and its expression is regulated by the *trpR* and *tyrR* regulator gene (Ely and Pittard, 1979; De Feyter *et al.* 1986; Heatwole and Sommerville, 1992). The *tyrR* protein is of particular interest since it modulates the expression of at least eight unlinked operons (Cornish *et al.*, 1986). Seven of these operons are regulated in response to changes in

the concentration of the three aromatic amino acids, suggesting that these amino acids bind as co-factors to the *tyrR* protein to form an active regulatory molecule (Cornish *et al.*, 1986).

In contrast to that of SKII, the activity of SKI in the cell is independent of both the amount of extracellular aromatic amino acids and the level of *tyrR* gene product (Ely and Pittard, 1979). One other *aro* gene which is not subject to *tyrR* repression or to end product repression is *aroB* (Table 1-1), the gene for dehydroquinate synthase, the second enzyme of the shikimate pathway. Based on that observation it was possible to locate the *aroK* gene, encoding SKI, in the region upstream of *aroB* (Løbner-Olsen and Marinus, 1992; Whipp and Pittard, 1995). Both genes form another two gene operon.

The gene for SKII has been cloned and overexpressed and the enzyme was found to be a monomer with a calculated mass of 18 937 Da (Millar *et al.*, 1986; De Feyter and Pittard, 1986a). The apparent K_m for shikimate was 200 μM and 160 μM for ATP (De Feyter and Pittard, 1986b). Type I and type II enzymes share 30% sequence identity and are monomers of exactly the same length (173 amino acids) (Whipp and Pittard, 1995). However, in comparison to the type II enzyme the type I enzyme has a very much lower affinity for shikimate (K_m=20mM) and its expression is constitutive. Therefore, SKII appears to play a dominant role in the shikimate pathway. The role of SKI is less clear. It has been suggested that in *E. coli* SKI has been displaced by the catalytically more efficient and better regulated SKII (De Feyter and Pittard, 1986b). Another possible explanation might be that SKI normally carries out other functions in the cell and that it phosphorylates shikimate only fortuitously (De Feyter and Pittard, 1986b). This is consistent with the observation that SK I is associated with sensitivity to the antibiotic mecillinam which clearly implies an alternative biological role for SK I (Vinella *et al.*, 1996).

1.1.4.2 Shikimate kinases and the A-motif

From sequence comparison and crystallographic data it has been shown that a considerable number of proteins that bind ATP or GTP share a number of more or less conserved sequence motifs (Walker *et al*, 1982; Moller and Amons, 1985; Fry *et al.*, 1986; Dever *et al.*, 1987; Saraste *et al.*, 1990). The best conserved of these motifs is a glycine-rich region, which forms a loop between a beta-strand and an alpha-helix (Mueller and Schulz, 1992). This motif is generally referred to as the A-motif or P-loop and its consensus pattern is: [AG]-x(4)-G-K-[ST] (Walker *et al.*, 1982). This loop forms the ATP-binding site of

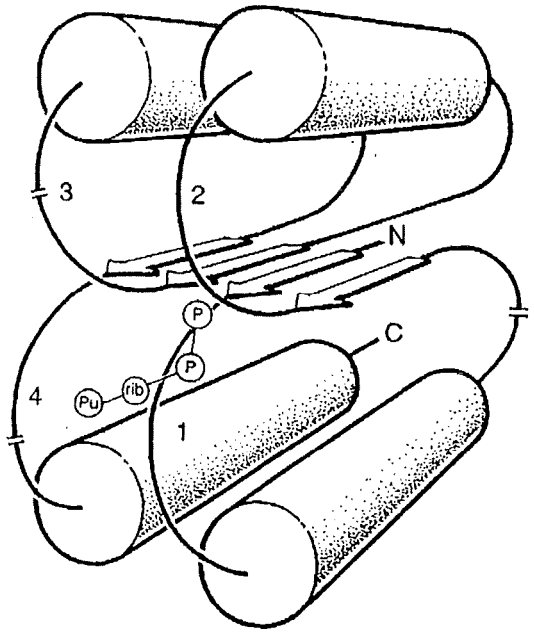


Figure 1-3 Core motif for ATP/GTP binding proteins as proposed by Milner-White *et al.* (1991).

these proteins. The peptide-mainchain atoms in the loop contribute to the formation of hydrogen bonds with the phosphate groups of the nucleotides.

There are numerous ATP- or GTP- binding proteins which contain the A motif; examples are the ATP dependent helicases (Lindner *et al.* 1989). Both SKI, SKII and the shikimate kinase domain of the AROM complex possess the type A motif. However, not all ATP-or GTP-binding proteins are picked up by this motif. A number of proteins escape detection because the structure of their ATP-binding site is different from that of the A-motif, for example the glycolytic kinases (Bennett and Steitz, 1980).

It has been proposed by Milner-White *et al.* (1991) that this A-motif could be extended to a common core structure for ATP/GTP binding proteins. This core region (Fig. 1-3) consists of about 100 amino acids which form four central, parallel beta-strands and-four alpha-helices. The A motif is in the loop connecting the N-terminal beta-sheet to the first alpha-helix.

1.1.4.3 Shikimate kinases in other species

SK activity has been detected in several plant species (Koshiba, 1979; Mousdale and Coggins, 1985). Schmidt *et al.* (1990) purified SK from spinach chloroplasts to near homogeneity. The enzyme was shown to be a monomer with a M_r of 31 000 . The gene for SK has been cloned from of *Lycopersicon esculentum L.* (tomato) (Schmid *et al.*, 1992). The open reading frame has the capacity to encode a peptide of 300 amino acids and the deduced amino acid sequence shows homology to bacterial and fungal SKs. The SK domains of the AROM complexes also show homology to the *E. coli* monofunctional enzymes. This is particularly obvious in the region around the A-motif.

The SK of *Erwinia chrysanthemi,* studied in this work, is a SKII enzyme and has 53% amino acid sequence homology to the *E. coli* SKII. The enzyme was cloned and sequenced by Minton *et al.* (1989). The molecular weight of the enzyme, as derived from the nucleotide sequence, is 18 955 Da. The enzyme was predicted to have a three-dimensional structure similar to adenylate kinase, which also contains the A-motif (Matsuo and Nishikawa, 1994).

1.2 Phosphoglycerate mutase (PGAM)

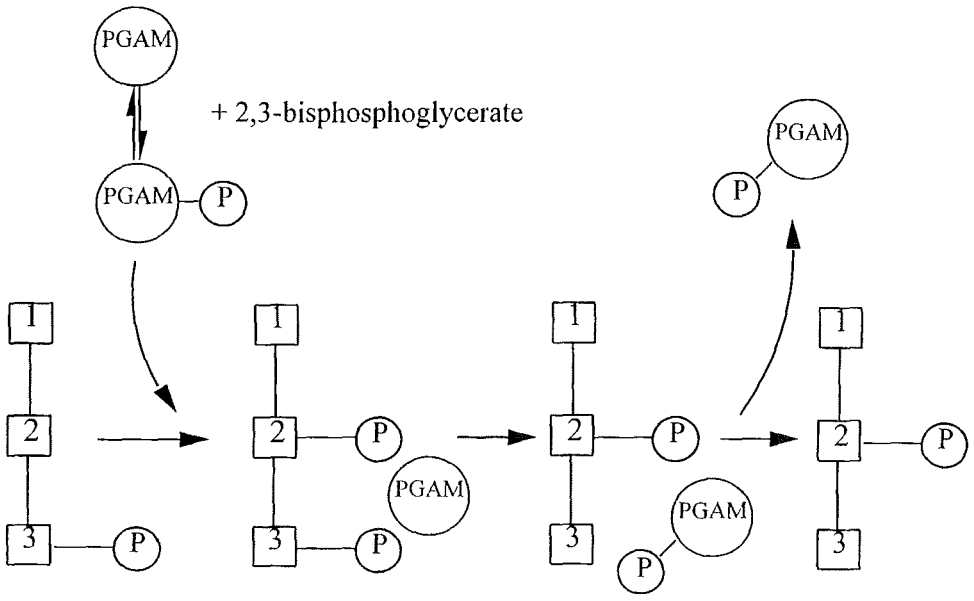


Figure 1-4 Conversion of 3-phosphoglycerate into 2-phosphoglycerate by 2,3-BPG dependent phosphoglycerate mutases (PGAM).

Phosphoglycerate mutase (PGAM) is a glycolytic enzyme which catalyses the interconversion of 3-phosphoglycerate and 2-phosphoglycerate. There are two classes of enzymes; enzymes which need a catalytic amount of 2,3-bisphosphoglycerate (2,3-BPG) to be active and others which are active in the absence of 2,3-BPG (Fothergill-Gilmore and Watson, 1989). Enzyme species studied in this work belong to the first group. For those enzymes 2,3-BPG donates either its 2- or 3-phosphoryl group to a histidine residue at the active site of the enzyme to generate the active (phosphorylated) form of the enzyme (Rose, 1970; Rose *et al.*, 1975).

The isomerisation reaction (Fig. 1-4) begins with the binding of 3-phosphoglycerate to the phosphorylated enzyme, phosphohistidine of PGAM transfers its phosphoryl group to the 2-OH of the bound substrate to form 2,3-BPG. The 3-phosphoryl group is then transferred to the histidine to generate 2-phosphoglycerate. Apart from the mutase activity PGAM exhibit a phosphatase activity (Fothergill-Gilmore and Watson, 1989).

Chapter 4 describes the measurement of the rate of dephosphorylation of the phosphorylated form of PGAM using electrospray mass spectrometry.

1.3 Methodology

1.3.1 Site specific chemical modification

Site specific chemical modification is a useful technique to obtain information about structure and function of enzymes. A chemical reagent is placed in contact with the enzyme and a chemical reaction occurs in which the reagent will bind covalently to amino acid side chains and changes in biological activity are measured. Of the 20 amino acids, only those possessing a polar side chain are normally the object of chemical modification. The chemical reactivity of these groups is basically a function of their nucleophilicity. The hyper-reactivity of certain of these groups appears to be related to their catalytic function (Lundblad and Noyes, 1984). Few investigators have evaluated structural changes as a consequence of chemical modification; but in such instances only minor changes were seen (Plapp *et al.*, 1978). It can be argued that protein engineering techniques have taken the place of chemical modification experiments to study enzymes, but it's clear that the combined use of both techniques is extremely powerful (Czupryn *et al.*, 1995; Robertson *et al.*, 1994). Very useful information about active site residues of enzymes can be obtained by differential labelling (i.e. modification in the presence and absence of substrate) (Singer, 1967; Urushibara *et al.* 1992).

To obtain information from chemical modification experiments a relationship between the loss of biological activity and the rate of the modification of amino acid residues needs to be established (Ray and Koshland, Jr., 1961). Therefore, it is critical to establish the stoichiometry of reaction. If the enzyme-reagent complex has a specific spectral absorbance the extent of reaction can be determined spectrophotometrically (Bray and Clarke, 1995). Other possibilities include the use of radiolabelled reagents (Gadda *et al.*, 1994) or amino acid analysis (Ray and Koshland, Jr., 1962) of the modified protein. However, it is thereby assumed that the modified protein is homogenous. This assumption is not always justified and can cause errors in the interpretation of the obtained measurements.

Another limitation of this technique is the question of the reaction specificity. Ideally, the reagent should be selective to react with only one type of residue, but most of the common reagents are known to undergo side-reactions with other types of residues. Therefore, interpretation of chemical modification data needs considerable caution (Lundblad and Noyes, 1984).

The following chemical reagents were used in this work.

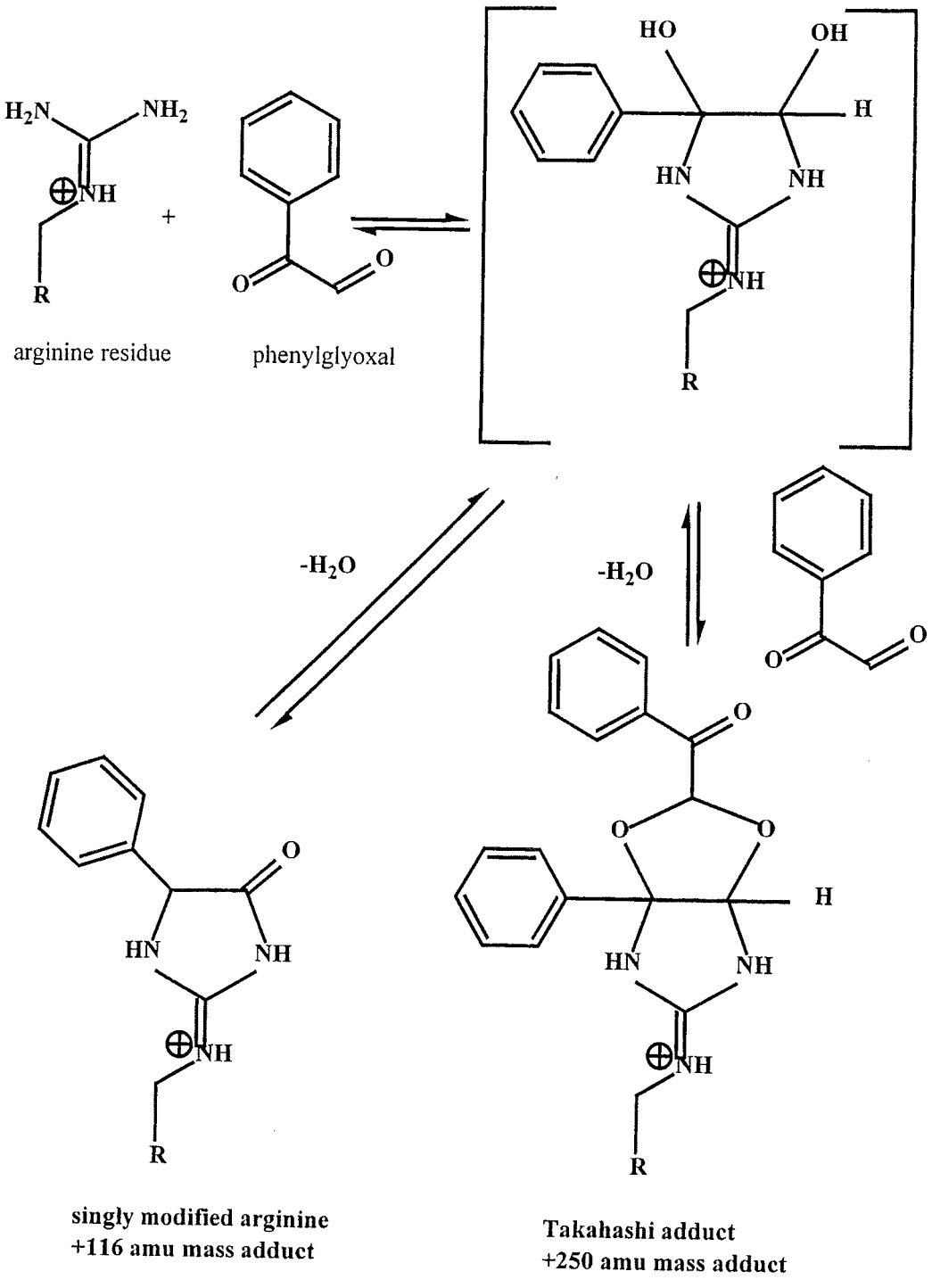


Figure 1-5 The stoichiometry of reaction between phenylglyoxal and arginine residues.

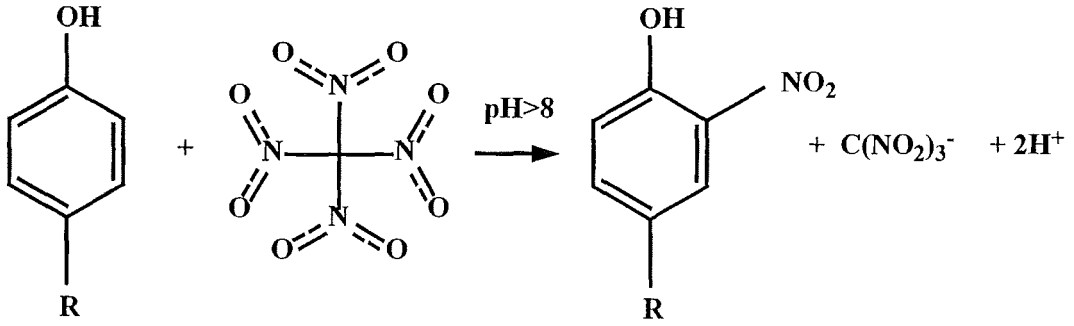
1.3.1.1 Phenylglyoxal (PGO)

PGO is an arginine specific reagent and was first introduced as a modification reagent by Takahashi (1968). He demonstrated that two molecules of PGO condense with one arginine residue (Fig. 1-5). Since no mono-substituted derivatives were found he concluded that the reaction of arginine with the first molecule of PGO is rate-limiting. Werber *et al.* (1975) showed that the modification is not homogeneous since he observed two modification sites after the incorporation of 1.5 mole PGO per mole enzyme and proposed the existence of stable 1 PGO : 1 arginine adducts (Fig. 1-5). In the same year Borders and Riordan (1975) reported the detection of a 1 : 1 adduct in borate buffer which appears to stabilise the 1 : 1 adduct. In the following years 1 : 1 adducts were detected even in non-borate buffer (Tedeschi *et al.*, 1992) while other authors reported an uncertain reaction stoichiometry in non-borate buffers (Eum and Miles, 1984). Despite these contradictory findings it is still common practice to calculate the number of modified arginine residues from the amount of incorporated ^{14}C PGO using Takahashi's 2 : 1 ratio (Kubiseski *et al.*, 1994; Qamar *et al.*, 1996)

The increased interest in establishing the functional role of arginine residues probably arises from the proposal of Riordan and co-workers (Riordan *et al.*, 1977; Riordan, 1979) that arginines function as general anion recognition sites in proteins. This theory was extended by Patthy and Thesz (1980) who suggested that arginine residues in anion binding sites have a lower pK_a than typical arginine residues due to a special micro-environment in the binding site. Therefore, arginines in substrate binding sites appear to be hyper-reactive towards dicarbonyles such as PGO. There is substantial experimental evidence confirming that theory (Borders and Riordan, 1975; Mohan *et al.*, 1988; Tedeschi *et al.*, 1992)

Chemical modification data obtained with PGO have to be discussed with caution since a secondary reaction of PGO with α-amino groups has been reported (Takahashi, 1968).

1.3.1.2 Tetranitromethane (TNM)

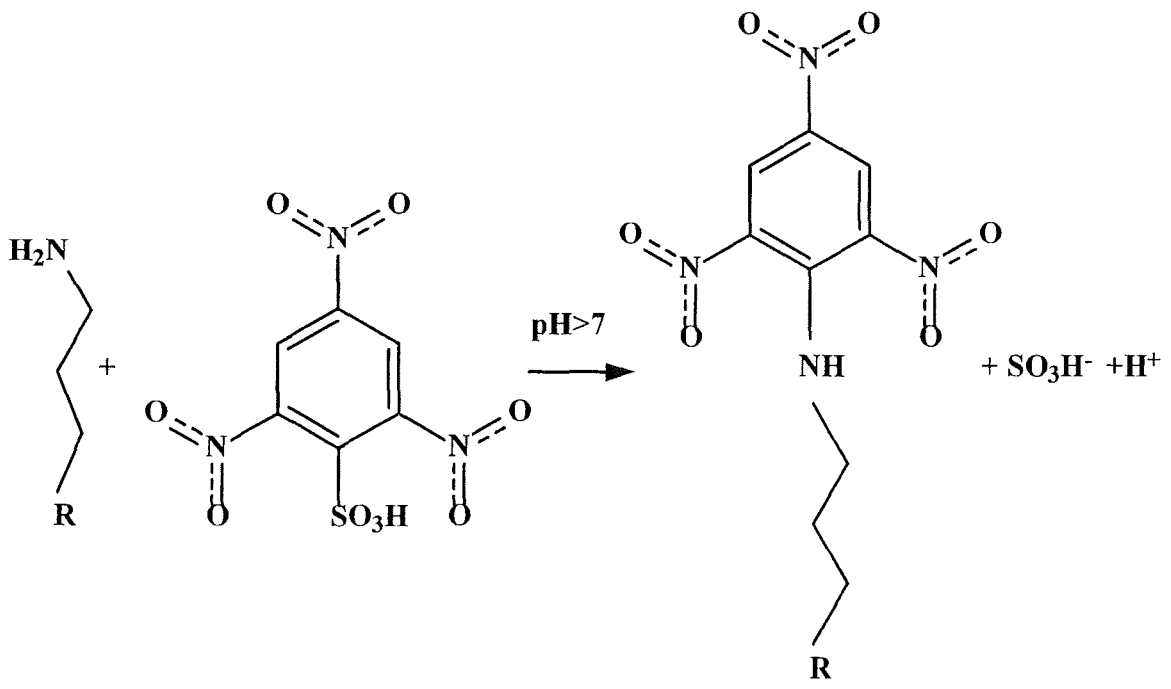


tyrosine residue tetranitromethane **mass adduct of 45 amu**

Tetranitromethane (TNM) is a widely used reagent for tyrosine modification studies. (Sokolovsky *et al.*, 1966). TNM nitrates tyrosine residues producing 3-nitrotyrosine which is a chromophore (characteristic absorbance at 428 nm) (Lundblad and Noyes, 1984). This offers the possibility of following the reaction spectrophotometrically. Essential tyrosine residues are mostly thought to be involved in substrate/coenzyme binding by establishing a hydrogen bond with an electrophile (Robertson *et al.*, 1994) or to be part of the catalytic mechanism as a proton donor (Mueller *et al.*, 1995). The presence of a strong electron withdrawing group at the 3-position lowers the p*K*a of the hydroxyl group of the tyrosine residue which will affect its potential to maintain a hydrogen bond. Alternatively, substitution of a bulky nitro group can make tyrosine less or even not available by steric hindrance.

The reaction of TNM with proteins is reasonably specific for tyrosine residues, however the oxidation of sulfhydryl groups has been reported (Sokolovsky *et al.*, 1969) and must be considered in interpreting the data.

1.3.1.3 Trinitrobenzenesulfonate (TNBS)



lysine residue TNBS **mass adduct of 211 Da**

The reaction of TNBS is a useful tool to study the function and reactivity of the ε-amino group of lysine residues (Goldfarb, 1966). The modification with TNBS is easy to follow by spectral analysis at 420 nm or 367 nm (Lundblad and Noyes, 1984). Reactivity is a sensitive measure of the basicity of an amino group. Adjacent charged amino acid residues can create a special micro-environment around the lysine residue and thereby altering its p*K*a. An increase in the rate of reaction was observed for lysine residues in a positively charged environment and a decrease for lysines in a negatively charged environment (Means *et al.*, 1972).

1.3.2 Mass spectrometry

Mass spectrometry has been used for quite some time in biochemical research to measure the molecular weight of small and stable molecules. Recent advances in biophysical research made it possible to apply this technique for larger and less stable molecules such as proteins. Mass spectrometry is based on the production, differentiation and detection of ions in the gas phase. The transfer of small and stable molecules into the gas phase has traditionally been accomplished by thermal vaporisation. However, this method is obviously of little use to transfer larger and labile molecules (proteins) into the gas phase. The search for a suitable ionisation source resulted in the development of ES-MS (electrospray mass spectrometry) (Fenn *et al.*,

1989) and MALDI-MS (matrix-assisted laser desorption ionisation mass spectrometry) (Karas and Hillenkamp, 1988).

1.3.2.1 Electrospray mass spectrometry (ES-MS)

ES-MS is based on the generation of singly or multiply charged gaseous ions directly from an aqueous or aqueous/organic solvent by creating a fine spray of highly charged droplets in the presence of a strong magnetic field. Drying gas and heat are applied to evaporate the solvent causing the droplets to decrease while surface charge increases. Multiply charged ions are transferred into the mass analyser (usually a quadrupole analyser) which is able to separate ions according to their mass-to-charge ratio (not according to their mass). The series of multiply charged ions is then used to calculate the molecular weight (Siuzdak, 1996).

ES-MS is now a very powerful method to determine the molecular weight of proteins and peptides. Conventional techniques, such as SDS-PAGE or ultracentrifugation produce measurements of molecular weights that are typically accurate to only ±5%, whereas ES-MS is generally better than ±0.01%. Furthermore ES-MS is a very rapid and sensitive technique. A typical molecular weight determination, including processing takes only 15 min and as little as 1 pmole of pure protein is needed for analysis (Siuzdak, 1996).

Another application that has created excitement involves the ability of ES-MS to analyse biological non-covalent complexes in the gas phase. Since the publication of the first two papers describing ES-MS as a tool for the observation of non-covalent complexes (Katta and Chait, 1991; Ganem *et al.*, 1991), this particular technique found a broad application in the study of molecular interaction, such as the observation of the catalytic antibody-hapten complex (Siuzdak *et al.*, 1994) or leucine zipper peptides (Li *et al.*, 1993).

ES-MS can also be used to probe protein conformational changes (Robinson *et al.*, 1994; Katta and Chait, 1993). This method is based on the mass-spectrometric measurement of hydrogen-deuterium exchange that occurs in different protein conformers over a time.

The main technical limitation of ES-MS is the sensitivity to the presence of non-volatile compounds such as buffers (Mann and Wilm, 1995). Ionic buffers disturb the spraying process and compete with analyte molecules for charges. Therefore it is essential to use volatile buffers or if possible only water. Other non-volatile substances can disturb the desorption process by forming a solid core in the evaporating droplet from which protein molecules cannot escape. The upper limits for non-volatile buffer substances like Tris/HCl is 5 mM but the quality of the spectra is always compromised (Pitt, 1996).

The processing of ES-MS raw data (series of peaks differing in their mass/charge ratio) using the Maximum Entropy (MaxEnt) deconvolution algorithm (Ferrige *et al.*, 1992) allows the semi-quantitative analysis of different protein species present in the analysed sample. The output of a MaxEnt deconvolution is a plot of relative abundance against molecular mass (not mass/charge as in the raw data) and quantification is achieved by integrating the area under those peaks.

1.3.2.2 Liquid chromatography ES-MS (LC/ES-MS)

All the applications mentioned above require protein solutions containing only a restricted number of distinct protein species. To work with more complex mixtures of proteins or peptides it is necessary to link a liquid-based separation system such as HPLC to the mass spectrometer. Monitoring such separations by ES-MS provides a very powerful means of analysing proteins and peptides.

This set-up is ideal to generate peptide maps of proteins. The protein is firstly fragmented into small peptides using a suitable protease, the peptide mixture is separated on a LC system and the eluate directly injected into the mass spectrometer. This technique was used successfully to localise post-translational modifications like phosphorylation (Meyer *et al.*, 1993; Taniguchi *et al.*, 1994) and glycosylation (Schindler *et al.*, 1995). LC/ES-MS based peptide mapping is very efficient compared with classical biochemical methods, in which each collected chromatographic fraction had to be sequenced by automated Edman degradation.

1.4 Aims of this project

- The main aim of the project was to identify active site residues in the two types of dehydroquinases and in shikimate dehydrogenase using site specific chemical modification. The feasibility of using ES-MS to measure the amount and to establish the stoichiometry of the chemical modification was tested. The sites of modification were localised using LC/ES-MS in combination with proteolytic digestion. Further experiments were carried out to elucidate the mechanistic role of active site residues identified (chapter 3 and 6).

- A secondary aim was to explore the scope of ES-MS to analyse other chemically modified enzyme species. In particular the method was used to characterise a phospho enzyme intermediate in the phosphoglycerate mutase reaction and to measure of the rate of hydrolysis of phospho enzyme and mutant phospho enzyme (chapter 4).

- The third aim was to characterise the P-loop lysine of shikimate kinase. This required purification and characterisation of the enzyme from *Erwinia chrysanthemi*. The *E. chrysanthemi* enzyme was also crystallised and its three dimensional structure determined (chapter 5).

2. Materials and Methods

2.1 Protein estimation

Protein concentrations were determined by the method of Bradford (1976) using bovine serum albumin as a standard, or spectrophotometrically at 280 nm in a Phillips Model PU8700 spectrophotometer using quartz cuvettes. The extinction coefficients were calculated according to Gill and von Hippel (1989); the values agreed to within 5-10%.

2.2 SDS-Polyacrylamide gelelectrophoresis (PAGE)

Electrophoresis in the presence of SDS was performed by the method of Laemmli (1970), with a 3% stacking gel and a 15% running gel in running buffer comprising Tris/HCl (3 g/l), glycine (15 g/l) and SDS (1 g/l). The ratio of acrylamide : bisacrylamide in all PAGE experiments was 30 : 0.8 and polymerisation was induced by 0.03% (v/v) TEMED and 0.05% (w/v) ammonium persulphate. Samples were denatured by boiling for 5 min after dilution in loading buffer comprising 60 mM Tris/HCl, pH 6.8, 2% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol and 0.0025% (w/v) bromophenol blue.

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

2.3 Synthesis of ammonium dehydroquinate

The substrate for dehydroquinases dehydroquinate was synthesised from quinate according to the procedure of Grewe and Haendler (1966) and stored as the ammonium salt in a sealed container at -20°C.

To check the purity 2 mg of the synthesised substrate was dissolved in 10 µl of water and separated on a HPLC Organic Acid Analysis column which consists of Aminex Ion exchange resin , 300 x 7.8 mm (Bio-Rad, Richmond, CA). The compounds were eluted with a constant flow (1 ml/min) of 10 mM acetic acid. Absorbances were recorded at wavelengths of 215 nm, 234 nm (absorbance of dehydroshikimate) and 280 nm. Fractions were collected and tested for dehydroquinate using the DHQ assay. For experiments only substrate batches were used with relative dehydroquinate concentrations of more than 90%.

2.4 Enzyme assay

2.4.1 Dehydroquinases

Enzyme activity was determined by monitoring the formation of 3-dehydroshikimate at 234 nm (ε=12x10^3 $M^{-1}cm^{-1}$) at 25°C. The assay mixture for the type I enzyme contained 100 μM ammonium dehydroquinate as substrate in 100 mM potassium phosphate (pH 7.0); for the type II enzyme of *Aspergillus nidulans* the assay mixture contained 1 mM ammonium dehydroquinate in 50 mM Tris/acetate (pH 7.0); the assay for the type II enzyme of *Streptomyces coelicolor* was carried out in 50 mM Tris/acetate, (pH 8.0) containing 2 mM substrate.

2.4.2 Shikimate kinase

Shikimate kinase was assayed at 25°C by coupling the release of ADP to pyruvate kinase (EC 2.7.1.40) and lactate dehydrogenase (EC 1.1.1.27). Shikimate dependent oxidation of NADH was monitored at 340 nm (ε=6180 $M^{-1}cm^{-1}$). The assay mixture contained 50 mM triethanolamine hydrochloride/KOH buffer, pH 7.0, 50 mM KCl, 5 mM $MgCl_2$, 1.6 mM shikimic acid, 5 mM ATP, 1 mM phosphoenolpyruvate, 0.1 mM NADH, 3 units of pyruvate kinase/ml and 2.5 units of lactate dehydrogenase/ml.

2.5 Enzyme purification

2.5.1 Dehydroquinases

Type I DHQ was purified from an overproducing strain of *Escherichia coli* according to the procedure of Chaudhuri *et al.* (1986). The type II DHQs from *Aspergillus nidulans,* overexpressed in *E. coli,* was purified as described previously (Beri *et al.*, 1990). DHQ from *S. coelicolor*, overexpressed in *E. coli*, was purified as follows:

All manipulations following cell breakage were performed at 4°C, unless otherwise stated.

Step 1: Extraction and centrifugation. A 20 g batch (wet weight) of *E. coli* BL21(DE3)pDHQpLysS was suspended in 20 ml 50 mM Tris/HCl, pH 7.5 (buffer A) and broken by two passages through a French pressure cell. This material was then centrifuged at 100 000 x g for 1 hour. DHQ was purified from the resulting cell-free extract.

Step 2. Anion exchange chromatography on DEAE-Sephacel. The supernatant was then applied (flow rate of 15 ml/hour) to a DEAE-Sephacel anion exchange column (250 x 21 mm, diam.) equilibrated in buffer A. The column was washed (flow rate of 50 ml/hour) with the same buffer until the A_{280} of the eluate was less than 0.3. The column was then washed (50 ml/hour) with buffer A containing 100 mM NaCl until the A_{280} of the eluate was again less than 0.3. Protein was eluted (50 ml/hour) with a gradient of 100-350 mM NaCl in 800 ml of buffer A and fraction (10 ml) collected and assayed as described.

Step 3. Hydrophobic interaction chromatography on Phenyl-Sepharose. Pooled fractions were made up to 1 M ammonium sulfate by addition of solid ammonium sulfate and applied to a Phenyl-Sepharose hydrophobic interaction column (100 x 9 mm, diam) equilibrated in buffer A containing 1 M ammonium sulfate. The column was washed with the same buffer over night (20 ml/hour). A 1-0 M ammonium sulfate gradient was run in a volume of 120 ml and 2 ml fractions collected and assayed as before.

Step 4. Anion exchange chromatography on Resource Q. This procedure was then carried out at room temperature using a Pharmacia FPLC system (Uppsala, Sweden). The enzyme solution was applied (1 ml/min) to an analytical Resource Q (Pharmacia, Uppsala, Sweden) anion exchange column (volume 6 ml) equilibrated in buffer A. Following washing with buffer A, protein was eluted (2 ml/min) with a gradient of 0-500 mM NaCl in buffer A and 1 ml fractions collected. After assaying for DHQ activity appropriate fractions were run on a 12% SDS-PAGE gel, dialysed against buffer A containing 30% glycerol and stored at -20°C.

2.5.2 Shikimate kinase

The gene encoding shikimate kinase of *Erwinia chrysanthemi* (Minton *et al.*, 1989) was cloned into pTB361 (Zeneca Pharmaceuticals, European patent application No. 92301456.8) and expressed in *E. coli* BL21(DE3)pLysS. The enzyme was purified as specified in chapter 5.

2.6 Site specific chemical modification of enzymes

Enzyme inactivations were carried out in a volume of 2 ml, at 25°C with constant stirring. During the inactivation enzyme aliquots were taken for enzyme assay. The percentage of inactivation was calculated as the ratio of enzyme activity after a certain time of treatment to the enzyme activity at time 0.

2.6.1 Dehydroquinases

2.6.1.1 Inactivation with phenylglyoxal (PGO)

Aliquots of DHQ (10-20 μg/ml) were incubated in 100 mM sodium bicarbonate buffer, pH 9.4 for 5 minutes and then PGO was added (freshly made up 50 mM stock solution in water) to a final concentration of 0.5-4.0 mM.

2.6.1.2 Inactivation with tetranitromethane (TNM)

Type II DHQ (10-20 μg/ml) was pre-incubated in 100 mM Tris/HCl, pH 8.0 for 5 min and then TNM (freshly made up stock solution of 15 mM in 95% ethanol) was added to a final concentration of 5-50 μM.

2.6.1.3 Inactivation with diethylpyrocarbonate (DEPC)

Inactivation was carried out by incubating DHQ (10-20 μg/ml) with DEPC in 50 mM potassium phosphate buffer, pH 6.0. A 50 mM solution of the reagent was freshly made up in ice-cold absolute ethanol before each experiment.

2.6.2 Shikimate kinase

2.6.2.1 Inactivation with PGO

Aliquots of SK (5 μg/ml) were incubated in 100 mM sodium bicarbonate buffer, pH 9.1 for 5 minutes and then PGO was added (freshly made up 50mM stock solution in water) to a final concentration of 0.5-4.0 mM.

2.6.2.2 Inactivation with trinitrobenzenesulfonic acid (TNBS)

The inactivation with TNBS was carried out in the dark. The enzyme (4 μg/ml) was pre-incubated for 5 min in 50 mM borate buffer, pH 9.2. A 500 μM TNBS solution in the same buffer was prepared and aliquots added to a final TNBS concentration of 0.5-2.5 μM.

2.6.2.3 Substrate protection against inactivation with PGO and TNBS

For substrate protection experiments with SK 50 mM stock solutions of shikimate, ATP with the molar amount of $MgCl_2$ and shikimate with ADP and molar amounts of $MgCl_2$were

made up in the inactivation buffer and the pH readjusted using 4M KOH. Defined aliquots of the stock solution were added to the pre-incubation mix and then the group specific reagent added.

2.6.3 Termination of inactivation reaction

The reaction was stopped by applying the reaction mixture to a G-50 Sephadex column (200 mm x 17 cm, flow rate 15 ml/min), equilibrated with 10 mM ammonium bicarbonate. The enzyme was eluted and 2 ml fractions were collected. Enzyme containing fractions were pooled and concentrated in Centricon 10concentrators (Amicon, Stonehouse, Gloucestershire, UK).

2.7 Proteolytic digest of DHQ

2.7.1 Digest of native and modified type II DHQ with chymotrypsin

All steps were performed at 25°C. The enzyme in water was denatured in 6 M GdnHCl in water for 15 min. The solution was diluted with ammonium bicarbonate (5 g/l) to a GdnHCl concentration of 1 M and incubated with 2% (w/w) chymotrypsin (stock solution of 1 g/l chymotrypsin in 1 mM HCl) for 45 min. The reaction was stopped by freezing the sample to -80°C.

2.7.2 Digest of native and modified type I DHQ with trypsin

Modified and native type I DHQ was digested at 37°C with continuous stirring. The enzyme in 0.4 M ammonium bicarbonate was denatured in 8 M urea for 1 hour. Water (37°C) was added to lower the urea concentration to 2 M and the mixture was incubated with 3% (w/w) trypsin (TPCK treated; 1 g/l stock solution in 1 mM HCl) for 4 hours. Afterwards another 1% (w/w) trypsin was added and the incubation continued for one hour. The reaction was stopped by freezing the sample to -80°C.

2.8 Electrospray mass spectrometry (ES-MS)

2.8.1 Sample preparation

To remove low molecular weight contaminants and buffer substances prior to mass spectrometry the enzyme samples were washed twice by diluting 50-fold with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse,

Gloucestershire, UK). Where proteins precipitated in water the same procedure was carried out using 1-5 mM ammonium bicarbonate.

2.8.2 The standard set-up

Mass spectrometry was performed on a VG Platform quadrupole mass spectrometer (2-3000 amu range) fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG Mass-Lynx software (VG Biotech. Ltd, Altricham, Cheshire, UK). Carrier solvent (1 : 1 (v/v) acetonitrile : water) infusion was controlled at 10 µl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA). Protein samples were dissolved in carrier solvent at a concentration of 20 pmol/µl , centrifuged at 5000 x g for 2 min and then 10-20 µl samples injected into the carrier stream. Capillary voltages were between 2.8 and 3.2 kV, extraction cone voltages 20-30 V, and the focusing cone voltage offset by +10 V. The source temperature was set at 65°C, the nebulising gas flow at 10 l/h, and the drying gas flow at 250 l/h. Lens stack voltages were adjusted to give maximum ion currents. The m/z range 700-1500, which contained >95% of the signal intensity for protein samples, was scanned at least 10 times with a sweep time of 5s. The instrument was calibrated over this M_r range immediately before use with horse heart myoglobin.

2.8.3 Processing of raw spectra

Raw spectra were processed using the MassLynx software (VG Biotech. Ltd, Altricham, Cheshire, UK).

2.8.3.1 Molecular weight determination

To determine the molecular weight of a protein the recorded raw spectra were processed using an algorithm which subtracted, smoothed and centred raw spectra followed by the actual mass calculation from the centred spectra. The following parameters were found to be suitable for processing a clean spectra with baseline resolution:

subtract:	10-15% below curve
smooth:	peak width of 0.5-1 Da, 1 smooth
centre:	2-4 Da peak width at half height
component detection:	manual, peak window 0.5-1.0 Da

2.8.3.2 Quantitative analysis

The Maximum Entropy (MaxEnt) deconvolution procedure (Ferrige *et al.*, 1992), was applied for quantitative analysis of raw data using 1.0-1.5 Da peak width and 1 Da/channel resolution. Typically, the MaxEnt procedure was stopped after 8-10 iterations.

2.8.4 Liquid chromatography ES-MS (LC/ES-MS)

The proteolytic digests were separated by HPLC on a C-4 reverse phase column (Delta-Pak^{tm} HPI C4, 2.0 x 150 mm, Waters, Watford, Hertfordshire, UK) using 2% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid as the initial solvent (flow rate 0.3 ml/min); after an 8 min wash to remove GdnHCl a linear gradient of 2-70% acetonitrile (v/v) in 0.1% trifluoracetic acid was applied to elute the peptides. The column eluate was introduced directly into the mass spectrometer with the drying gas flow of 400 l/h and the source temperature set at 100°C. The absorption profile of the eluted peptides was recorded at 214 nm and centroid spectra in the range 400-1800 m/z were recorded at 4 s intervals.

2.9 Circular dichroism spectrometry (c.d.)

The circular dichroism (c.d.) spectra were recorded in a Jasco J - 600 spectropolarimeter at the University of Stirling, BBSRC-funded c.d. facility by Dr. Sharon M. Kelly. Spectra in the far u.v. (260-190 nm) were recorded in cylindrical quartz cells of pathlength 0.02 cm and spectra in the near u.v. region (320-260 nm) were recorded in cells of pathlength 0.5 cm. The protein concentration was typically 0.5-0.6 mg/ml for far u.v. work and 5-6 mg/ml for the near u.v. c.d. experiments. All protein solutions were dialysed against 10 mM Tris/HCl, pH 7.5. The content of secondary structure elements was determined from the far u.v. spectrum using the CONTIN procedure (Provencher and Gloeckner, 1981).

2.10 Isothermal titration calorimetry (ITC)

Experiments were carried out in collaboration with Dr. Alan Cooper, University of Glasgow, Biophysical Chemistry Group. Binding of equilibrated substrate/product mixtures to native and mutant DHQ samples at 25°C was determined using a Microcal OMEGA isothermal titration calorimeter following standard ITC procedures (Wiseman *et al.*, 1989; Cooper and Johnson, 1994). Protein samples (ca.1 mg/ml), exhaustively dialysed against 10 mM Tris/HCl, pH 7.6, were degassed briefly before loading into the ITC cell (ca. 1.4 ml). Substrate (ammonium dehydroquinate, 50 mM) was dissolved in dialysis buffer together with a trace of native DHQ

and incubated at room temperature for 15 hours to give the equilibrated substrate/product mixture prior to loading into the ITC injection syringe. [Note: use of pure substrate in these experiments would result in large heat effects from the enzymatically catalysed reaction, rather the heats of binding required here.] A typical titration experiment involved a series of up to twenty-five 10 μl injections of ligand (substrate/product mix) into the enzyme solution at 3 min intervals, with continuos stirring. Integrated heat effects, after correction for dilution and mixing controls determined separately under identical conditions, were analysed by standard techniques using Microcal ORIGIN software assuming 1 : 1 enzyme : ligand complex formation.

2.11 Dynamic light scattering

The native molecular weight of enzymes was determined using a DYNA-PRO. 801 dynamic light scattering/molecular sizing instrument (protein solutions, Buckinghamshire, UK), the recorded data were processed using the AUTOPRO software. Protein samples were 1 mg/ml in 50 mM Tris/HCl, pH 7.5.

2.12 Protein crystallography

2.12.1 Crystallisation of Shikimate kinase

Enzyme samples were dialysed exhaustively into 20 mM Tris/HCl, pH 7.6. Afterwards shikimic acid (freshly made up solution in water, pH adjusted to 7.6 using 4 M HCl) and magnesium chloride were added to a final concentration of 5 mM and ADP (freshly made up solution in water, pH adjusted to 7.6 using 4 M HCl) was added to a final concentration of 10 mM. This increased the protein solubility from approximately 5 to at least 70 mg/ml. Enzyme was concentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, UK) to a concentration of 16 mg/ml. Crystallisation was achieved at 293K by the sitting-drop vapour-diffusion technique. In the final condition for crystallisation the reservoir solution contained 2.16 M Sodium chloride, 100 mM Hepes buffer, pH 6.9. Protein samples (6 μl) were mixed with equal amounts of reservoir solution and allowed to equilibrate. Crystals appeared after 10-12 days and continued to grow as tetragonals up to a maximum size of 0.7 x 0.2 x 0.2 mm.

2.12.2 The collection of a native data set

A complete native data set was collected at beamline 9.6 at the CLRC Daresbury Laboratory at 0.87 nm wavelength using a MAR Research imaging-plate scanner. The crystal was soaked in crystallisation buffer containing 12.5% (v/v) glycerol and mounted in a loop. The crystal was immediately frozen to 100K using a Oxford cryosystems cryocooler. A native data set diffracting to 1.9 Å resolution was collected as 1° and 1.5°oscillation frames.

2.13 Site-directed mutagenesis

2.13.1 Strains, plasmids and synthetic oligonucleotides

The bacterial strains, plasmids and synthetic oligonucleotides used in this project are listed in the tables below. Plasmid pDHQ was a kind gift of Prof. I.S. Hunter, University of Strathclyde, Glasgow.

2.13.2 Strategy

Three mutants of *Streptomyces coelicolor* DHQ (R23K, R23Q, R23A) were prepared using PCR site-directed mutagenesis. The procedure is summarised in Figure 2-1. Primers were synthesised on a Applied Biosystems Model 280A DNA synthesiser. The 3 mutating primer (Table 2-3) contain the mismatch codon (highlighted in bold, GCG for R) and cover the *Bgl*II site which is located 10 bp downstream of arginine residue 23. The reverse-primer (Table 2-3) was complementary to a region 190 bp upstream containing a *Eco*RV site. The PCR products were cloned back into the *Bgl*II/*Eco* RV site of pDHQ. For protein expression the entire mutated sequences were cloned into the *Nde*I/*Hind* III site of pT7-7 and overexpressed in *E. coli*BL21(DE3)plysS. All clones were verified by DNA sequence analysis using the Sanger methodology (Sanger *et al.*, 1980) and Sequenase Version II.

Table 2-1 Bacterial strains and their genotypes.

Bacterial Strain	Genotype
Escherichia coli DH5α	F- Φ80d*lacZ*ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17(r_K-m_K+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF*)U169
Escherichia coli BL21(DE3)	F- *ompT* r_B-,m_B- λ_{DE3}

Table 2-2 Plasmids used.

Plasmid	Remarks	Reference
pIBI25	T7 expression plasmid	(Dente *et al.*, 1983)
pDHQ	pIBI25 containing the DHQ gene of *S. coelicolor*	text
pT7-7	T7 expression plasmid	(Gilbert, 1991)
pLysS	T7 lysozyme plasmid	(Moffatt and Studier, 1987)
pDHQ1-3	pIBI25 containing the mutated DHQ genes of *S. coelicolor*	text
pTK1-3	pT7-7 containing the mutated DHQ genes of *S. coelicolor*	text

Table 2-3 Synthetic oligonucleotides used.

Primer	Sequence	Location in pDHQ	purpose
DHQ1	5' CCAGATATCAAATTAATACGACTCACTATAGG 3'	232-264	PCR rev. primer
DHQ2	5' GTAGATCTCCGGCTG**CTT**CTGGCCGAGC 3'	379-407	PCR (R23K)
DHQ3	5' GTAGATCTCCGGCTG**CGC**CTGGCCGAGC 3'	379-407	PCR (R23A)
DHQ4	5' GTAGATCTCCGGCTG**CTG**CTGGCCGAGC 3'	379-407	PCR (R23Q)
DHQ5	5' TACTCGCACACGTCCGTC 3'	569-587	sequencing
DHQ6	5' TTGGAGATGTGGACCTCC 3'	631-649	sequencing

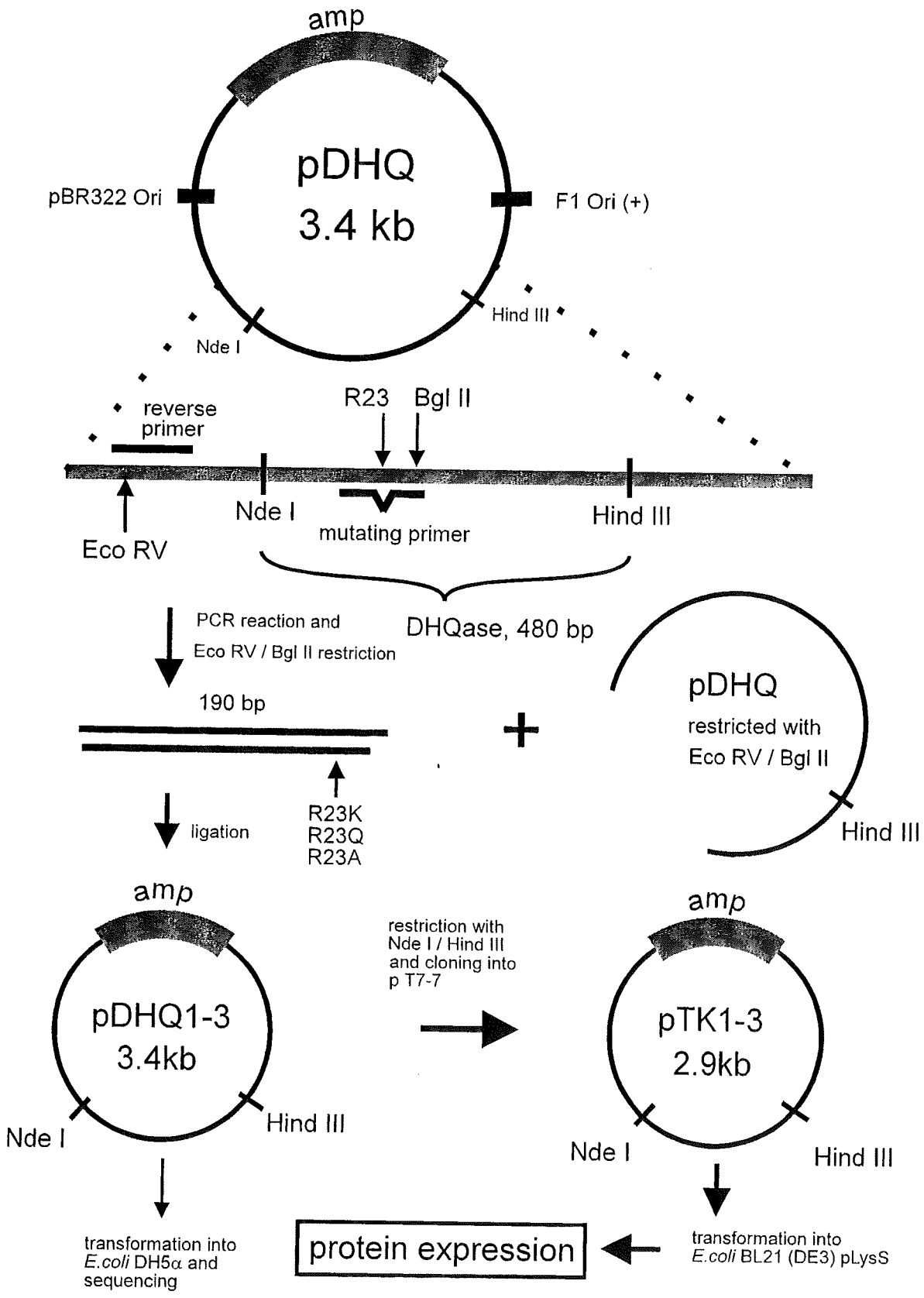


Figure 2-1 Site-directed mutagenesis of R23 of *S. coelicolor* DHQ.

2.13.3 Culture media

Bacteria were routinely cultured in Luria-Bertani medium (LB)(Bacto tryptone 10 g/l, yeast extract 5 g/l, NaCl 10 g/l) and grown overnight at 37°C with shaking. For culture on solid media, Bacto agar (Difco, Detroit, USA) was added at 1.5% (w/v) to the above medium. Antibiotics were added to the culture medium at the following concentrations: ampicillin 100 μg/ml, tetracycline 12.5 μg/ml, chloramphenicol 17 μg/ml (all from Sigma). Stock solutions were made up in water or ethanol as appropriate and filter-sterilised.

2.13.4 Storage of strains

The long term storage of bacteria was accomplished by making glycerol stocks of overnight cultures with the addition of sterile 80% (v/v) glycerol to a final concentration of 15% (v/v). These were stored at -80°C and never defrosted. For inoculation an upper piece of the frozen culture was chipped or melted off. Short-term storage was accomplished with the use of agar plates sealed with tape and maintained at 4°C.

2.13.5 PCR reaction

The PCR reactions were performed in 100 μl volumes (PCR tubes) using Vent polymerase (New England Biolabs, USA). The following components were added sequentially: 10 μl of 10 x Vent buffer, water, 10 μl of 10 x dNTPs (30 mM), 100 ng template DNA and 100 pmol of each primer. The mixture was overlaid with paraffin oil and kept on ice. The cycle profile used was: 94°C for 1min, 55°C for 1min, 72°C for 30 sec, with 25 cycles being performed. The PCR tubes were placed into the thermal cycler once it had reached 94°C.

2.13.6 Recovery of PCR products from reactions

20 μl of PCR product was run on a 1% (w/v) agarose gel to determine whether the PCR reaction had produced the desired amplion and whether other secondary PCR products were present. When only the desired amplion was observed after agarose gel electrophoresis a Wizard PCR clean-up kit (Promega, USA) was used with elution of the DNA into sterile distilled water.

2.13.7 DNA cloning procedures

2.13.7.1 Restriction and purification of plasmid DNA

5 µg of plasmid was digested in a volume of 50 µl with 10-20 U of restriction enzyme in the corresponding buffer at 37°C for one hour. Restricted plasmid was separated on a low melting point gel and the band excised. An approximately equal volume of TE (10 mM Tris/HCl, pH 7.2, 1 mM EDTA), pH 7.2 was added together with an equal volume of TE saturated phenol, pH 7.2 and heated to 65° for 5 min, vortexed and spun at full speed for 5 min. The aqueous layer was then promptly removed and further extractions of the aqueous layer with phenol, TE-saturated phenol/chloroform and TE saturated chloroform were performed. The DNA was recovered by adding an equal volume of isopropanol and 1/10 th volume of 1 M NaCl, vortexing and spinning at full speed in a microfuge for 15 min. The pellet was then dried and resuspended in a small volume of sterile, distilled water.

2.13.7.2 Alkaline phosphatase treatment of restricted plasmid

Calf intestinal alkaline phosphatase (CIAP) was added to the digested plasmid at a concentration of 1.0 U per pmol of 3' or blunt ends and 0.1 U per pmol of 5' ends. The molar amount of plasmid ends was calculated according to the following formula:

$$\frac{\text{mass of plasmid [g] x No. of bases}}{\text{Mr of plasmid}}$$

CIAP buffer was added and the volume made up to 100 µl and incubated at 37°C for 1 hour. The enzyme was then subsequently removed using a Magic DNA clean-up column (Promega, USA) with elution of the purified DNA in water.

2.13.7.3 Ligation of cut moieties

Prior to ligation, concentrations of DNA were either measured accurately by spectrophotometry or by estimation from gel electrophoresis. Several ratios of vector DNA and insert DNA (1 : 3 to 3 : 1 (w/w)) were used in separate ligations. The total amount of DNA per ligation was between 200-400 ng. The ligations were performed in 10 µl volumes with 1-3 units of T4 ligase and incubated overnight at 4°.

2.13.7.4 Preparation of competent cells and transformation

Competent cells were made on the day they were required. A single colony of the appropriate strain of *E. coli* was picked from an agar plate or a loop was used to melt a small amount of culture from the top of a glycerol stock. This was used to inoculate 20 ml of LB medium which was incubated at 37°C overnight with shaking. 30 ml of fresh LB medium was inoculated with 600 µl of overnight culture and incubated at 37°C for around two hours until the A_{600} reached around 0.3. The cells were then spun down at 3000 x g for 5 min and completely resuspended in one half volume of ice cold sterile 100 mM $CaCl_2$. The cells were left on ice for around 20 min followed by centrifugation at 3000 x g for 5 min followed by resuspension in one tenth volume of ice cold $CaCl_2$. The cells were usually left on ice for several hours prior to transformation.

For transformation of the cells, 5 µl of ligation mix was diluted with 45 µl of sterile, ice cold TNE buffer (50 mM Tris/HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA) on ice and 100 µl of competent cells were added. Transformation controls were performed at the same time and were made up to an equal volume. The mixture was incubated on ice for 30 min ensuring the contents were thoroughly mixed. The tubes were then transferred to a 37°C waterbath for 90 sec without agitation and then immediately returned to ice for a further 30 min. 1 ml LB was added and the tubes were incubated at 37°C for 1-2 hours to allow expression of antibiotic resistance. Suitable aliquots were then plated onto selective plates and incubated overnight at 37°C.

2.13.8 Plasmid DNA sequencing

Plasmid DNA sequencing was performed on DNA isolated using a Quiagen plasmid kit (Quiagen, Germany). Typically 2.5-3 µg of DNA was used for each reaction using a Sequenase 2.0 kit (Amersham, England). To denature the DNA: for each reaction, 8 µl of plasmid DNA (i.e. total volume contains 2.5-3 µg) and 2 µl of 1 M NaOH, 1 mM EDTA were mixed in an eppendorf tube and incubated at room temperature for 5 min. 3 µl of 3 M NaAc, 17 µl distilled water and 30 µl were then added. The tubes were centrifuged at full speed for 15 min, drained, vacuum desiccated and the DNA dissolved in 7 µl of water. For each annealing reaction, 2 µl of Sequenase reaction buffer and 1 µl of primer (1-2 pmol/µl) was added. This was heated at 100°C for 1 min before transferring the tubes to a 37°C waterbath for 20 min to allow the primer to anneal.

For the sequencing reactions: the ddNTP mixes were dispensed into wells of a microtitre plate (Nunc plates are heat resistant). For 4 reactions 10 µl of ddNTP mix and 1 µl of DMSO

were mixed together with 2.5 µl amounts dispensed per well. The labelling mixture (L) was made up by adding, for four reactions, 2 µl of labelling mixture, 1 µl of DMSO and 7 µl of water. To each template/primer mix 1 µl of 0.1 M DTT, 0.5 µl of α^{35}[dATP] and 2 µl of mixture L was added. The enzyme mixture (S), for 4 reactions, was prepared by mixing 10.6 µl of enzyme dilution buffer and 1.4 µl of Sequenase enzyme. 2 µl of this was then dispensed into each template/primer/label mix and incubated at room temperature for 5 min. Finally, 3.5 µl of the complete mixture was dispensed into the 4 corresponding ddNTP wells for each reaction and incubated at 37°C for 5 min. 4 µl of stop/loading dye was added and mixed before storing at -20°C until loading.

2.13.9 Sequencing gel and sample loading

Sequencing gels were poured the day before running and stored at 4°C overnight. A 6% sequencing gel was prepared by mixing 63 g of urea, 22.5 ml of 40% (w/v) acrylamide solution, 15 ml of 10 x TBE (900 mM Tris/HCl, 900 mM boric acid, 25 mM EDTA, pH 8.3) made to 150 ml with distilled water. 750 µl of 10% (w/v) ammonium persulphate and 125 µl of TEMED were then added, the solution mixed and poured without degassing. The samples were heated to 75°C-80°C and placed on ice immediately before loading. 3 µl of sample was loaded per well and the gel run at 40 mA for 4 hours or until the gel front reached the base of the gel. After electrophoresis the gel was fixed for 30 min in 2 litres of 10% (v/v) methanol, 10% (v/v) acetic acid then transferred to a sheet of 3MM Whatman paper and dried at 80°C under vacuum on a gel drier. An autoradiograph was then set up using Hyperfilm MP (Amersham, England) with exposure for around 18 hours and development using Kodak X-Omat.

2.13.10 Expression in pT7-7

The expression plasmid pT7-7 contains an ampicillin resistance gene and a T7 RNA polymerase promotor. A number of host *E. coli* strains were used during studies on the expression of genes using the T7 expression plasmids. *E. coli* BL21(DE3) was found to give the best level of expression. The latter strain was also used in conjunction with pLysS and pLysE plasmids. The inclusion of the gene for T7 lysozyme (which binds and inhibits T7 RNA polymerase) on plasmid pLysS provides means of controlling the basal levels of expressed protein prior to induction with IPTG.

For expression 20 ml of an overnight culture was grown in LB plus antibiotic(s). 50 ml of LB plus antibiotic(s) were then inoculated with 5 ml of overnight culture and grown to an A_{600}

of 0.6. The culture was then made 0.4 mM with respect to IPTG and grown for a further 5-6 hours removing 1 ml samples at time-points. The A_{600} of the samples were measured and then they were centrifuged at full speed for 1min in a microfuge. The cells were resuspended in 10 µl of SDS gel sample buffer for every 0.1 absorbance unit and boiled for 5 min before loading 20 µl per well onto a SDS gel.

2.13.11 Expression for protein purification

Large scale growth of bacteria for subsequent enzyme purification was accomplished by growing 500 ml cultures in 2 l flasks in a flat-box shaker at 37°C. Typically, flasks containing 500 ml of LB plus appropriate antibiotics were inoculated with 25 ml of overnight culture. The antibiotics concentrations used were:

*E. coli*BL21(DE3)pTB361SKpLysS - 12.5 µg/ml tetracycline and 17 µg/ml chloramphenicol

*E. coli*BL21(DE3)pDHQpLysS and
*E. coli*BL21(DE3)pTK1-3pLysS - 17 µg/ml chloramphenicol and 100 µg/ml ampicillin

Growth for 2-3 hours at 37°C was usually sufficient before induction with IPTG and growth for a further 4-5 hours at 37°C for expression.

3. Active site studies on type I and type II Dehydroquinases.

The use of electrospray mass spectrometry to identify active site arginine residues in type II dehydroquinases.

Tino Krell, Andrew R. Pitt and John R. Coggins (1995) *FEBS Lett.* **360**, 93-96.

Localisation of the active site of type II dehydroquinases. Identification of a common arginine-containing motif in the two classes of dehydroquinases.

Tino Krell, Malcolm J. Horsburgh, Alan Cooper, Sharon M. Kelly and John R. Coggins (1996) *Journal of Biological Chemistry* **271**, 24492-24497.

3.1 The use of electrospray mass spectrometry to identify active site arginine residues in type II dehydroquinases

3.1.1 Abstract

The arginine-specific reagent phenylglyoxal has been used to identify a hyper-reactive arginine residue which is essential for activity in the type II dehydroquinases of *Streptomyces coelicolor* and *Aspergillus nidulans*. Electrospray mass spectrometry was used both to characterise the phenylglyoxal modified protein, and to identify the phenylglyoxal modified peptides following enzymatic digestion. The advantages of using electrospray mass spectrometry for monitoring arginine modification aimed at identifying functional residues in proteins are discussed.

3.1.2 Introduction

The type II dehydroquinases (3-dehydroquinate dehydratases, EC 4.2.1.10) catalyse the conversion of dehydroquinate to dehydroshikimate. This reaction, which occurs on both the biosynthetic shikimate pathway and the catabolic quinate pathway [1-3], involves the *trans* elimination of water [4]. Little is known about the structure and mechanism of the type II dehydroquinases which are clearly mechanistically and structurally different from the better characterised class I enzymes [5,6]. The class I enzymes are exclusively biosynthetic [5], have a conserved active site lysine residue [7], and catalyse a *cis* elimination via an imine intermediate [4,8] with the participation of a conserved histidine residue as the general base [9]. In contrast the fungal type II enzymes have an exclusively catabolic role [2,10] while the bacterial enzymes may be exclusively biosynthetic [11,12] or be involved in both biosynthesis and catabolism [13]. There is preliminary chemical modification evidence for a role for histidine in the type II enzymes [5] but there are no conserved lysine residues and this together with the different stereochemistry and the failure of experiments to inhibit the type II enzymes with substrate and sodium borohydride [5] emphasises their mechanistic distinction from the class I enzymes.

A type II dehydroquinase has recently been crystallised [7] and to facilitate the structure determination we have been using group specific chemical modification reagents to identify amino acid residues in the active site. Many enzymes which use carboxylic acids as substrates utilise an arginine residue for carboxylate recognition [14] and in several cases the functional arginine has proved to be hyper-reactive [15]. The reagent phenylglyoxal (PGO) has been widely used to demonstrate the involvement of arginine residues in enzyme function [16,17]. However, the

identification of specific arginine residues is complicated by the relatively low stability of the adducts formed [17], their variable stoichiometry [17] and the necessity to use radiolabelled reagent. Also in the case of the type II dehydroquinases the identification of active site residues is further complicated by the relatively low affinity of the enzymes for both substrate and product and the lack of tight binding competitive inhibitors [5]. To determine whether there was an active site arginine a method was needed for analysing enzyme in the early stages of modification and correlating the modification of particular residues directly with the loss of activity. We have used electrospray mass spectrometry to monitor the reaction of two type II dehydroquinases with PGO. This has permitted the simple characterisation of singly and multiply modified enzyme molecules and, through the direct analysis of enzymatic digests of the modified enzyme, the direct location of the modification sites without the necessity for sequencing or the use of radioactive or other labels. By this means we have identified a single hyper-reactive arginine residue in the substrate binding site of the type II dehydroquinases.

3.1.3 Experimental

The type II dehydroquinases from *Streptomyces coelicolor* [11] and *Aspergillus nidulans* [18] were overexpressed in *Escherichia coli* and purified as described previously [19,18]. DHQ activity was determined as described previously [5] and protein concentrations were determined spectrophotometrically at 280 nm using extinction coefficients calculated from the amino acid compositions [20].

Enzyme samples in 100 mM sodium bicarbonate buffer, pH 9.4 were pre-incubated for 5 min at 25°C and then PGO (freshly made up 50mM stock solution in water) was added to a final concentration of 0.5-4.0 mM. Aliquots were removed at various times for enzyme assay. Enzyme inactivated to different extents for mass spectrometry was prepared by stopping the reaction by gel filtration on a Sephadex G50 column (200 mm X 17 mm, flow rate 15 ml h^{-1}) equilibrated with 10mM ammonium bicarbonate. To remove low molecular weight contaminants prior to mass spectrometry the enzymes samples were washed twice by diluting 50-fold with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, U.K.).

To prepare peptides from the native and modified enzyme protein samples were first denatured in 6 M GnHCl for 15 min and then diluted with 0.5% (w/v) ammonium bicarbonate to a GnHCl concentration of 1 M and incubated with 2% (w/w) chymotrypsin (stock solution 1 g/l chymotrypsin in 1 mM HCl) for 45 min at 25°C. The digestion was stopped by freezing the samples at -80°C.

Mass spectrometry was performed on a VG Platform quadrupole mass spectrometer fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG MassLynx software (VG Biotech Ltd., Altrincham. Cheshire, U.K.). Carrier solvent [1:1 (v/v) acetonitrile : water, 0.2% formic acid] infusion was controlled at 10µl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, Mass., U.S.A.). Protein samples were dissolved in carrier solvent at a concentration of 20 pmol/µl, centrifuged at 5000 x g for 2 min and then 10-20 µl samples injected directly into the carrier stream. MaxEnt deconvolution [21] was applied for quantitative analysis of the raw data using 1.0 Da peak width and 1 Da/channel resolution.

The peptide digests were separated by HPLC on a C-18 reverse phase column (mBondapak, Waters, Watford, Hertfordshire, U.K.) using 2% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid as the initial solvent (flow rate 0.5 ml/min); after an 8 min wash to remove GdnHCl a linear gradient of 2-70% acetonitrile (v/v) in 0.1% (v/v) trifluoracetic acid was applied to elute the peptides. The column eluate was introduced directly into the mass spectrometer with a drying gas flow of 400 l h^{-1} and the source temperature set at 100°C. The absorption profile of the eluted peptides was recorded at 214 nm and centroid mass spectra in the range 400-1800 amu were recorded at 4 sec intervals.

3.1.4 Results and Discussion

The inactivation of the *S. coelicolor* type II dehydroquinase with PGO followed pseudo-first order kinetics (Fig. 3-1). A secondary plot of the observed pseudo-first order rate constants against PGO concentration was linear (data not shown) and gave a second order inactivation rate constant of 89 $M^{-1}min^{-1}$.

Similar results were obtained for inactivation of the *A. nidulans* enzyme by PGO; in this case the second order rate constant was 150 $M^{-1}min^{-1}$. These observations suggested that arginine residue(s) were required for the type II dehydroquinase reaction. The relationship between the extent of inactivation and the number of arginine residues modified by PGO was monitored by mass spectrometry; the spectra observed for three samples inactivated to different extents are shown in Fig. 3-2. During the early stages of inactivation (exemplified by 15% inactivation) the major peak was of unmodified enzyme and the only significant modified species had a mass difference of +116 (M_r=16 666) (Fig. 3-2). This mass difference corresponds to the incorporation of one PGO per site with the loss of water (Fig. 3-3). For low extents of inactivation there was a good correlation between the extent of inactivation and the relative size of this +116 peak (Table 3-1).

Figure 3-1 Inactivation of the type II dehydroquinase of *S. coelicolor* by PGO.

Semi-logarithmic plot of residual activity as a function of time. PGO concentrations: 0.5 mM (▼), 1.0 mM (◓), 2.0 mM (▲) and 4 mM (●).

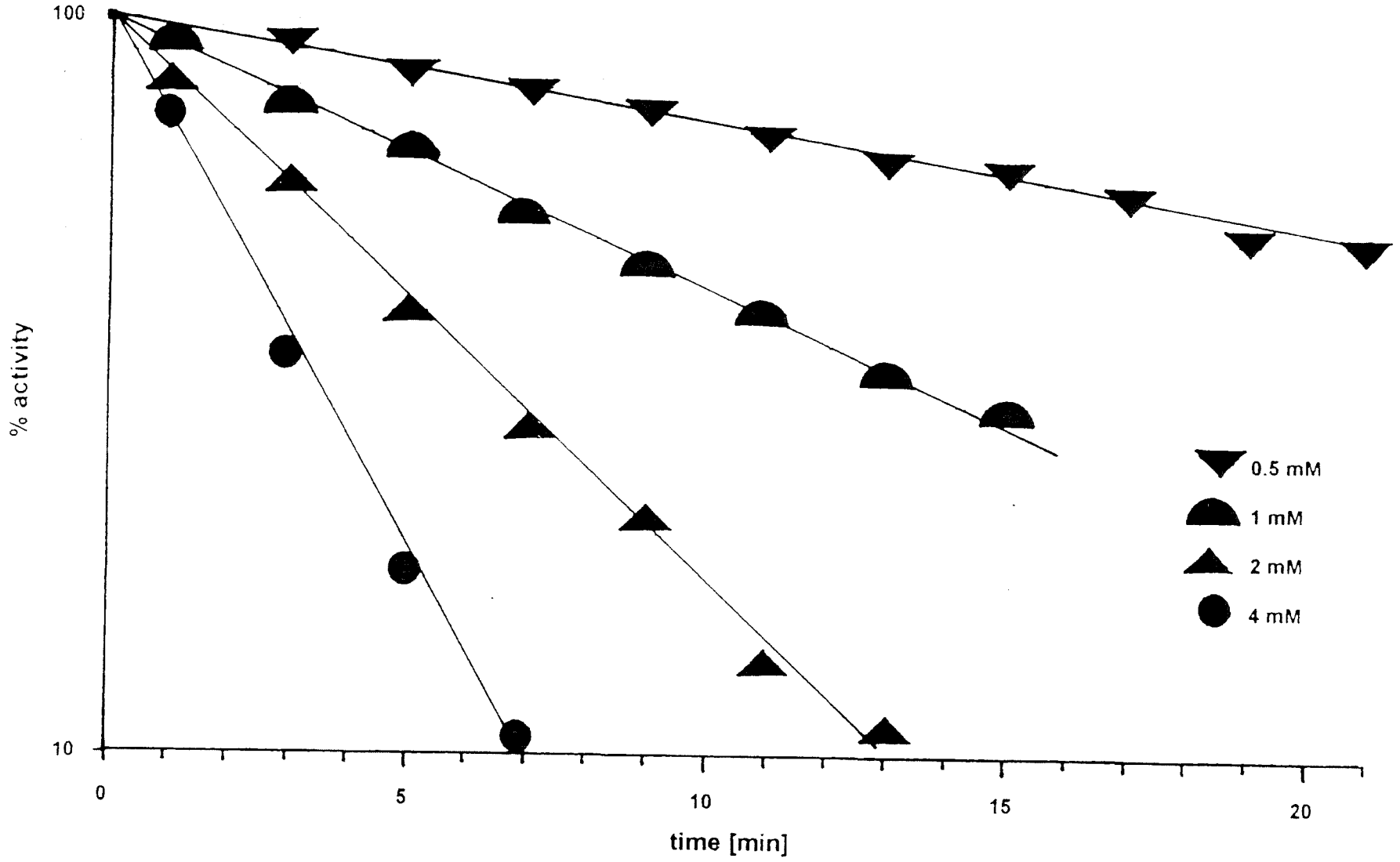

100
10
% activity
0
5
10
15
20
time [min]
0.5 mM
1 mM
2 mM
4 mM

Figure 3-2 Electrospray mass spectra of type II dehydroquinase from *S. coelicolor* at different extents of modification with phenylglyoxal.

a) 85% active, b) 60% active and c) 5% active.

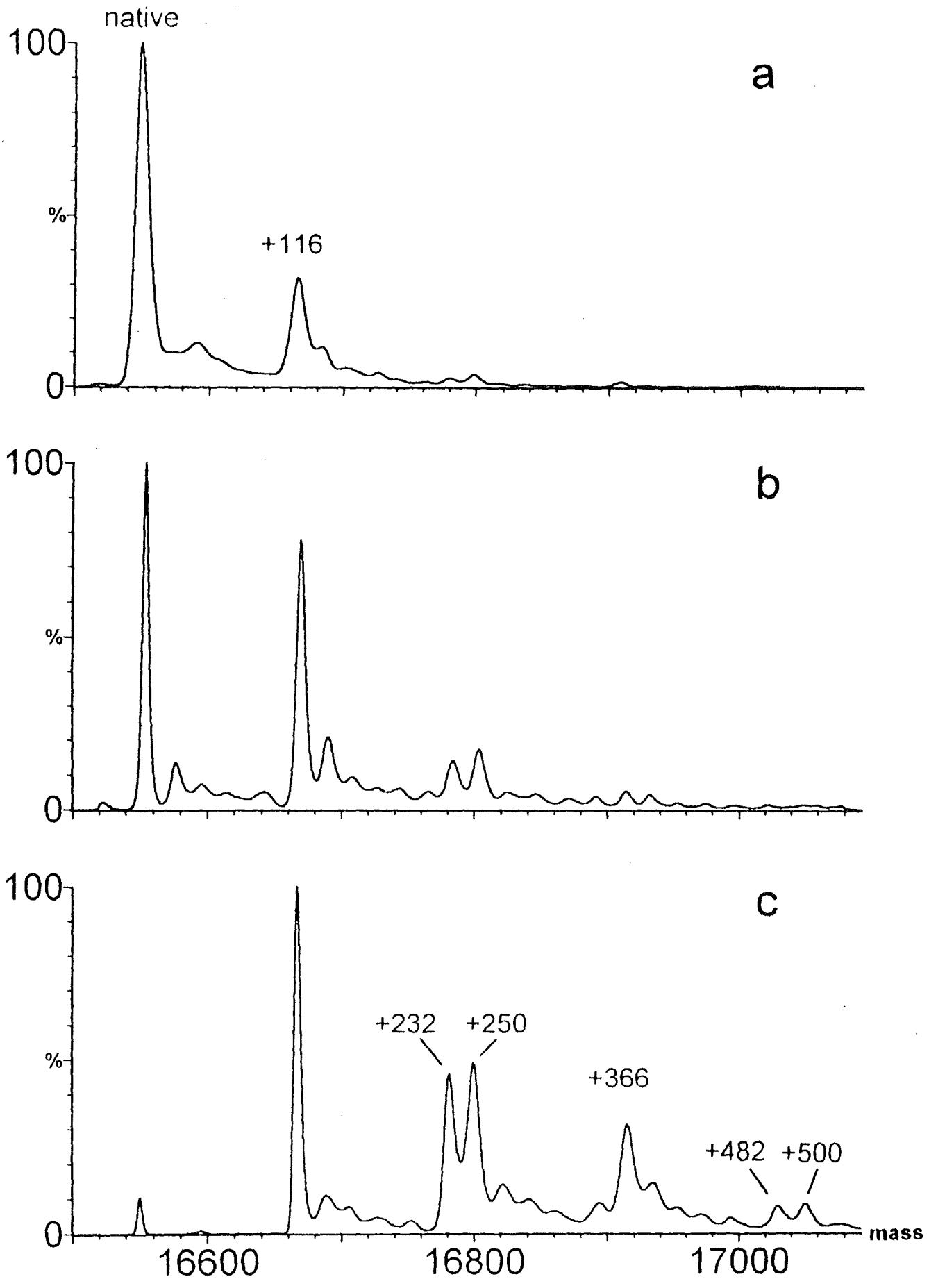

native
100
%
0
+116
a
100
%
0
b
100
%
0
c
+232
+250
+366
+482
+500
mass
16600
16800
17000

Table 3-1 Relative amounts of adducts formed between PGO and the type II dehydroquinases.

Samples were inactivated by treatment with PGO for various times, assayed for residual activity and then analysed by electrospray mass spectrometry (see Experimental section).

Enzyme source	% inactivation	Relative percentage of adducts						
		Native	+116	+232	+250	+366	+482	+500
S. coelicolor	15	77	23	-	-	-	-	-
	60	44	35	10	11	-	-	-
	95	2	31	19	21	18	4	5
A. nidulans	10	82	18	-	-	-	-	-
	30	66	34	-	-	-	-	-
	90	16	58	14	12	-	-	-

In the sample that was 60% inactivated the two major peaks due to the +116 species and to unmodified protein together accounted for 80% of the material (Fig. 3-2). A number of minor modified species were also present including one with a mass difference of +250 (M_r=16 800), corresponding to the Takahashi adduct with two PGO's per arginine (see Fig. 3-3), and another with a mass difference of +232 (M_r=16 782) corresponding to two different arginines, each modified with a single phenylglyoxal. These minor peaks were more significant in 95% inactivated enzyme (Fig. 3-2); in this sample very little unmodified protein remained, the singly modified +116 species was the most prominent peak and there were also small peaks due to multiply modified species for example at +232, +250, +366 (250 plus 116), +482 (250 plus 116 plus 116) and at +500 (250 plus 250). These data suggest that the initial inactivation of the enzyme is due to the formation of a simple 1:1 adduct with phenylglyoxal followed by dehydration.

To identify the primary site of modification samples of *S. coelicolor* dehydroquinase, modified to different extents with PGO, were digested with chymotrypsin and the peptides separated and analysed by reverse phase chromatography/electrospray mass spectrometry. In the 15% inactivated sample there was a single modified peptide of mass 3 219 which corresponds to residues 1-28 of the enzyme +116 mass units. This peptide contains 2 arginine residues (R2 and R23). In the aligned type II dehydroquinase sequences the *A. nidulans* enzyme has no residue corresponding to R2

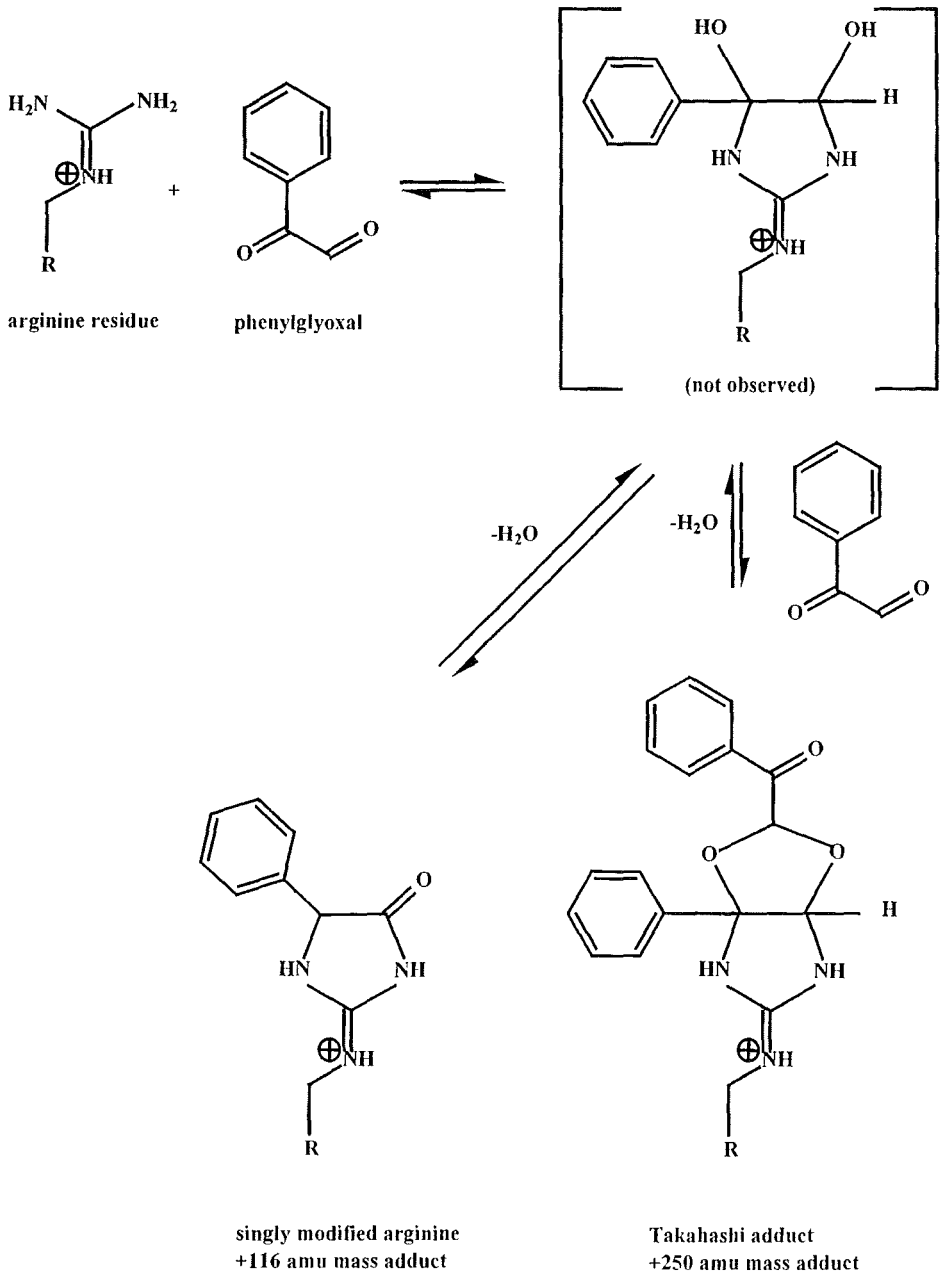


Figure 3-3 Structures of adducts formed by the reaction of PGO with arginine residues.

(*S. coelicolor*) but there is a so far totally conserved arginine residue, R19 (R23 in *S. coelicolor*) and so it seemed very likely that R23 was the site of initial modification in the *S. coelicolor* enzyme.Consistent with this observation the similarly modified *A. nidulans* enzyme was found to contain a peptide of mass 2852 which corresponds to residues 1-24 of the enzyme plus 116 mass units; this peptide contains a single arginine residue R19 which corresponds to R23 in the *S. coelicolor* enzyme.

In samples of 80% inactivated *S. coelicolor* enzyme the singly modified peptide 1-24 was found and in addition singly modified peptides corresponding to residues 84-116 (R113) and 139-156 (R144 and R155) were also observed. At this extent of modification it was not possible to detect any peptides containing the Takahashi adduct. The *A. nidulans* enzyme contains no arginine residues corresponding to R144 and R155. It does contain a residue corresponding to R113 but even at 95% inactivation no modification at this site was observed. The type II dehydroquinases contain an additional conserved arginine residue (R117 in the *S. coelicolor* enzyme). No evidence that this residue is modified by phenylglyoxal has been otained but it should be noted that peptides containing this residue have proved difficult to detect by mass spectrometry.

These results suggests that the conserved residue R23 in the *S. coelicolor* enzyme (R19 in the *A. nidulans* enzyme) is essential for enzyme function. This arginine is clearly hyper-reactive and on this basis is very likely to be involved in substrate binding as has been shown for hyper-reactive arginine residues in other enzymes [15]. Much of the earlier work on arginine modification of enzymes with phenylglyoxal has assumed that the major initial reaction involved the formation of the 2:1 Takahashi adduct [17] and the extent of modification was usually estimated by monitoring the incorporation of radioactive phenylglyoxal and assuming this 2:1 stoichiometry. Our results show that the initial phenylglyoxal modification of the type II dehydroquinases does not involve formation of the 2:1 Takahashi adduct but instead a 1:1 adduct forms which is then very rapidly dehydrated (Fig. 3-3).

Our observations also confirm that the adducts formed between phenylglyoxal and proteins are sufficiently stable to be analysed by the standard conditions used for the electrospray mass spectrometry of proteins. A major difficult in locating the sites of phenylglyoxal modification is the poor stability of the phenylglyoxal modified peptides during purification, especially during reverse phase HPLC in the presence of 0.1% trifluoroacetic acid. Although, under the conditions we describe there is undoubtedly significant hydrolysis of the phenylglyoxal peptide adducts the high sensitivity of electrospray masspectrometry still permits the simple and rapid identification of the major sites of arginine modification in proteins.

3.1.5 References

[1] Haslam, E. (1993). Shikimic Acid, Metabolism and Metabolites. J. Wiley, Chichester.

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

[3] Bentley, R. (1990) *Critical Reviews in Biochemistry and Molecular Biology* **25**, 307-384.

[4] Harris, J., Kleanthous, C., Coggins, J.R., Hawkins, A.R. & Abell, C. (1993) *J. Chem . Soc. Chem. Comm.* 1080-1081.

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

[6] Gourley, D.G., Coggins, J.R., Isaacs, N.W., Moore, J.D., Charles, I.G. & Hawkins, A.R. (1994) *J. Mol. Biol.* **241**, 488-91.

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

[8] Shneier, A., Kleanthous, C., Deka, R., Coggins, J.R. & Abell, C. (1991) *J. Am. Chem. Soc.* **113**, 9416-9418.

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

[10] Hawkins, A. R, Lamb, H. K., Moore, J. D., Charles, I. G.& Robert , C.F. (1993) *J. Gen. Microbiol.* **139**, 2891-2899.

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

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

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

[14] Riordan, J.F. (1979) *Molecular & Cellular Biochemistry* **26**, 71-91.

[15] Patthy, L. & Thesz, J. (1980) *Eur. J. Biochem.* **105**, 387-393.

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

[17] Lundblad, R.L. & Noyes, C.M. (1984) *Chemical Reagents for Protein Modification*, Volume II, pp 1-45.

[18] White, P.J., Young, G., Krell, T., Hunter, I.S. & Coggins, J.R. Manuscript in preparation

[19] Beri,R.K., Grant, S., Roberts, C.F., Smith, M. & Hawkins, A.R. (1990) *Biochem. J.* **265**, 337-42.

[20] Gill, C.S. & von Hippel, P.H. (1989) *Anal. Biochemistry* **64**, 2090-2095.

[21] Ferrige, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A. & Skilling, J. (1992) *Rapid Commun. Mass Spectrom.* **6**, 707-711.

3.2 Localisation of the active site of type II dehydroquinases. Identification of a common arginine-containing motif in the two classes of dehydroquinases.

Tino Krell, Malcolm J. Horsburgh, Alan Cooper, Sharon M. Kelly and John R. Coggins (1996) *Journal of Biological Chemistry* **271**, 24492-24497.

Localisation of the active site of type II dehydroquinases. Identification of a common arginine-containing motif in the two classes of dehydroquinases.

3.2.1 Summary

A novel method based on electrospray mass spectrometry (Krell, T., Pitt, A. R. and Coggins, J. R. (1995) *FEBS Lett.* **360**, 93-96) has been used to localise active site residues in the type I and type II dehydroquinases. Both enzymes have essential hyper-reactive arginine residues and the type II enzymes have an essential tyrosine residue. The essential hyper-reactive R23 of the *Streptomyces coelicolor* type II enzyme has been replaced by lysine, glutamine and alanine residues. The mutant enzymes were purified and shown by c.d. spectroscopy to be structurally similar to the wild-type enzyme. All three mutant enzymes were much less active, for example the k_{cat} of the R23A mutant was 30 000-fold decreased. The mutants all had reduced K_m values, indicating stronger substrate binding, which was confirmed by isothermal titration calorimetry experiments. A role for R23 in the stabilisation of a carbanion intermediate is proposed. Comparison of the amino acid sequence around the hyper-reactive arginine residues of the two classes of enzymes indicates that there is a conserved structural motif which might reflect a common substrate binding fold at the active centre of these two classes of enzyme.

3.2.2 Introduction

It is generally believed that enzymes have evolved to catalyse reactions by optimal mechanisms. With the exception of the four classes of proteinases (1) and the two classes of aldolases (2) examples of mechanistically different pairs of enzymes that catalyse the same reaction are very rare. This unusual situation is found with the two classes of dehydroquinase (EC 4.2.1.10) which catalyse the dehydration of 3-dehydroquinate to form 3-dehydroshikimate (3). The reaction is common to two metabolic pathways: the biosynthetic shikimate pathway which is used for the synthesis of aromatic compounds in plants and micro-organisms (4) and the catabolic quinate pathway which enables fungi and some other micro-organism to use quinate as a carbon and energy source (5). The type I dehydroquinases catalyse a *cis* elimination and are only involved in the biosynthesis of shikimate (6) whereas the type II enzymes, which catalyse a *trans* elimination (1), have been found to have either a biosynthetic (7) or a catabolic role (8) and in at least one species a dual role (9).

Besides being mechanistically distinct the two classes of DHQ have very different biophysical properties and are apparently unrelated at the level of primary structure (10,11). The type I enzymes are dimers with a M_r of about 46 000 , they are heat labile and use a mechanism which involves the formation of a Schiff-base intermediate between the substrate and a lysine residue of the enzyme followed by the abstraction of a proton by a general base (12,13,14,15). In the case of the type I *E. coli* enzyme the lysine residue has been located (K170) by trapping the Schiff-base intermediate by borohydride reduction (12,13) and this residue is conserved in all type I sequences (12) as is His-143 (15,16) which is the general base. Both Met-23 and Met-205 (17) have also been identified as active site residues although their role in substrate binding or catalysis has not been established.

In contrast the type II enzymes are heat stable dodecamers with a subunit molecular weight of about 16 000 Da (11), they catalyse a *trans* elimination of the elements of water (3), there are no conserved lysine residues and they are resistant to inhibition by borohydride treatment (C. Kleanthous, R. K. Deka and T. Krell, unpublished results). Clearly the type II enzymes do not use the Schiff-base mechanism and it has been suggested that the two classes of DHQs are the result of convergent or parallel evolution (18).

Both the type I and type II enzymes have been crystallised (19,20) and to aid structural analysis we have been using group-specific chemical modification to localise the active sites. Recently we have described a new method, based on electrospray mass spectrometry for monitoring the modification of proteins by the arginine-directed reagent phenylglyoxal (PGO) (21). This method allows the direct measurement of the amounts of individually modified enzyme species. PGO rapidly inactivates the type II DHQs (21). In the early stages of inactivation only one modified species could be detected which correlated directly with the activity loss. The single site of modification was identified by HPLC-electrospray mass spectrometry based peptide mapping to be the hyper-reactive residue R23. Here we report the use of this methodology to identify further active site residues in the DHQs. Inactivation with TNM has identified an essential tyrosine in the type II enzymes and inactivation with PGO has identified a hyper-reactive arginine in the type I enzymes. The essential role of the active site arginine in the type II enzymes has been confirmed by site-directed mutagenesis and a number of mutant enzymes characterised kinetically, by circular dichroism (c.d.) and by isothermal titration calorimetry (ITC). This has led to the identification of a structural motif containing these essential amino acids which is common to both the type I and II DHQs.

3.2.3 Experimental Procedures

Purification and Assay of Type I and Type II DHQs-The type II DHQs from *S. coelicolor* (22) and *Aspergillus nidulans* (23) were overexpressed in *E. coli* and assayed and purified as described previously (24,25). Protein concentrations were determined spectrophotometrically at 280 nm using extinction coefficients calculated from the amino acid compositions (26).

The *E. coli* type I DHQ was purified from an overproducing strain according to the procedure of Chaudhuri *et al.* (6). Enzyme activity was determined by monitoring the formation of 3-dehydroshikimate at 234 nm (ε = 12 x 10^3 $M^{-1}cm^{-1}$) at 25°C. The assay mixture for the type I enzyme contained 100 μM ammonium dehydroquinate as substrate in 100 mM potassium phosphate (pH 7.0); for the type II enzyme of *A. nidulans* the assay mixture contained 1 mM ammonium dehydroquinate in 50 mM Tris/acetate (pH 7.0). The assay for the type II enzyme of *S. coelicolor* was carried out in 50 mM Tris/acetate (pH 8.0) containing 2 mM substrate.

Inactivation with Phenylglyoxal (PGO)-Enzyme inactivation reactions were carried out at 25°C in a volume of 2 ml with continuous stirring. Samples of type I and type II DHQ (3 nM) in 100 mM sodium bicarbonate buffer, pH 9.4 were pre-incubated for 5 min at 25°C and then PGO (freshly made up 50 mM stock solution in water) was added to a final concentration of 0.5-4.0 mM. Aliquots were removed at various times for enzyme assay.

Inactivation with Tetranitromethane (TNM)-Type II DHQ was pre-incubated in 0.1 M Tris/HCl, pH 8.0 for 5 min at 25°C and then TNM (freshly made up stock solution of 15 mM in 95% ethanol) was added to a final concentration of 5-50 μM. Aliquots were removed at various times for enzyme assay.

Preparation of inactivated Enzyme Samples for ES-MS-Enzyme inactivated to different extents was prepared for mass spectrometry by stopping the reaction by gel filtration on a Sephadex G50 column (200 mm x 17 mm, flow rate 15 ml/h) eluted with 0.5% (w/v) ammonium bicarbonate. To remove low molecular mass contaminants prior to mass spectrometry the enzyme samples were washed twice by diluting 50-fold with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, UK).

Proteolytic Digests-Modified and native DHQ were digested at 37°C. The enzyme in 0.4 M ammonium bicarbonate was denatured in 8 M urea for 1 hour. Water (37°C) was added to lower the urea concentration to 2 M and the mixture was incubated with 3% (w/v) trypsin (1 g/l stock solution in 1 mM HCl) for 4 hours. Afterwards, another 1% (w/v) trypsin was added and incubation continued for one hour. Digestion was stopped by freezing the samples at -80°C.

Monitoring the Amount and localising the Site of Modification using ES-MS-Mass spectrometry was performed on a VG Platform quadrupole mass spectrometer fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG MassLynx software (VG Biotech Ltd., Altrincham. Cheshire, UK). Carrier solvent [1:1 (v/v) acetonitrile/water, 0.2% formic acid] infusion was controlled at 10 µl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA). Protein samples were dissolved in carrier solvent at a concentration of 20 pmol/µl, centrifuged at 5000 x g for 2 min and then 10-20 µl samples injected directly into the carrier stream. MaxEnt deconvolution (27) was applied for quantitative analysis of the raw data using a 1.0 Da peak width and 1.0 Da/channel resolution.

The protein digests were separated by HPLC on a C-4 reverse phase column (2.0 x 150 mm; Delta-pac^{tm}, Waters, Waterford, Hertfordshire, UK) using 2% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid as the initial solvent (flow rate 0.25 ml/min); after an 8 min wash to remove GdnHCl a linear gradient of 2-70% acetonitrile (v/v) in 0.1% (v/v) trifluoracetic acid was applied to elute the peptides. The column eluate was introduced directly into the mass spectrometer with a drying gas flow of 400 l/h and the source temperature was set at 100°C. The absorption profile of the eluted peptides was recorded at 214 nm and centroid mass spectra in the range 400-1800 Da/z were recorded at 4 s intervals.

Site-directed Mutagenesis-The plasmid containing the coding sequence of *S. coelicolor* DHQ cloned into the *Sma*I site of pIBI25 (IBI/Eastman Kodak Co., New Haven, CT) (pDHQ) was a gift from Professor Ian S. Hunter, Strathclyde University, Glasgow. Three site-directed mutants (R23K, R23Q, R23A) were prepared using PCR site-directed mutagenesis (28). Primers were synthesised on an Applied Biosystems Model 280A DNA synthesiser. The three reverse PCR primers which contained the mismatch codon (highlighted in bold) overlap the *Bgl*II site which is located 10 bp downstream of arginine residue 23:

5' GTAGATCTCCGGCTG**CTT**CTGGCCGAGC 3'(R23K)

5' GTAGATCTCCGGCTG**CGC**CTGGCCGAGC 3'(R23A)

5' GTAGATCTCCGGCTG**CTG**CTGGCCGAGC 3' (R23Q).

The forward-primer was complementary to a region of the polylinker 190 bp upstream and contained an *Eco*RV site : 5' CCAGATATCAAATTAATACGACTCACTATAGG 3'.

The PCR products were cloned back into pDHQ replacing the wild-type *Bgl*II-*Eco*RV fragment. The presence of only the desired mutations was confirmed by DNA sequencing on both strands (29). For protein expression the entire mutated sequences were cloned into the *NdeI*/*Hind*III site of the expression-plasmid pT7-7 (30) and overexpressed in *E. coli* BL21(DE3)plysS (31). The

mutant enzyme was purified by following the same protocol as the wild-type enzyme. The expression strain used has a native level of type I DHQ. The background activity could be detected as a distinct peak in the eluate of the Phenylsepharose column (second column in enzyme purification) and separated completely from mutant enzyme. As a control a fraction of the pure mutant enzyme was incubated at 75°C for 20 min (type II native and mutant enzymes are heat stable, type I enzyme is not). The enzyme activity before and after the heat treatment was identical. Otherwise, the mutant enzymes were purified and assayed as described above for the native enzyme. Pure mutant enzyme was subjected to ES-MS analysis.

Circular Dichroism-The circular dichroism (c.d.) spectra of each protein were recorded in a Jasco J-600 spectropolarimeter. Spectra in the far u.v. (260-190 nm) were recorded in cylindrical quartz cells of pathlength 0.02 cm and spectra in the near u.v. region (320-260 nm) were recorded in cells of pathlength 0.5 cm. The protein concentration was typically 0.5-0.6 mg/ml for far u.v. work and 5-6 mg/ml for the near u.v. experiments. All protein solutions were dialysed against 10 mM Tris/HCl, pH 7.5. The secondary structure content was determined from the data in the far u.v. over the range 190-260 nm using the CONTIN procedure (32).

Isothermal Titration Calorimetry (ITC)-Binding of equilibrated substrate/product mixtures to native and mutant DHQ samples at 25°C was determined using a Microcal OMEGA isothermal titration calorimeter following standard ITC procedures (33,34). Protein samples (ca. 1 mg/ml), exhaustively dialysed against 10 mM Tris/HCl, pH 7.6, were degassed briefly before loading into the ITC cell (ca. 1.4 ml). Substrate (ammonium dehydroquinate, 50 mM) was dissolved in dialysis buffer together with a trace of native DHQ and incubated at room temperature for 15 hours to give the equilibrated substrate/product mixture prior to loading into the ITC injection syringe. [Note: use of pure substrate in these experiments would result in large heat effects from the enzymatically-catalysed reaction, rather the heats of binding required here.] A typical titration experiment involved a series of up to twentyfive 10 μl injections of ligand (substrate/product mix) into the enzyme solution at 3 min intervals, with continuous stirring. Integrated heat effects, after correction for dilution and mixing controls determined separately under identical conditions, were analysed by standard techniques using Microcal ORIGIN software assuming 1:1 enzyme-complex formation (32).

3.2.4 Results

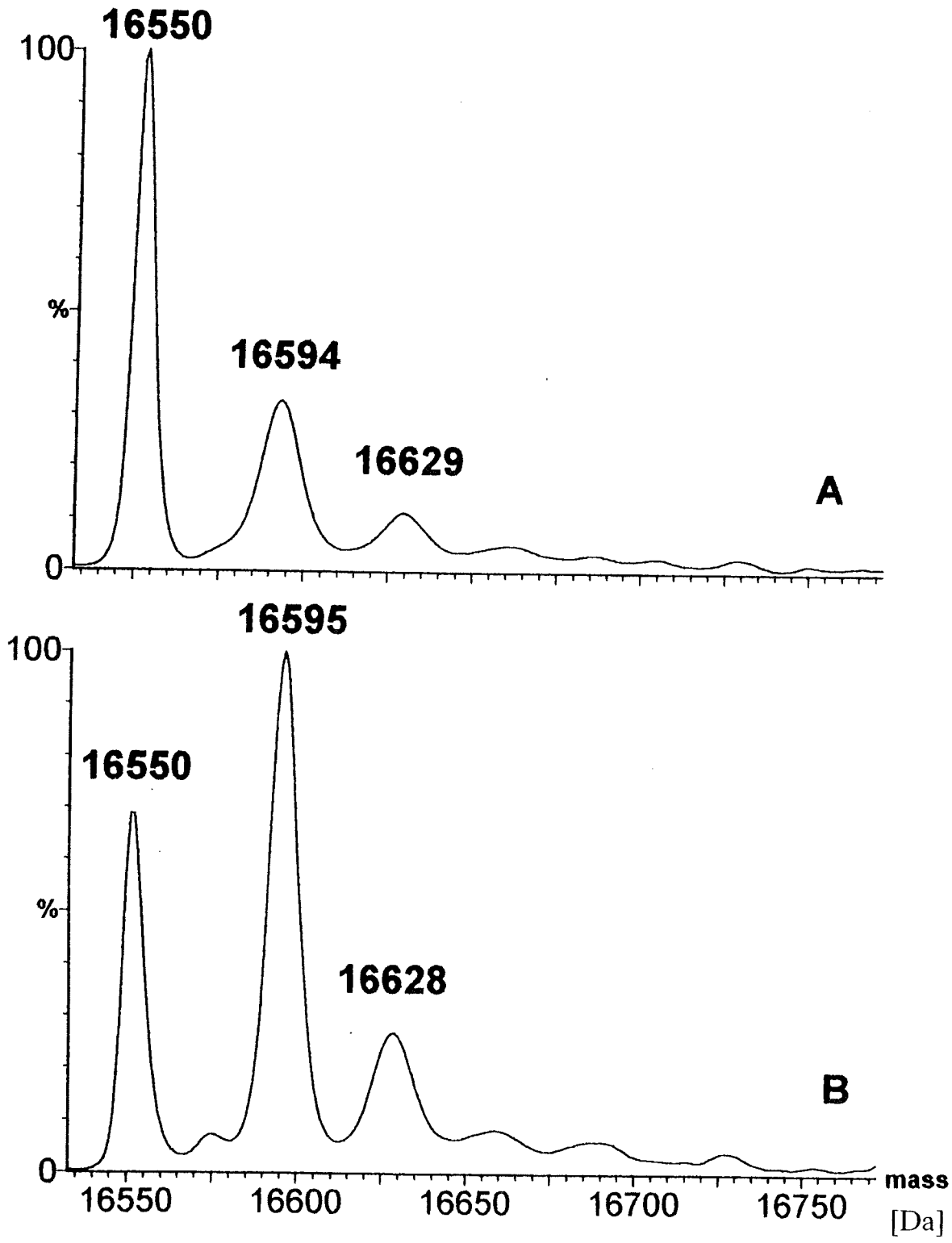
Identification of an essential Tyrosine Residue (Y28) in the type II DHQ of *S. coelicolor* DHQ-*S. coelicolor* DHQ can be inactivated by treatment with the nitrating reagent tetranitromethane (TNM). This reagent is reasonably specific for tyrosine residues (35). Inactivation followed pseudo-first order kinetics (data not shown). The nitration of the aromatic ring of a tyrosine residue results in a mass increase of +45 Da. Enzyme inactivated to different extents showed only one modified tyrosine residue in the mass spectrometer (Fig. 3-4, exemplified by 30% and 70% inactivated species). A minor peak which corresponds to an oxidised species could also be detected. To localise the site of modification a 30 % inactivated sample was digested with trypsin. A peptide with the mass of 2096.3 Da could be detected which was not present in the tryptic digest of the native enzyme (data not shown). This mass corresponds to peptide 24-42 +45 Da. The sequence contains only one tyrosine residue (Y28) which is conserved in all the type II DHQs (Fig. 3-5).

Identification of a hyper-reactive Arginine (R213) in the Type I DHQ of *E. coli*-The type I DHQs, like the type II enzymes, can be inactivated by PGO; in the case of the *E. coli* enzyme the inactivation follows pseudo-first order kinetics (data not shown). 20% inactivated enzyme showed only one modified species with a mass difference of +116 Da which corresponds to the incorporation of one PGO per reacting site. Native and 20% inactivated enzyme were digested with trypsin, peptides separated by HPLC and injected directly into the electrospray mass spectrometer. Trypsin cleaves at unmodified arginine and lysine residues; no cleavage is expected after a modified arginine residue. Recorded spectra were therefore scanned for the theoretical tryptic peptides containing a single internal PGO-modified arginine residue (+116 Da). This tryptic digestion allows the precise location of the site of reaction. Three modified peptides were detected. From the peak size on the HPLC trace it was apparent that two of these modified peptides were only minor species whereas the third modified peptide was much more abundant. This abundant arginine-containing peptide corresponds to residues 208-229 (Fig. 3-6, A), and identifies R213 as the hyper-reactive arginine residue (Fig. 3-5). Peptides 208-213 and 214-229 are shown to demonstrate that the protease cuts the unmodified enzyme species at R213 (Fig. 3-6, B and C). The double peak in Figure 3-6, C is due to the occurrence in the tryptic digest of a second peptide (26-31) which has exactly the same mass as peptide 208-213; these peptides have different mobilities on reverse phase HPLC.

Figure 3-4 MaxEnt deconvolution of electrospray mass spectra of *S. coelicolor* DHQ.

(native molecular weight: 16 550 Da) inactivated with tetranitromethane (TNM).

The nitration of a single tyrosine residue results in a mass species of 16 595 Da (+ 45 Da), the minor species with the mass of 16 628 Da is likely to be due to oxidised enzyme containing methionine sulfone; A, 30% inactivated enzyme; B, 70% inactivated enzyme.


16550
100
%
16594
16629
A
0
16595
100
16550
%
16628
B
0
16550
16600
16650
16700
16750
mass
[Da]

The secondary sites of modification are at residues 38 and 48. The tryptic digest of native enzyme did not contain species with masses corresponding to the PGO-modified peptides.

Site-directed Mutagenesis of the hyper-reactive Arginine Residue in the Type II DHQ of *S. coelicolor* -The hyper-reactive arginine (R23) of *S. coelicolor* DHQ (21) was replaced by a lysine, glutamine and alanine residue. The masses of purified mutant proteins were determined using ES-MS. All three masses were found to be very close to the theoretical values.

Far u.v. c.d.. spectroscopy was employed to estimate the influence of the amino acid exchange on the secondary structure of the enzyme (Fig. 3-7A, Table 3-2). The R23K and R23A substitutions caused no significant structural perturbation compared with the wild-type. The R23Q mutation resulted in a small decrease in the percentage of β-sheet (Table 3-2), but this conclusion should be treated with caution in view of the noise in the spectrum below 195 nm. The near u.v. c.d. spectra (Fig. 3-7, B) also confirm that the tertiary structures of the mutant enzymes are very similar to the wild-type, although there are some small differences in the 270 to 285 nm region suggesting a subtle change in the environment of one or more tyrosine side chains.

Table 3-2 Estimation of secondary structure elements of native and mutant *S. coelicolor* DHQ

[using the CONTIN procedure (32)].

	α-Helix	β-sheet	remainder
	%	%	%
wild-type	21±1.1	41±1.1	39±2.0
R23K	20±0.8	40±0.9	40±1.6
R23Q	21±0.9	35±1.1	44±1.8
R23A	22±1.2	41±1.3	36±2.3

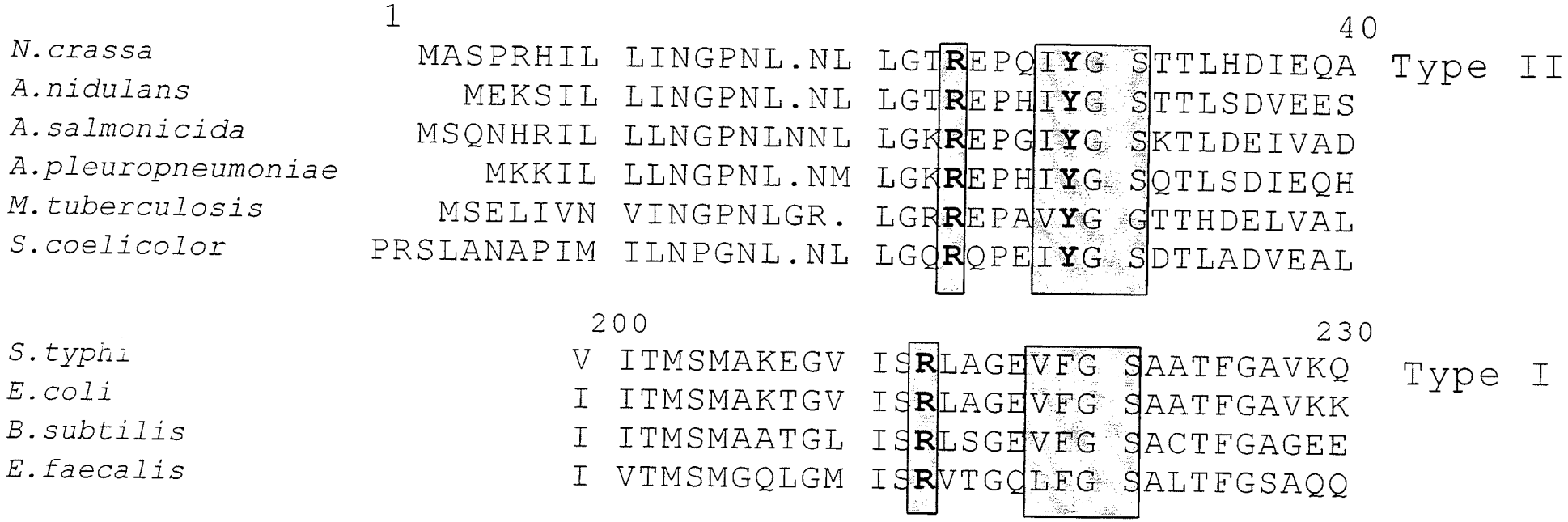


Figure 3-5 Sequence alignment of currently available type II and monofunctional type I DHQs.

Boxed amino acids form the common motif. Highlighted are the hyper-reactive arginine residues and the essential tyrosine.

Figure 3-6 Reverse phase chromatography/electrospray mass spectrometry (RP/ES-MS) data of a tryptic digest of *E. coli* DHQ.

20% inactivated after phenylglyoxal treatment. The enzyme sample is a mixture of 80% native enzyme and 20% single-site modified enzyme, centroid spectra were collected every 4 sec.; traces A, B and C show the relative abundance of mass species in all the recorded spectra against retention time; A, scan within all the recorded spectra for a mass of modified peptide 208-229 + 116 Da (1 PGO attached); B, scan for a mass of unmodified peptide 214-229; C, scan for the mass of unmodified peptide 208-213 (the double peak in this trace is due peptide 26-31 which has exactly the same mass as peptide 208-213; these two peptides have different HPLC-mobilities); D, simple HPLC trace at 215 nm against retention time (min).

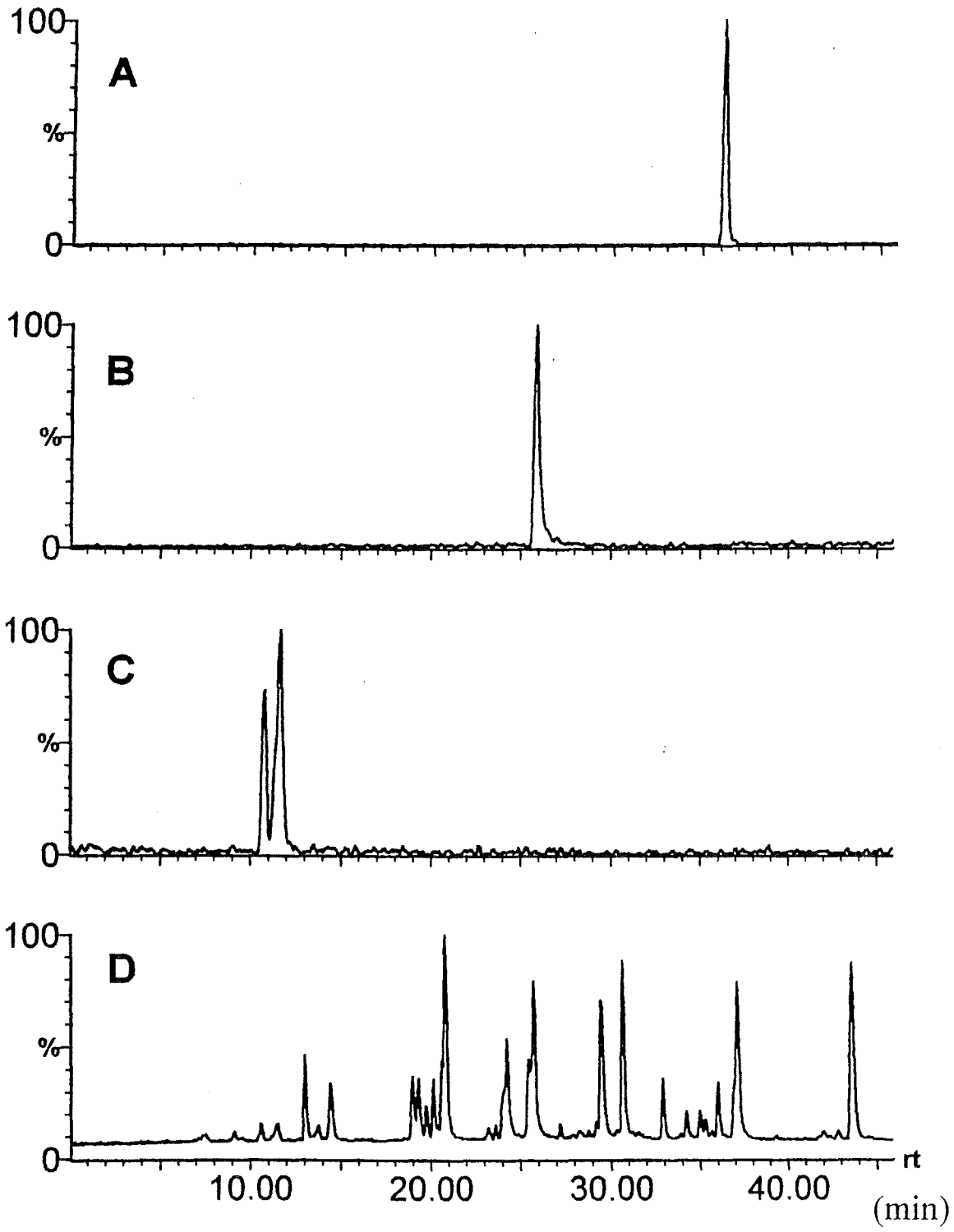
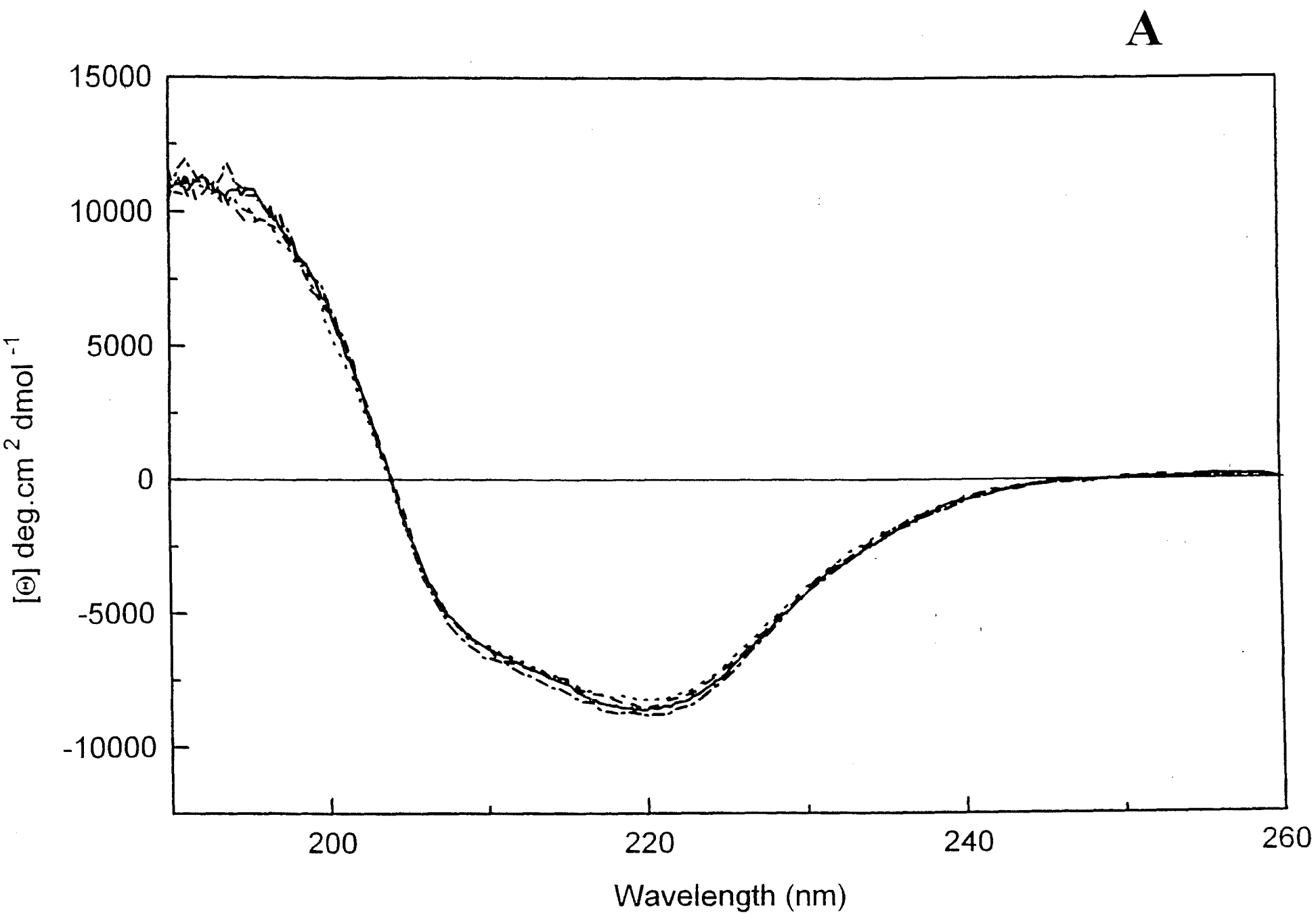

100
A
%
0
100
B
%
0
100
C
%
0
100
D
%
0
10.00
20.00
30.00
40.00
rt
(min)

Figure 3-7 Superimposed circular dichroism spectra of native and mutant *S. coelicolor* DHQ.

(——) wild-type; (-----) R23A; (•••••) R23K; (– • – • –) R23Q; A, far u.v.; B, near u.v.


A
15000
10000
5000
0
-5000
-10000
[Θ] deg.cm 2 dmol -1
200
220
240
260
Wavelength (nm)

B

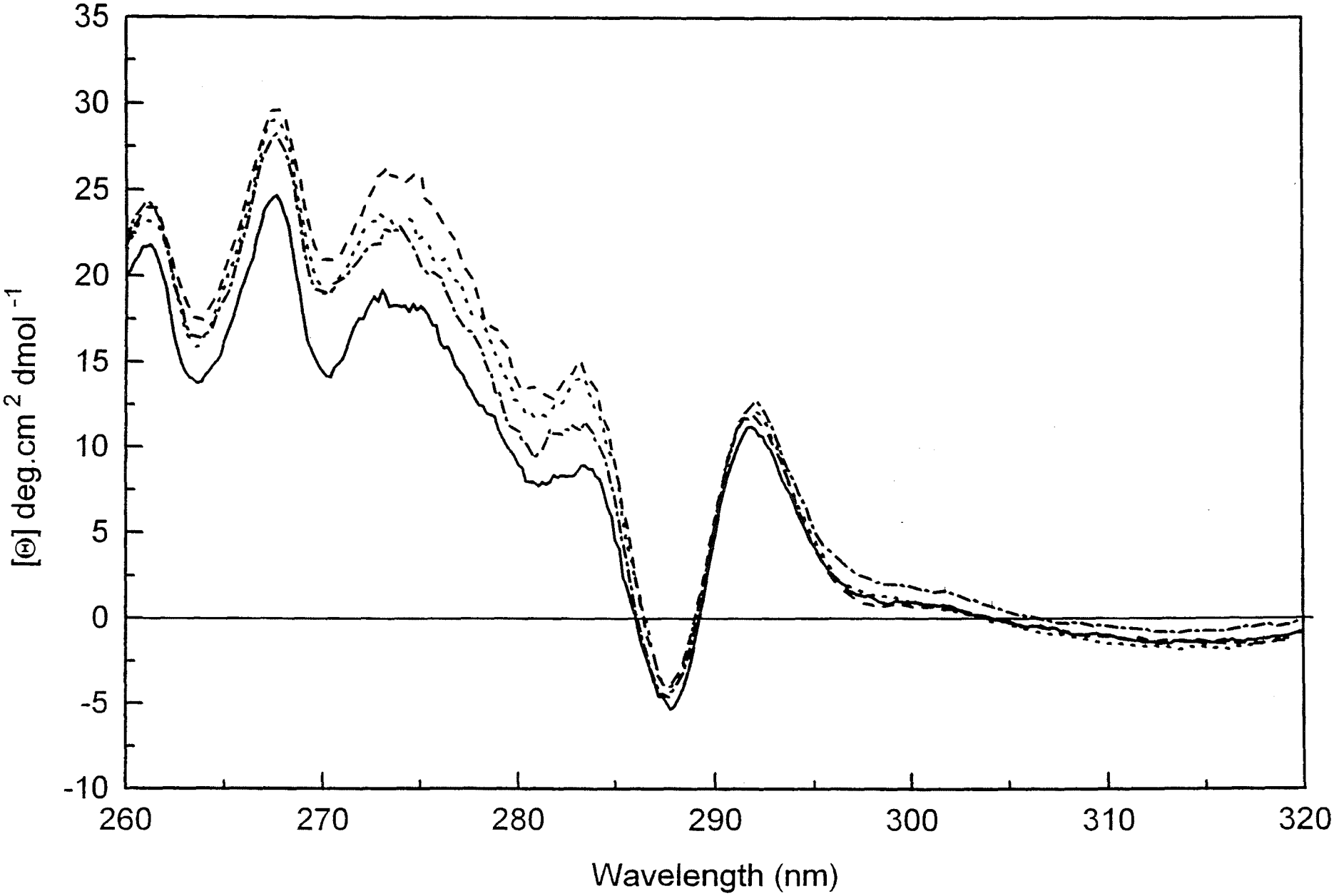

35
30
25
20
15
10
5
0
-5
-10
260
270
280
290
300
310
320
[Θ] deg.cm 2 dmol -1
Wavelength (nm)

Figure 3-8 Isothermal titration calorimetry curve - raw data

Raw ITC data for successive 10 μl injections of equilibrated substrate/product mixture (50 mM) into native or mutant DHQ (40-80 μM). The dilution control shows data for identical injections of substrate/product mixture into buffer. Traces are off-set for clarity. Negative-going peaks indicate exothermic binding.

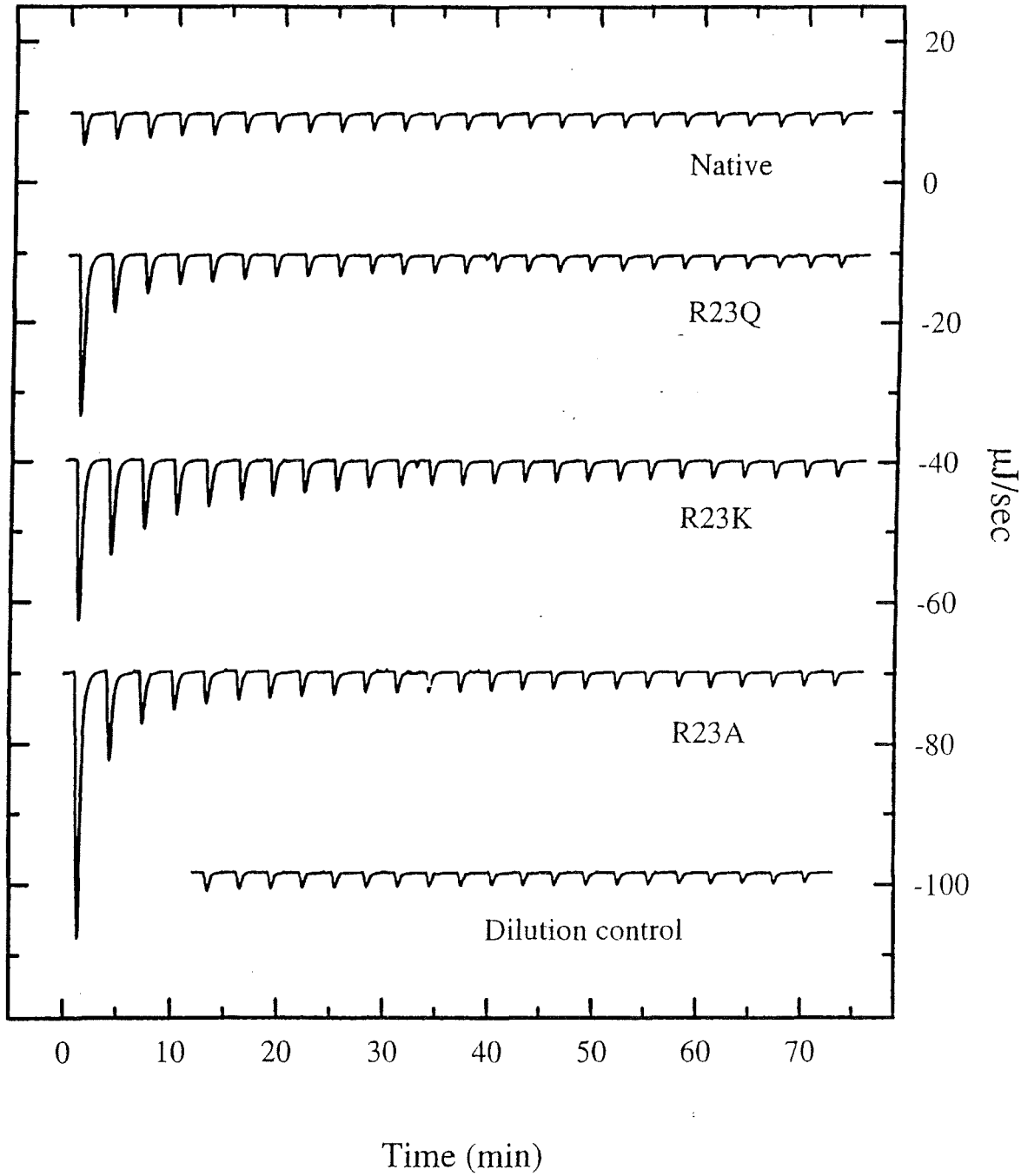
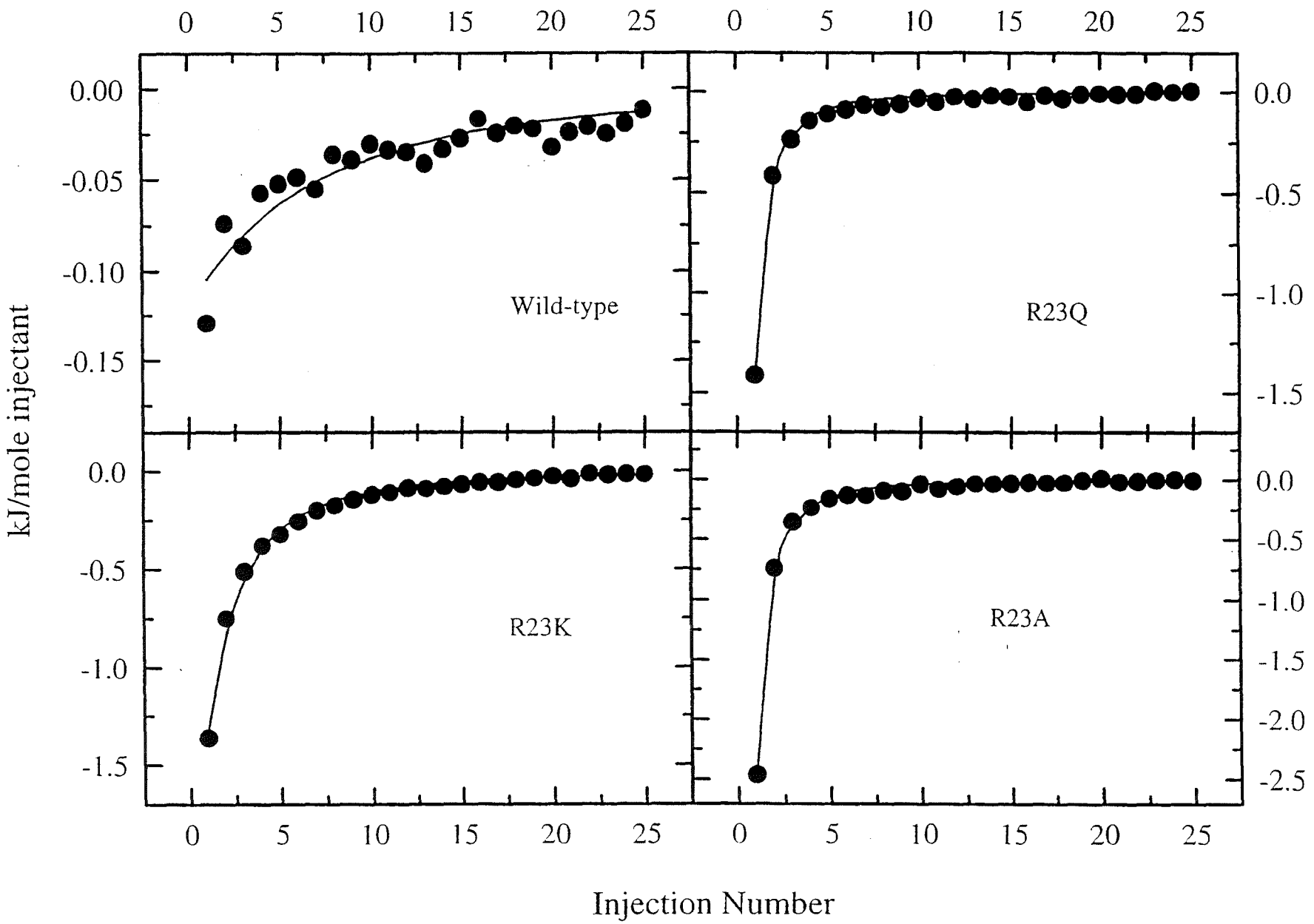

20
0
-20
-40
-60
-80
-100
μJ/sec
Native
R23Q
R23K
R23A
Dilution control
0
10
20
30
40
50
60
70
Time (min)

Figure 3-9 Isothermal titration calorimetry curve.

Comparison of integrated heat of binding data for native and mutant DHQ after correction for dilution controls. Lines are theoretical curves for 1:1 binding with parameters given in Table 3-4. Note the expanded scale for the wild-type enzyme.


Wild-type
R23Q
R23K
R23A
kJ/mole injectant
Injection Number
0
5
10
15
20
25
0.00
-0.05
-0.10
-0.15
0.0
-0.5
-1.0
-1.5
-2.0
-2.5

The kinetic parameters were determined as shown in Table 3-3. The K_m values were calculated using a Lineweaver-Burk plot and the mutant enzyme activity was determined as described above for the wild-type enzyme. Surprisingly, the K_m values of the mutant enzymes were found to be lower than those of the wild-type enzyme.

Table 3-3 The kinetic parameters of native and mutant *S. coelicolor* DHQ.

Assays were performed in triplicate with the values within 10% for R23Q and R23K and within 20% for R23A.

	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$s^{-1}M^{-1}$
wild-type	960	1100	8.7×10^5
R23K	0.35	250	1400
R23Q	0.31	135	2300
R23A	0.032	170	188

This apparent increased affinity for substrate was confirmed by isothermal titration experiments of enzyme species with an equilibrated substrate/product mixture as illustrated in Fig. 3-8; 3-9 and Table 3-4.

Table 3-4 Apparent DHQ-substrate binding parameters of wild-type and mutant *S. coelicolor* DHQ.

determined from calorimetric titrations assuming 1:1 complex formation.

enzyme	K_{app}	ΔH_{app}
	mM	$kJ\ mol^{-1}$
wild-type	5.40	-11
R23K	1.10	-25
R23Q	0.33	-24
R23A	0.30	-29

With wild-type enzyme the heat effects were relatively small but, after correction for dilution heat effects, consistent with weak exothermic 1:1 enzyme-substrate complex formation

(K_{app}≈5 mM, see Table 3-4). In marked contrast, addition of substrate to mutant proteins gave significantly more exothermic initial heat effects, which decreased rapidly with subsequent injections in a manner consistent with much tighter binding (K_{app}≈0.3-1 mM; Fig. 3-9 and Table 3-4).

3.2.5 Discussion

Both chemical modification and site-directed mutagenesis of the conserved hyper-reactive arginine residue in the type II DHQs (R23 in the *S. coelicolor* enzyme) lead to loss of enzyme activity. Arginine residues with hyper-reactivity towards α-1,2 dicarbonyls such as PGO are often found to be involved in carboxylate binding. Such residues are located in special microenvironments and have lower pK_a's than other arginine residues (36,37). Three mutants were made of the hyper-reactive R23 of *S. coelicolor* DHQ: R23K, R23Q, and R23A.

Mass spectrometry and c.d. analysis revealed that the mutants were of the expected size and had the same or very similar secondary structure as the wild-type enzyme (Fig. 3-7; Table 3-2). All three mutants showed a much smaller turnover number (k_{cat}), with R23K being in the same range as R23Q and with R23A still being another order of magnitude less active (Table 3-3). In contrast a decrease in K_m values was observed, by a factor of 4 for the R23K mutant and by factors of 8 and 6 for the R23Q and R23A proteins respectively (Table 3-3). This behaviour was also confirmed more directly by calorimetric binding studies, which showed that binding of substrate was significantly enhanced by replacement of R23 with K, Q or A, as indicated by a 5 to 10-fold decrease in the apparent binding constant (K_{app}) and significantly more exothermic binding. Direct comparison of K_{app} with K_m values is not strictly appropriate because K_m's relate to initial substrate binding, rather than the equilibrated substrate/product mixture used for calorimetric experiments. They are, in any case, only related to substrate binding affinities within the validity of the Michaelis-Menten approximation. Nevertheless, K_m and K_{app} are of the same order of magnitude and, more importantly, show the same trends in the mutant enzyme.

In other enzymes, where arginine residues are known to be involved in the recognition of substrates containing carboxyl groups, arginine mutation leads to a large increase in K_m (38,39). Our results therefore suggest that R23 is not simply involved in substrate recognition but must have a catalytic role. One possibility is that a positive charge is required to stabilise a negative transition state such as an enolate. Although the R23K mutant retains the positive charge the

position of this charge will be significantly altered in the mutant. The full interpretation of this result must await the determination of the 3D structure of the enzyme.

Chemical modification has also identified Y28 as a residue in or near the active site. Both R23 and Y28 are conserved in all type II DHQs (Fig. 3-5). The small changes observed in the near u.v. c.d. spectra (Fig 3-7) of the arginine mutant proteins can be explained by a change in the environment of a tyrosine residue. The proximity in the sequence of R23 and Y28 and the occurrence of these small spectral changes when the arginine is altered are both consistent with these residues being spatially close.

Chemical modification experiments on the type I DHQs have also identified a hyper-reactive arginine (R213 in the *E. coli* enzyme) which appears to be essential for activity and is near the methionine residue previously identified as an active site residue (17). Comparison of the amino acid sequences around the hyper-reactive arginine residues of the two classes of enzyme (Fig. 3-5) indicates that there is a conserved structural motif, extending over 9 residues, shown by the shaded region in Fig. 3-5. In the type I enzymes the essential tyrosine residue identified by TNM modification in the type II enzymes is replaced by a phenylalanine residue. Both the mechanistic experiments (3,13) and the sequence comparisons (11,40) have suggested that these two classes of enzyme are likely to be structurally very different and it will therefore be interesting to see whether this sequence similarity is reflected in a common substrate binding fold at the active centres of these two classes of enzyme.

3.2.6 Acknowledgements

The financial support of BBSRC and the University of Glasgow is acknowledged. The Biological Microcalorimetry Facility in Glasgow and the Biological Circular Dichroism Facility are supported by BBSRC and EPSRC. We would like to thank Dr. V. Math for the synthesis of oligonucleotides, Prof. I. Hunter for the kind gift of pDHQ, Mr. A. Elwell for computer support, Dr. A. R. Pitt for help with mass spectrometry and Prof. N. C. Price and Dr. C. Abell for helpful discussions.

3.2.7 References

1. Page, M. I., and Williams, A. (1987) *Enzyme Mechanisms*, The Royal Society of Chemistry, London.

2. Littlechild, J. A., and Watson, H C (1993) *TIBS* **18**, 36-39.

3. Harris, J., Kleanthous, C., Coggins, J. R., Hawkins, A. R., and Abell, C. (1993) *J. Chem. Soc. Chem. Commun.* 1080-1081.

4. Haslam, (1974) *The Shikimate Pathway*. J. Wiley, New York.

5. Giles, N. H., Case, M. E., Baum, J., Geever, R., Huiet, L., Patel, V., and Tyler, B. (1985) *Microbiol. Reviews* **49**, 338-358.

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

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

8. Hawkins, A. H., Giles, N. H., and Kinghorn, J. R. (1982) *Biochem. Genet.* **20**, 271-286.

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

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

11. Kleanthous, C., Deka, R., Davies, 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.

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

13. Shneier, A., Kleanthous, C., Deka, R., Coggins, J. R., and Abell, C. (1991) *J. Amer. Chem. Soc.* **113**, 9416-9418.

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

15. Leech, A.., James, R., Coggins, J., and Kleanthous, C. (1995) *J. Biol. Chem.* **270**, 25827-25836.

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

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

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

19. 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.

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

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

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

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

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

25. White, P. J., Young, G., Krell, T., Hunter, I. S., and Coggins, J. R. manuscript in preparation.

26. Gill, C. S., and von Hippel, P. H. (1989) *Anal. Biochem.* **64**, 2090-2095.

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

28. Higuchi, R., Krummel, B., and Saki, R. K. (1988) *Nucleic Acids Res.* **16**, 7351-7367.

29. Sanger, F., Coulson, A. R., Barrell, B, G., Smith, A. H., and Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161-178.

30. Gilbert, W. (1991), unpublished , obtained from VecBase 3.0, copyright IBI.

31. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol.* **189**, 113-130.

32. Provencher, S. W., and Glöckner, J. (1981) *Biochemistry* **20**, 33-37.

33. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L-N. (1989) *Anal. Biochem.* **179**, 131-137.

34. Cooper, A., and Johnson, C. M. (1994) *Methods in Molecular Biology, Vol. 22: Microscopy, Optical Spectroscopy, and Macroscopic Techniques*, Ed. C. Jones, B. Mulloy and A. H. Thomas, Hunana Press, Toronto, N.J.

35. Sokolovsky, M., Riordan, J. F., and Vallee, B. L. (1966) *Biochemistry* **5**, 3582 3589.

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

37. Riordan, J. F. (1979) *Mol. Cell. Biochem.* **6**, 71-92.

38. Taguchi, H., and Ohta, T. (1994) *J. Biochem.* **115**, 930-936.

39. Van Berkel, W., Westphal, A., Eschrich, K., Eppink, M., and De Kok, A. (1992) *Eur. J. Biochem.* **210**, 411-419.

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

4. Kinetic studies using the electrospray mass spectrometer: Measuring the rate of dephosphorylation of phosphoglycerate mutases.

The use of mass spectrometry to examine the formation and hydrolysis of the phosphorylated form of phosphoglycerate mutase.

Jacqueline Nairn, Tino Krell, John R. Coggins, Andrew R. Pitt, Linda A. Fothergill-Gilmore, Rebecca Walter and Nicholas C. Price (1995) *FEBS letters* **359**, 192-194.

Phosphoglycerate mutase from *Schizosaccharomyces pombe*; development of an expression system and characterisation of three histidine mutants of the enzyme.

Jacqueline Nairn, Nicholas C. Price, Sharon M. Kelly, Daniel Rigden, Linda A. Fothergill-Gilmore and Tino Krell (1996) *Biochimica et Biophysica Acta* **1296**, 69-75.

4.1 The use of mass spectrometry to examine the formation and hydrolysis of the phosphorylated form of phosphoglycerate mutase.

4.1.1 Abstract

Electrospray mass spectrometry has been used to study the formation and hydrolysis of the phopshorylated forms of two phosphoglycerate mutases. The half-life of the enzyme from *Saccharomyces cerevisiae* was 35 min at 20°C in 10 mM ammonium bicarbonate, pH 8.0. Addition of 1 mM 2-phosphoglycollate reduced this value by at least 100-fold. The phosphorylated form of the enzyme from *Schizosaccharomyces pombe* was much less stable with a half life of less than 1 min. The results are discussed in terms of the kinetic properties of the enzymes. Mass spectrometry would appear to be a powerful method to study the formation and breakdown of phosphorylated proteins, processes which are of widespread significance in regulatory mechanisms.

4.1.2 Introduction

The advent of electrospray mass spectrometry with its ability to measure molecular masses with a precision of $\pm$0.01% has made it much easier to detect and characterise both post-translational and chemical modifications of proteins [1-3]. The introduction of the phospho group ($-OPO_3^{2-}$) in place of -H would lead to a mass increase of 78 units and thus be readily detectable. This approach has been used, for instance, in the delineation of the sites of phosphorylation in glycogen synthase [4] after separation of the modified peptides. However, examination of an intact phosphorylated protein by mass spectrometry does not appear to have been widely studied. In this paper we describe the use of electrospray mass spectrometry to monitor the formation of the phosphorylated forms of two phosphoglycerate mutases (PGAMs) and to examine the stability of these phosphorylated enzymes towards hydrolysis.

The catalytic cycle of phosphoglycerate mutase is thought to proceed via an enzyme-substitution pathway involving the enzyme phosphorylated on a histidine side chain (His-8 in the case of *S. cerevisiae* PGAM) [5]. The phosphorylated PGAM is slightly unstable towards hydrolysis leading to a low level of phosphatase activity (approx. 0.002% that of the mutase activity in the case of the *S. cerevisiae* enzyme). It has been suggested [5] that the flexible C terminal segment (14 amino acids) of this enzyme may be important in preventing access of water to the active site and thus maintain a high level of mutase to phosphatase activity. Recently

the monomeric PGAM from the fission yeast *Schizosaccharomyces pombe* has been shown to lack this flexible C-terminal tail sequence [6]. The kinetic properties of the *S. pombe* enzyme have not yet been explored in detail.

4.1.3 Experimental

PGAM from an overexpressing strain of *S. cerevisiae* was isolated as described previously [7], with the addition of a final FPLC Superose-12 gel-filtration step. The concentration of the enzyme was determined spectrophotometrically assuming a value of 1.45 for the A_{280} of a 1 mg/ml solution [8].

PGAM from *S. pombe* was produced using the PGK-based vector pMA91 [9] for the high level expression of recombinant *GPMSP* in a transformed null mutant strain of *S. cerevisiae* (S150-gpm::HIS3) [7]. The overexpressed *S. pombe* PGAM was purified in a similar manner to over-expressed *S. cerevisiae* PGAM [7]; routinely 10-15 mg of enzyme of at least 95% purity on SDS-PAGE [10] could be obtained per litre of cells. Full details of the expression system and enzyme purification will be published elsewhere (Nairn, Fothergill-Gilmore and Price, in preparation).

The concentrations of the *S. pombe* enzyme were determined by a Coomassie Blue binding method [11] using bovine serum albumin as a standard or spectrophotometrically using a value for the A_{280} (1.40 for a 1 mg/ml solution) calculated from the aromatic amino acid content of the enzyme [6,12]; the values agreed to within 5%. The assays of mutase, phosphatase and synthase activities were performed as described previously [7].

The phosphorylated forms of the PGAMs from *S. cerevisiae* and *S. pombe* were prepared by mixing the enzymes with 2,3-bisphosphoglycerate (BPG) in 10 mM Tris/HCl, pH 8.0, followed by rapid gel filtration on NAP 5 columns (Pharmacia) equilibrated with 10 mM ammonium bicarbonate, pH 8.0, to remove free mono- and bis-phosphoglycerates. This procedure allowed samples to be studied by mass spectrometry within 3 min of the mixing.

Mass spectrometry was performed on a VG Platform quadrupole mass spectrometer (2 - 3000 amu range) fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG MassLynx software (VG Biotech. Ltd, Altrincham, Cheshire, U.K.). Carrier solvent (1:1 (v : v) acetonitrile : water) infusion was controlled at 10 ml/min using a Harvard syringe pump (Harvard Apparatus, South Natic, Mass., U.S.A.). Capillary voltages were between 2.8 and 3.2 kV, extraction cone voltages 20 - 30 V and the focusing cone voltage offset by + 10 V. The source temperature was set at 65°C, the nebulising gas flow at 10 l/h and the

drying gas flow at 250 l/h. Lens stack voltages were adjusted to give maximum ion currents. The M_r range 700 - 1500, which contained > 95% of the signal intensity for both unmodified and phosphorylated forms of PGAM, was scanned with a sweep time of 5 s. The instrument was calibrated over this M_r range immediately before use with horse heart myoglobin (Sigma). Samples for analysis were diluted with an equal volume of 4% (v : v) formic acid in acetonitrile and 10 - 20 ml aliquots injected directly into the carrier stream. The MaxEnt deconvolution procedure [13] was applied for quantitative analysis of the raw data using 1.0 Da peak width and 1 Da/channel resolution.

4.1.4 Results and Discussion

***S. cerevisiae* PGAM**-The specific activities of the *S. cerevisiae* enzyme in the mutase and the phosphatase assays (970 and 0.020 mmol/min/mg respectively) and the effect of 1 mM substrate analogue 2-phosphoglycollate on the latter activity (18 fold stimulation) were very similar to those described previously [7].

The mass spectrum of the *S. cerevisiae* enzyme shows a single peak of M_r 27478.9 $\pm$ 1.0 consistent with the subunit M_r (27 477) calculated from the cDNA-derived sequence of the enzyme [14]. After addition of 0.8 molar equivalents (expressed per active site) of BPG, followed by rapid gel filtration, the mass spectrum clearly shows the formation of the (mono)phosphorylated enzyme, with a mass increase of 79 amu (Fig. 4-1, A).

Under these conditions, 60% of the total enzyme was present in the phosphorylated form. There was little or no ($\leq$0.05 molar equivalents per active site) BPG or monophosphoglycerates either free or enzyme bound after the gel filtration. By increasing the molar ratio of BPG to 10-fold more than 95% of the enzyme could be isolated in the phosphorylated form (data not shown). When the enzyme which had been prepared by reaction with the sub-stoichiometric amount of BPG and then gel filtered was subsequently incubated at 20^{o}C, there was a slow loss of the phospho group from the enzyme. Fig. 4-1, B shows the mass spectrum of the sample taken after 18.5 min incubation. The data from three independent experiments expressed as a semi-logarithmic plot are shown in Fig. 4-2; in each case the proportion of the enzyme in the phosphorylated form is expressed relative to the initial proportion as 100%. From the semi-logarithmic plot the rate constant for the hydrolysis of the phosphorylated enzyme is 0.02 min^{-1}, corresponding to a half-life of approximately 35 min. This direct estimate half-life of the phosphorylated form of *S. cerevisiae* PGAM is rather longer than the value (1-2 min) quoted by Britton *et al.* [15] on the basis of unpublished work.

Figure 4-1 Formation and hydrolysis of the phosphorylated form of *S. cerevisiae* PGAM monitored by mass spectrometry.

The spectra over the M_r range shown (27440 - 27680) were obtained by applying the MaxEnt deconvolution procedure to the raw data. In each case the size of the major peak is set as 100%. The peaks labelled E and E-P represent the dephosphorylated and phosphorylated forms of the enzyme respectively. (a) Mass spectrum of the sample immediately after gel filtration. (b) Mass spectrum recorded after a further 18.5 min incubation.

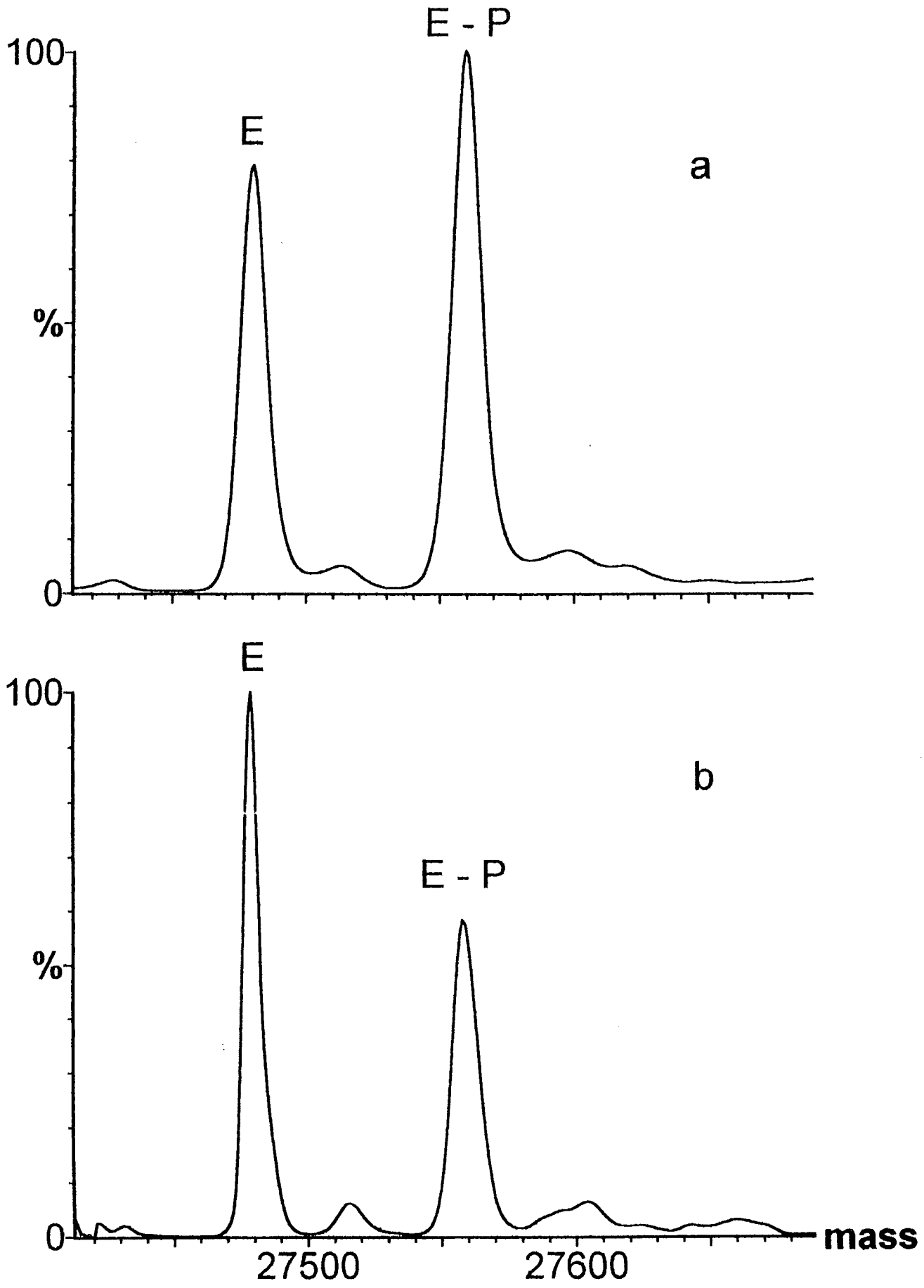
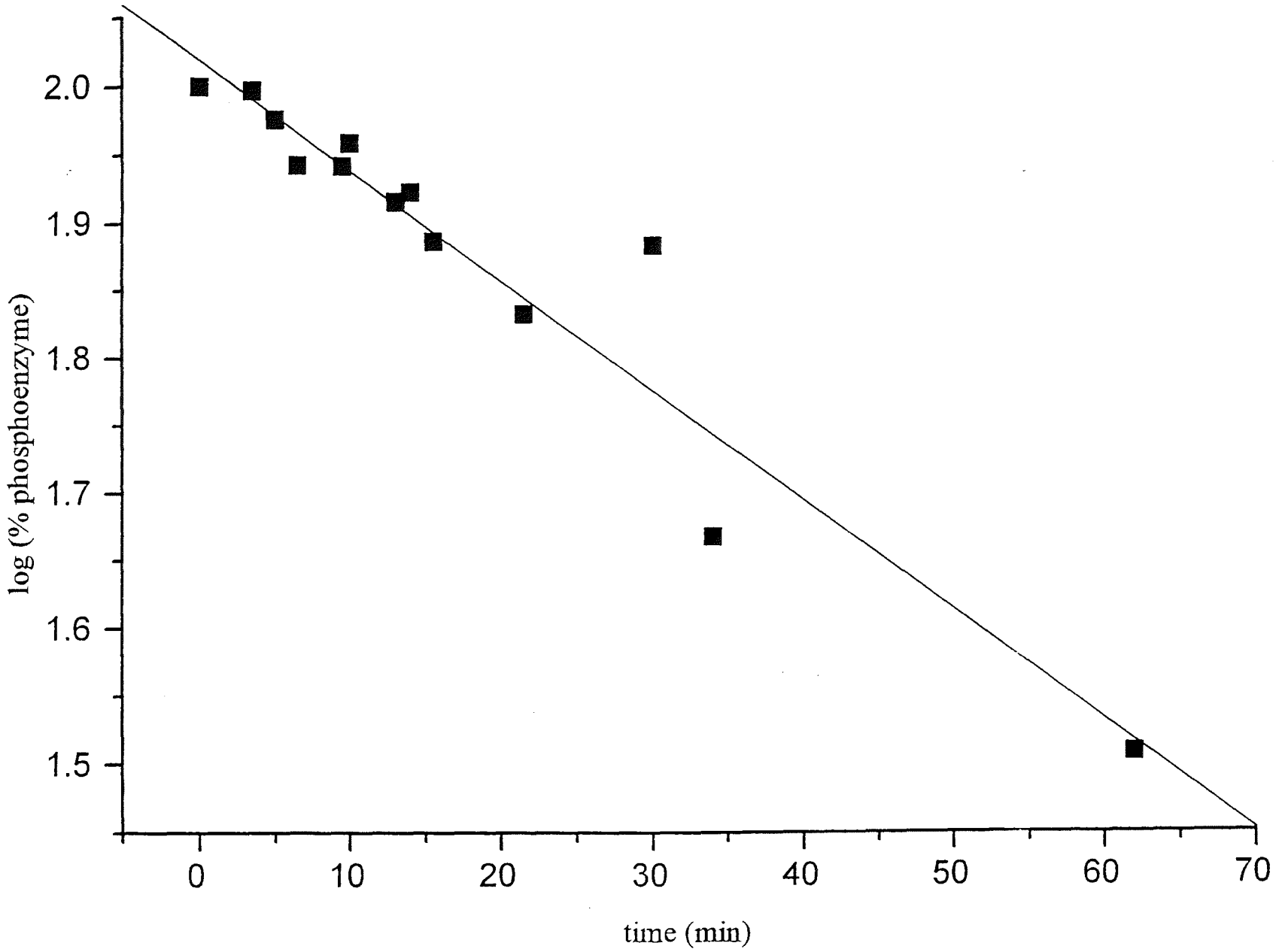

E - P
100
E
a
%
0
E
100
b
E - P
%
0
mass
27500
27600

Figure 4-2 Semi-logarithmic plot to show the conversion of the phosphorylated form to the dephosphorylated form of the enzyme.

Data from three independent experiments are plotted; in each case the initial proportion of phosphoenzyme (immediately after gel filtration) is scaled to 100%.


2.0
1.9
1.8
1.7
1.6
1.5
log (% phosphoenzyme)
0
10
20
30
40
50
60
70
time (min)

It is however clear from preliminary work that the measured phosphatase activity is markedly influenced by factors such as ionic strength and the presence of phosphorylated substrates and analogues, and this may well account for at least some of the observed differences in stability of the phosphorylated enzyme, i.e. the phosphorylated enzyme would appear to be much less stable when the enzyme is turning over.

When the mass spectrometry experiment was repeated with 1 mM 2-phosphoglycollate added to the phosphorylated enzyme immediately after gel filtration, it was found that within 1 min, less than 5% of the enzyme remained in the phosphorylated form (data not shown). This result indicates that in the presence of 2-phosphoglycollate the half-life of the phosphorylated enzyme is less than 20 s (i.e. the rate of the dephosphorylation reaction is accelerated at least 100-fold in the presence of this substrate analogue; a somewhat greater effect than the 18-fold stimulation of phosphatase activity).

In the presence of acetonitrile (50% (v : v)), the phosphorylated enzyme showed no detectable breakdown after 70 min incubation (data not shown). Since this concentration of acetonitrile leads to a considerable loss of secondary structure (as shown by far u.v. c.d. measurements) it can be concluded that the denatured phosphorylated enzyme is considerably more stable towards hydrolysis than is the native phosphorylated enzyme. This conclusion would be consistent with earlier work in which negligible breakdown had been shown to occur over 120 min in the presence of 1.5% (w/v) SDS [16].

***S. pombe* PGAM**-In the mutase assays, the specific activity of *S. pombe* PGAM (215 mmol/min/mg) is about 20% of that of the *S. cerevisiae* enzyme, whereas in the phosphatase assay it is some 2.5-fold higher (0.06 mmol/min/mg). Thus the ratio of the phosphatase/mutase assays is 12-fold higher for the *S. pombe* enzyme. The degree of stimulation of the phosphatase activity by 2-phosphoglycollate is considerably lower in the case of the *S. pombe* enzyme (4.2-fold).

The mass spectrum of *S. pombe* PGAM shows a peak at with an M_r of 23 679.4 $\pm$ 1.5, corresponding to that calculated from the published sequence [6] assuming that the initiating Met has been removed and the N-terminal threonine acetylated. On addition of 0.8 molar equivalents of BPG followed by gel filtration (a process lasting 3 min), less than 5% of the enzyme was present in the phosphorylated form (data not shown). This was not due to an inability to form the phosphorylated enzyme since the mixture prior to gel filtration showed that 60% of the enzyme was present in the phosphorylated form. From these results, it could be concluded that the half-life of the phosphorylated form of *S. pombe* PGAM was less than 1 min (at least 35-fold shorter

than that of the *S. cerevisiae* enzyme). This greater instability of the phosphorylated form of the *S. pombe* enzyme is in qualitative agreement with the higher ratio of phosphatase to mutase activities for this enzyme.

In conclusion, mass spectrometry should prove to be a very useful technique for monitoring the phosphorylation and dephosphorylation of a number of enzymes and proteins, a process which has been recognised to be a key regulatory mechanism for a large number of crucial biological processes [17]. The mass spectrometric technique avoids the necessity of using radioactive isotopes. Further refinements would include increasing the time resolution, allowing the rates of faster processes to be monitored accurately.

4.1.5 Acknowledgements

We wish to thank the Welcome Trust and the B.B.S.R.C. for financial support and for the provision of a studentship (RW). We are grateful to Dr. M.F. White for preparing the S150-gpm::HIS3 strain of *S. cerevisiae* and to Dr. S. M. Kingsman for supplying the plasmid pMA91.

4.1.6 References

[1] Feng, R. and Konishi, Y. (1992) *Anal. Chem.* **64**, 2090-2095.

[2] 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.

[3] Akashi, S., Niitsu, U., Yuji, R., Ide, H. and Hirayama, K. (1993) *Biol. Mass Spectrom.* **22**, 124-132.

[4] Gibson, B.W. and Cohen, P. (1990) *Methods Enzymol.* **193**, 480-501.

[5] Fothergill-Gilmore, L.A. and Watson, H.C. (1989) *Adv. Enzymol. Relat. Areas Mol. Biol.* **62**, 227-313.

[6] Nairn J., Price, N.C., Fothergill-Gilmore, L.A., Walker, G.E., Fothergill, J.E. and Dunbar, B. (1994) *Biochem. J.* **297**, 603-608.

[7] White, M.F. and Fothergill-Gilmore, L.A. (1992) *Eur. J. Biochem.* **207**, 709-714.

[8] Edelhoch, H., Rodwell, V.W. and Grisolía, S. (1957) *J. Biol. Chem.* **228**, 891-903.

[9] Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983) *Gene* **24**, 1-14.

[10] Laemmli, U.K. (1970) *Nature* (London) **227**, 680-685.

[11] Sedmak, J.J. and Grossberg, S.E. (1977) *Anal. Biochem.* **79**, 544-552.

[12] Gill, S.C. and von Hippel, P.H. (1989) *Anal. Biochem.* **182**, 319-326.

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

[14] White, M.F. and Fothergill-Gilmore, L.A. (1988) *FEBS Lett.* **229**, 383-387.

[15] Britton, H.G., Carreras, J. and Grisolía, S. (1972) *Biochemistry* **11**, 3008-3014.

[16] Han, C.-H. and Rose, Z.B. (1979) *J. Biol. Chem.* **254**, 8836-8840.

[17] Hardie, D.G. (1993) *Protein Phosphorylation: a Practical Approach*, IRL Press, Oxford.

4.2 Phosphoglycerate mutase from *Schizosaccharomyces pombe*; development of an expression system and characterisation of three histidine mutants of the enzyme.

Jacqueline Nairn, Nicholas C. Price, Sharon M. Kelly, Daniel Rigden, Linda A. Fothergill-Gilmore and Tino Krell (1996) *Biochimica et Biophysica Acta* **1296**, 69-75.

Phosphoglycerate mutase from *Schizosaccharomyces pombe*; development of an expression system and characterisation of three histidine mutants of the enzyme.

4.2.1 Summary

The small, monomeric, phosphoglycerate mutase (PGAM) from *Schizosaccharomyces pombe* has been overexpressed in a strain of *Saccharomyces cerevisiae* in which the gene encoding PGAM has been deleted, with a yield of purified enzyme of 10-15 mg per litre cell culture. Three mutants in which histidine residues in *S. pombe* PGAM have been substituted by glutamine have been purified and characterised. Two mutants (H151Q and H196Q) have kinetic and structural properties very similar to wild-type enzyme, consistent with the proposed location of these (non-conserved) histidines on the surface of the enzyme. The third mutant (H163Q) involving a histidine thought to be part of the active site has greatly reduced mutase and phosphatase activities. Mass spectrometry shows that the phosphorylated form of the H163Q is several hundred times more stable towards hydrolysis than the phosphorylated form of wild-type enzyme. The H163Q mutant appears to be structurally quite distinct from wild-type enzyme. 600 MHz 1D proton NMR spectra of good quality have been obtained for wild-type enzyme and the H151Q and H196Q mutants.

4.2.2 Introduction

Phosphoglycerate mutase (PGAM) (EC 5.4.2.1) catalyses the interconversion of 2-phosphoglycerate and 3-phosphoglycerate. There are broadly two classes of PGAM; those which are active in the absence of 2,3-bisphosphoglycerate (BPG) and those which depend on BPG for activity [1]. The mechanism of the latter group is known to involve the formation of a phosphoenzyme, in which a histidine side chain is phosphorylated. The most extensively studied of the BPG-dependent PGAMs is the tetrameric enzyme from *Saccharomyces cerevisiae*, whose amino acid sequence [2] and X-ray structure at 0.28 nm resolution [3] have been published. Although relatively little is known about the mechanism of the BPG-independent enzymes and no stable phosphorylated enzyme intermediate can be isolated, there has been some recent evidence from both chemical modification and site-directed mutagenesis [4] that up to three histidine residues may be crucial for activity.

The BPG-dependent PGAM from the fission yeast *Schizosaccharomyces pombe* has been shown to be monomeric with a molecular mass of 23 kDa [5]. The amino acid sequence of the *S. pombe* enzyme is 43% identical with that of the *S. cerevisiae* enzyme, with conservation of the histidine which is phosphorylated during the reaction cycle (His-8 in the *S. cerevisiae* enzyme; His-15 in the *S. pombe* enzyme) and most of the other residues thought to be involved in the active site. However, significant differences exist in regions known to be important in the interactions between the subunits of the *S. cerevisiae* enzyme. Thus in the "D" (dimer) interface region of the *S. cerevisiae* enzyme, a five residue loop which interacts with the adjacent subunit (Leu^{74}-Trp-Ile-Pro-Val^{78}) is replaced by Asn-Leu-Glu-Thr-Ile in the aligned *S. pombe* sequence. A 25 residue loop (Pro-122 to Pro-146) in the *S. cerevisiae* enzyme, part of which is involved in interactions across the "T" (tetramer) interface is deleted in the aligned *S. pombe* sequence. These changes may well account for the monomeric nature of the *S. pombe* enzyme.

The small size of *S. pombe* PGAM puts it within the size range which is accessible to high resolution NMR spectroscopy. Application of this technique in conjunction with a programme of site-directed mutagenesis should yield new, detailed insights into the structure and dynamics of the active site of PGAMs. In this paper we describe the development of an expression system for *S. pombe* PGAM, and report on site-directed mutagenesis experiments aimed at examining the roles of histidine side chains. We also present preliminary 1D proton NMR spectra of the wild-type and mutant enzymes which demonstrate that the full assignment of the spectrum should be a feasible proposition with suitable isotopic labelling strategies.

4.2.3 Experimental

PGAM from cultures of *S. pombe* (i.e. chromosomal PGAM) was isolated as described [6], with the addition of a second chromatography step on Cibacron Blue-Sepharose eluted with 1 M NaCl in order to remove any traces of cofactor or substrate which might interfere with subsequent assays. The NaCl was subsequently removed by dialysis against the appropriate buffer. The concentrations of the *S. pombe* enzyme were determined by a Coomassie Blue binding method [7] using bovine serum albumin as a standard. This method gave values within 5% of those determined spectrophotometrically using a value for the A_{280} calculated from the aromatic amino acid content of the enzyme [8]. (The value calculated (1.40 for a 1 mg/ml solution) refers to the enzyme in 6 M GdnHCl, but for most proteins the value in buffer is generally within 10% of this value [8]).

The PGK-based vector pMA91 [9] was used for the high level expression of recombinant *GPM*^{sp} in *S. cerevisiae*. To generate *GPM*^{sp} for subcloning, two oligonucleotides were designed: one oligonucleotide corresponded to the upstream coding region of *GPM*^{sp}, with a *Bgl*II site added on to the 5' end of this ("sense") oligonucleotide (5'-GCCGAAGATCTATGACTACCGAAGCT-3'). The other ("anti-sense") oligonucleotide corresponded to the downstream coding region of *GPM*^{sp}, with a *Bgl*II site added on to the 3' end of this oligonucleotide (5'-GCCGAAGATCTCTAGTTGTCAATGAG-3'). These oligonucleotides served as primers for the amplification of *GPM*^{sp} such that a *Bgl*II site was introduced to allow direct cloning into the expression site of pMA91. Using the cDNA clone of *GPM*^{sp} (*pbii.GPM*^{sp}) [5] as the template, the PCR conditions used were as follows: 3 mM $MgSO_4$, 350 mM dNTPs, 120 pmol of each primer and 2 units of Vent polymerase (New England Biolabs, Bishops Stortford, Herts., U.K.) per 100 ml reaction. After 30 cycles of denaturation (95°C, 1 min), annealing (55°C, 1 min) and extension (72°C, 1 min), the resultant fragment was isolated from an agarose gel, digested with *Bgl*II and cloned into pMA91. pMA91-GPM^{sp} was transformed into the strain of *S. cerevisiae* (S150-gpm::HIS3) from which the chromosomal copy of the gene encoding *S. cerevisiae* PGAM has been deleted [10]. Overexpressed *S. pombe* PGAM was purified by ion-exchange on DE-52 in an analogous fashion to that described for the *S. cerevisiae* enzyme [10] with the addition of a final FPLC gel-filtration step on Superose-12. The Superose-12 column was eluted with 50 mM sodium phosphate, pH 8.0.

Site-directed mutagenesis of PGAM from *S. pombe* used the method outlined by Chen and Przybyla [11], involving two rounds of PCR with the appropriate primers. (In the following description, P1 and P2 are respectively the 5'→3' and 3'→5' oligonucleotides used to introduce *Bgl*II sites in the construction of pMA91.GPM^{sp}). The first round of PCR was carried out using the following reaction conditions:- 1 x *Pfu* buffer (20 mM Tris-HCl, pH 8.2, 10 mM KCl, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 0.1% (v/v) Triton X-100, 10 mg/ml BSA), 200 mM dNTPs, 100 pmol P1, 100 pmol mutagenic primer M, 500 pg *pbii.GPM*^{sp} and 2.5 units *Pfu* polymerase in 100 ml reaction volume. 25 cycles were carried out (95°C, 1 min; 53°C, 1 min; 72°C, 1 min). The product generated from this first round of PCR (PCR-1) was purified using Wizard^{TM} PCR Preps DNA Purification System (Promega). The second round of PCR was also performed for 25 cycles under the following conditions (95°C, 1 min; an extended period (1 min 15 s) of cooling to 53°C; 53°C, 1 min; 72°C, 1 min). A 100 ml reaction mixture contained 1 x *Pfu* buffer, 200 mM dNTPs, 100 pmol P2, 100 pmol PCR-1, 0.5 mg *pbii.GPM*^{sp} and 2.5 units *Pfu* polymerase.

The product generated from the second round of PCR was digested with *Bgl*II and purified from a low melting point agarose gel. This purified product was then ligated into *Bgl*II-treated pMA91. To confirm the desired mutation, the nucleotide sequence of the entire GPM^{sp} coding region was determined using the dideoxy chain termination method [12] on alkaline denatured ds DNA. The mutagenic primers (altered sequences underlined) were designed using the published nucleotide sequence of the gene encoding *S. pombe* PGAM [5]:- (a) for the H151Q mutant, to replace H in the sequence IVPHILK by Q, the primer was 3'-TAACAGGGAGTTTAGGAATTC-5'; (b) for the H163Q mutant, to replace H in the sequence IAAHGNS by Q, the primer was 3'-TAACGGCGAGTTCCATTGAGA-5'; (c) for the H196Q mutant, to replace H in the sequence PIVYHLD by Q, the primer was 3'-GGGTAACAGATGGTTAACCTG-5'.

The assays of mutase and phosphatase activities were performed at 30°C as described [10]. Mutase activity was measured in a system consisting of:- 30 mM Tris-HCl at pH 7.0, 20 mM KCl, 5 mM $MgSO_4$, 0.2 mM ADP, 0.15 mM NADH, 10 mM 3-phosphoglycerate, 0.2 mM BPG, 0.08 units/ml enolase and 0.5 units/ml of each of pyruvate kinase and lactate dehydrogenase. Phosphatase activity was measured in a similar system, except that the Tris-HCl buffer was at pH 8.0 and 3-phosphoglycerate was omitted.

A model of the *S. pombe* PGAM was built based on the X-ray structure of *S. cerevisiae* PGAM [3] using standard procedures [13]. SYBYL (SYBYL 6.0 Manual (1994) Tripos Associates, St. Louis, MO) was used for all molecular modelling steps except final model refinement for which XPLOR [14] was used. Secondary structure prediction was by the HSSP [15] and SOPM [16] methods. The alignment of PGAM sequences published previously [5] was used. Firstly, the portions of the crystal structure corresponding to the C-terminal tail and the 25 residue deletion were removed. Then amino acid substitutions were made so that the new side chain occupied, as closely as possible, the same space as the original side chain [17]. If the new side chain was larger than the original, backbone-based side chain preferences [18] were used. The structure was inspected visually and minimally adjusted to relieve unfavourable steric contacts. Three new proline residues were introduced by the modelling. In two cases the marked backbone preferences of prolines and their preceding residues [19] combined with secondary structure predictions suggested that the existing backbone in the vicinity was incorrect. For these regions and for the 25 residue deletion, database loop searches [20] as implemented in SYBYL were used for backbone modelling. Loops were selected based on packing considerations, particularly exposure of charged residues and burial of hydrophobic residues, and for absence of

unusual backbone torsion angles. Model refinement was by a combination of Energy Minimisation (EM) and limited Molecular Dynamics (MD). Initial EM comprised 200 cycles in which only side chains were allowed to move, 200 cycles during which the structure in the vicinity of remodelled backbone was allowed to move and 200 cycles in which the whole structure was free to move. MD was in three similar stages with 5 ps simulation at 300K at each stage. A weak restraint was imposed on active site residues throughout the MD. Finally, a further 200 cycles of EM were carried out to produce the final model. The model was assessed and compared to the crystal structure using PROCHECK [21], PROFILE [22] and PROSA [23].

CD measurements were performed on a JASCO J-600 spectropolarimeter. The molar ellipticities were calculated assuming a value of 112 for the mean residue weight of each enzyme [5]. Fluorescence spectra were recorded on a Perkin-Elmer LS50 spectrofluorimeter.

Mass spectrometry analysis was performed on a VG Platform quadrupole mass spectrometer (2-3000 amu range) fitted with a pneumatically assisted electrospray source and controlled via the VG Mass-Lynx software (VG Biotech Ltd, Altrincham, Cheshire, UK) as described in detail [24]. The MaxEnt deconvolution procedure [25] was applied for quantitative analysis of the raw data using 1.0 Da peak width and 1 Da/channel resolution.

Proton NMR spectra of wild-type and mutant PGAMs were recorded (at protein concentrations in the range 6-8 mg/ml) both in H_2O and D_2O buffers (50 mM sodium phosphate) at 25°C, using the Varian 600 MHz instrument at the Department of Chemistry, University of Edinburgh. The pH (pD) of the samples was adjusted to be in the range 5.8-6.3, and the residual HOD signal (4.7 ppm) was used as a reference. ^1H spectra were acquired as a first increment of a phase-sensitive 2D NOESY experiment to achieve better suppression of residual water. The spectral width was 9000 Hz and the acquisition time 0.228 s. All spectra were zero filled to 8192 points prior to Fourier transformation. The resolution of the spectra was enhanced by Gaussian multiplication. The residual water signal of aqueous solution samples (90% H_2O, 10% D_2O) was removed by post-acquisition processing using the FELIX routine with a sine shape convolution function. The number of transitions for each sample was between 256 and 2048 depending on the sample concentration.

4.2.4 Results

Molecular modelling procedures -The model of *S. pombe* PGAM seems to be of comparable quality to the crystal structure on which it is based. One of the best measures of the quality of the structure is the percentage of residues in the core regions of the Ramachandran plot

[26]. The PROCHECK analysis shows that the model has 59% of its residues in these regions. While this Figure compares only moderately with the 60-80% expected from an X-ray structure at 0.28 nm resolution, it appears to be considerably better than the *S. cerevisiae* structure [3] (46% of residues in the core regions). This improvement would arise during refinements of a model based on a crystal structure of only moderate resolution. In addition, by all the other backbone and side chain parameters measured, the model performs better than would be expected for a crystal structure at 0.28 nm resolution. The PROFILE and PROSA results for the model are comparable to those for the starting structure. There are 4 histidine residues in the model. Active site residues His-15 and His-163 (corresponding to His-8 and His-181 respectively in the *S. cerevisiae* enzyme) are present in the same "clapping hands" orientation seen in the crystal structure. Residues His-151 and His-196 are not conserved and are both fully solvent-exposed on the surface of the model. This suggests that they should both be amenable to replacement by any hydrophilic residue.

Expression system for *S. pombe* PGAM-GPM^{sp} was amplified and subcloned into the *Bgl*II expression site of the high efficiency *S. cerevisiae* expression vector pMA91. The resulting construct pMA91.GPM^{sp} (Fig. 4-3) was transformed into the *S. cerevisiae* strain S150-gpm::HIS3. As described by White and Fothergill-Gilmore [10], S150-gpm::HIS3 lacks GPM^{sc} and consequently cannot utilise glucose as a carbon/energy source. Once transformed with pMA91.GPM^{sp}, this strain could be grown in complex YEPD medium indicating that the small monomeric *S. pombe* PGAM complements the *S. cerevisiae* strain that lacks all phosphoglycerate mutase coding sequence. The *S. pombe* PGAM was expressed in high yield in S150-gpm::HIS3 and the yield of enzyme of at least 95% purity on SDS-PAGE amounted to 10-15 mg per litre cell culture. Similar yields were obtained of the three mutant enzymes.

Kinetic properties of *S. pombe* PGAM-Table 4-1 shows the kinetic properties of chromosomal *S. pombe* PGAM as well as recombinant wild-type and mutant enzymes. It should be noted that in all cases the mutase activity was BPG dependent; in the absence of this cofactor, the activity was at least 50-fold lower. (The very small level of activity observed in the absence of the cofactor may result from traces of BPG contamination in the 3PGA). The recombinant and chromosomal *S. pombe* PGAMs are very similar in terms of their kinetic properties. In the mutase assay, the specific activity of *S. pombe* PGAM is about 20% of that of the *S. cerevisiae* enzyme (Table 4-1). However, the phosphatase activity of the *S. pombe* enzyme is about 6.5-fold higher than that of the *S. cerevisiae* enzyme. Thus the ratio of the phosphatase/mutase activities

is some 30-times higher for *S. pombe* PGAM than for the *S. cerevisiae* enzyme. In the case of the *S. pombe* enzyme the stimulation of the phosphatase activity by 2-phosphoglycollate is 2.1 fold.

Table 4-1 Kinetic properties of wild-type and mutant PGAMs.

The assays of mutase and phosphatase activities were performed as described previously [10]. Assays were performed in triplicate with the values within 5% for the mutase activities and within 10% for the phosphatase activities.

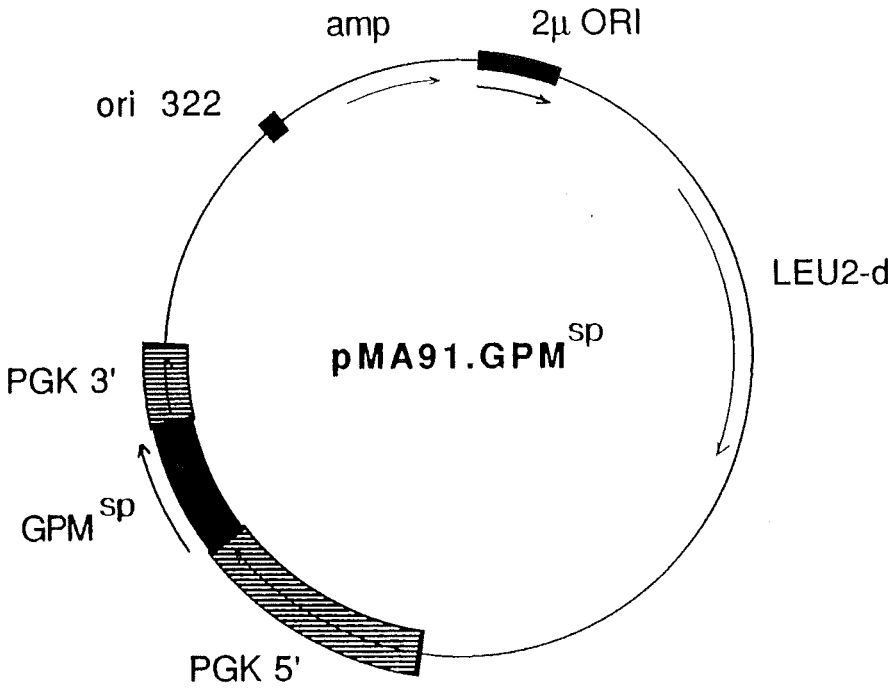
Enzyme	Mutase specific activity (mmol/min/mg)	Phosphatase specific activity (mmol/min/mg)	Phosphatase + 1 mM 2-PG (mmol/min/mg)
S. cerevisiae	970	0.02	0.35
S. pombe (chromosomal)	210	0.06	0.22
S. pombe (recombinant)	218	0.13	0.27
S. pombe (H151Q mutant)	221	0.14	0.55
S. pombe (H163Q mutant)	0.91	<0.003	<0.003
S. pombe (H196Q mutant)	238	0.14	0.68

As shown in Table 4-1, the H151Q and H196Q mutants have very similar kinetic properties to those of wild-type enzyme. This would be consistent with the proposal that these non-conserved histidines are located on the surface of the enzyme. The H163Q mutant has a substantially reduced mutase activity (some 220 fold lower than wild-type enzyme). This mutant had no detectable phosphatase activity (< 0.003 mmol/min/mg), which represents the lower limit of detection under the assay conditions employed). This is at least some 40 fold lower than that of wild-type enzyme. There was no effect of 1 mM 2-phosphoglycollate on this activity.

Mass spectral analysis of wild-type and mutant enzymes-The mass spectra of wild-type PGAM and the three mutants are shown in Fig. 4-4. The mass of the wild-type enzyme is consistent with that calculated (23,678 Da) from the amino acid sequence, assuming that the initiating Met has been removed and the N-terminal threonine acetylated [5,24]. The His→Gln mutation would be expected to lead to a decrease in mass of 9 Da; within experimental error, this is consistent with the observed masses of the mutants (Fig.4-4). All three mutants can apparently

Figure 4-3 The expression plasmid pMA91-GPM^{sp}.

The gene encoding PGAM from *S. pombe* is indicated by the solid bar. The hatched blocks indicate the *S. cerevisiae* PGK promoter (PGK 5') and the *S. cerevisiae* PGK termination sequence (PGK 3'). The *E. coli* origin of replication (ori 322) and the *S. cerevisiae* 2m origin of replication (2m ORI) are shown. The position and orientation of the defective LEU2-d gene and the b-lactamase gene (amp) are also shown.


amp
2μ ORI
ori 322
LEU2-d
pMA91.GPMsp
PGK 3'
GPMsp
PGK 5'

be phosphorylated by BPG, as shown by an increase in mass of 78 Da. While it is very likely that this phosphorylation takes place on His-15 (which corresponds to His-8, the known site of phosphorylation in the *S. cerevisiae* enzyme), it would be necessary to isolate and characterise the appropriate modified peptide to substantiate this conclusion. As previously reported [24], the $t_{1/2}$ of the phosphorylated form of the wild-type enzyme was found to be less than 1 min; this was also the case for the H151Q and H196Q mutants. However, in the case of the H163Q mutant, the phosphorylated form was found to be much more stable. As shown in Fig. 4-5, there was very little breakdown after 183 min, indicating a $t_{1/2}$ of at least 400 min for this process. There was no effect of 1 mM 2-phosphoglycollate on the stability of the phosphorylated form of the H163Q mutant (data not shown).

Spectral characterisation of wild-type and mutant enzymes:

CD spectra-The far UV CD spectra of wild-type *S. pombe* PGAM and the three mutants are shown in Fig. 4-6. It is clear that the H151Q and H196Q mutants have very similar secondary structures to wild-type enzyme, whereas the H163Q mutant has a significantly different structure. There was no detectable effect of BPG (25 mM) on the CD spectrum of the wild-type enzyme or the H163Q mutant.

Fluorescence spectra-When excited at 290 nm, the emission maxima of wild-type PGAM and the H151Q and H196Q mutants are at 339 nm. The emission maximum of the H163Q mutant is blue-shifted by some 2 nm, with the peak intensity being within 5% of that of the wild-type enzyme.

NMR spectra-The 1D proton NMR spectra for the wild-type and mutant PGAMs (recorded in 90% H_2O) are shown in Fig. 4-7. The wild-type enzyme yields a spectrum with chemical shift dispersion and line widths consistent with a folded protein of molecular mass 20-25 kDa. The spectra of the H196Q and H151Q mutants are broadly similar to that of the wild-type enzyme, although the somewhat broader line widths in the latter suggest that there may be an increased tendency to aggregate. When transferred into D_2O (99.9%), the wild-type enzyme showed a number of sharp peaks in the aromatic region which could arise from the histidine protons (data not shown). However, at the concentrations used it was not possible to observe correspondingly distinct peaks from the H196Q and H151Q mutants, thereby precluding any assignment of the peaks in the wild-type spectrum. The lack of dispersion and broad lines in the spectrum of the H163Q mutant (Fig. 4-7) are indicative of poorly folded and/or unstable protein.

Figure 4-4 Mass spectra of *S. pombe* PGAM.

Spectra were recorded on a VG platform quadrupole mass spectrometer as described in the text. Spectra of wild-type and mutant enzymes. Traces A, B, C and D correspond to H196Q, H151Q, H163Q and wild-type enzyme respectively. In each case the small peak to the right of the main peak represents the K^+ adduct of the protein.

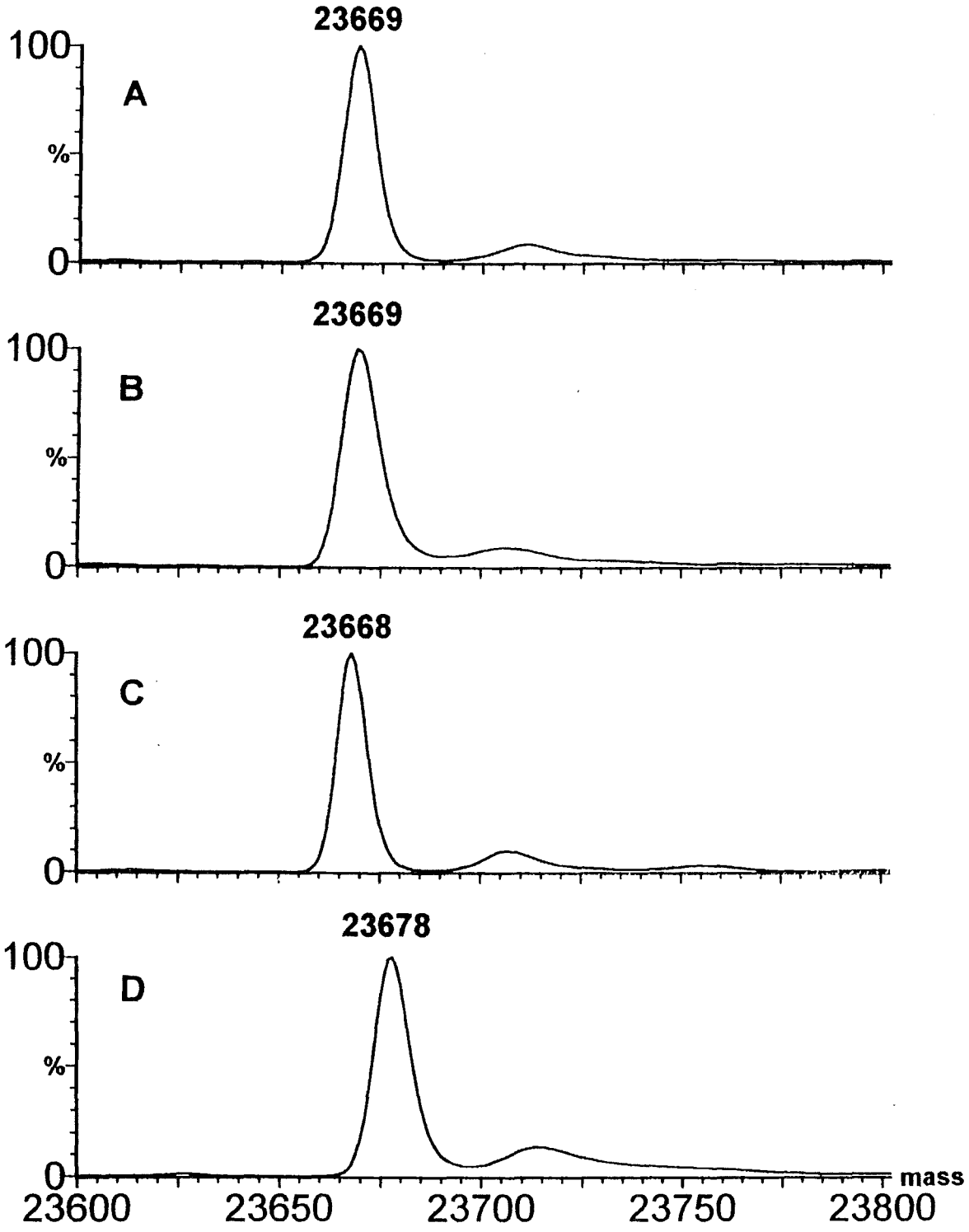

23669
100
A
%
0
23669
100
B
%
0
23668
100
C
%
0
23678
100
D
%
0
mass
23600
23650
23700
23750
23800

Figure 4-5 Spectra of the H163Q mutant enzyme.

Recorded immediately after addition of 7 molar equivalents (100 mM) BPG and subsequent rapid gel filtration to remove excess BPG [24] (A) and after a further 183 min incubation (B). The peaks corresponding to the phosphorylated (E-P) and non-phosphorylated (E) forms of the enzyme are indicated.

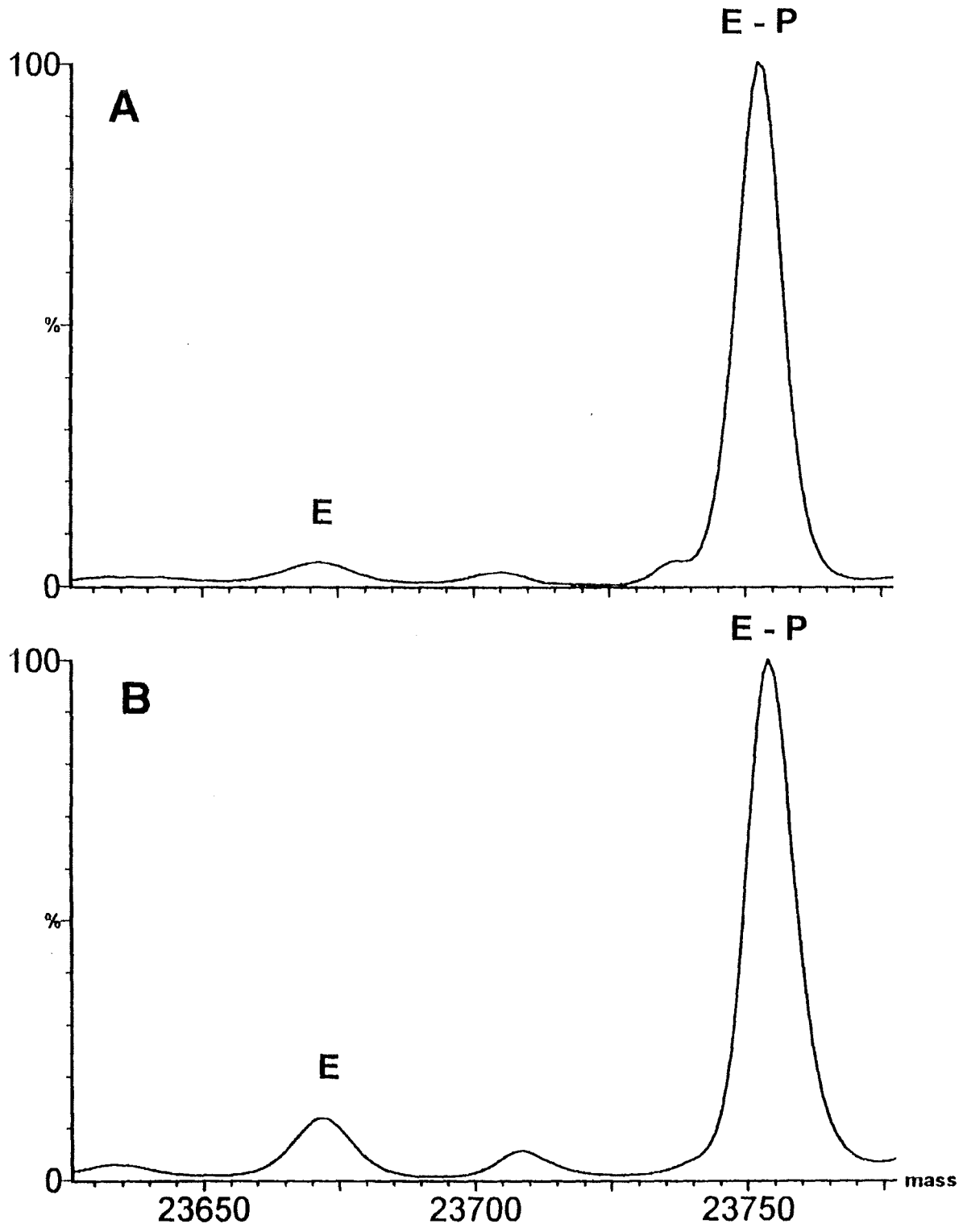

E - P
100
A
%
E
0
E - P
100
B
%
E
0
mass
23650
23700
23750

Figure 4-6 Far u.v. c.d. spectra of *S. pombe* PGAM.

Spectra were recorded in cells of pathlength 0.02 cm at a protein concentration of 0.3 mg/ml.

The spectra are depicted as: wild-type, (——); H151Q, (· · · ·); H196Q ,(— - — - —); H163Q, (- - -).

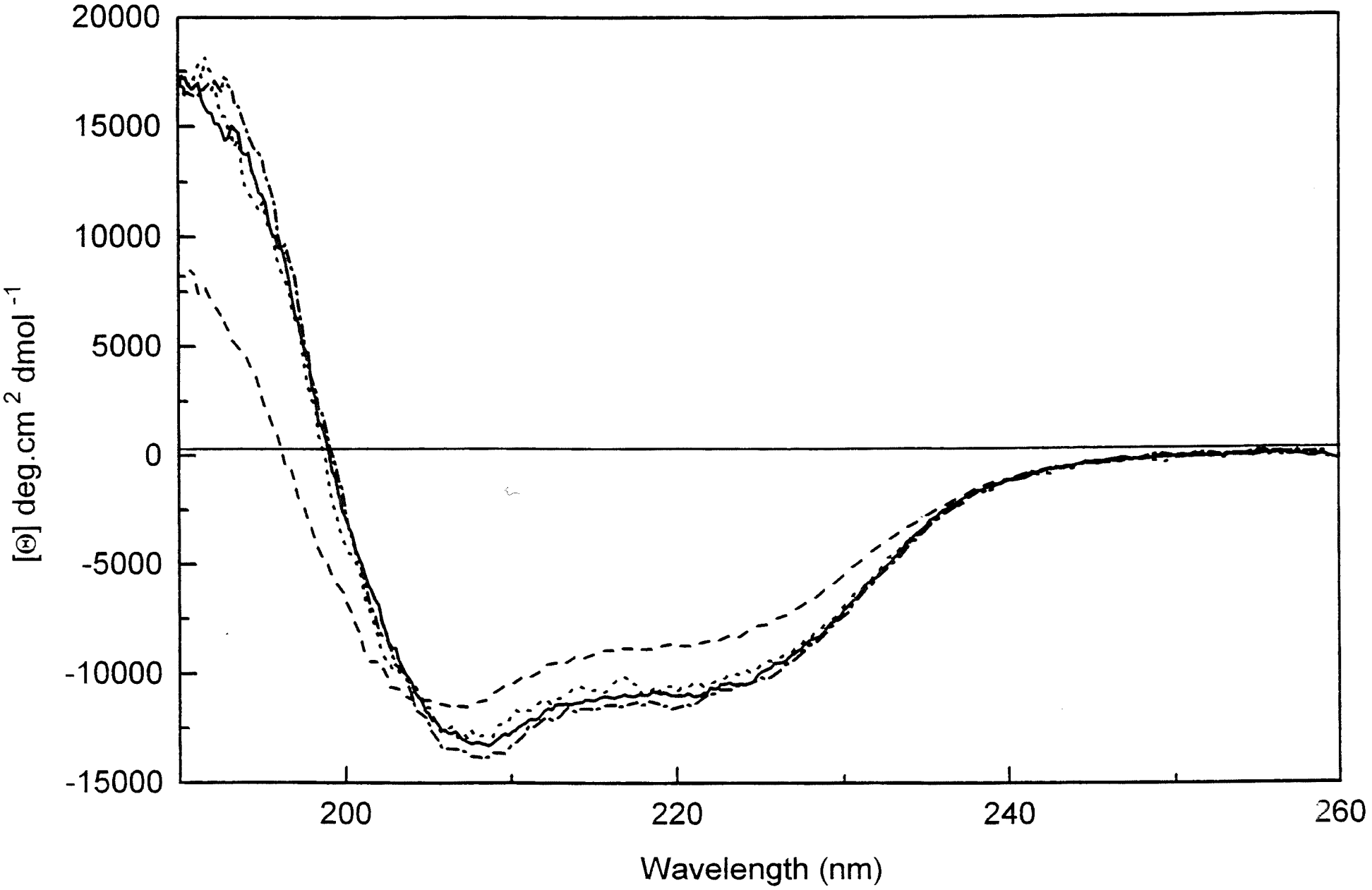
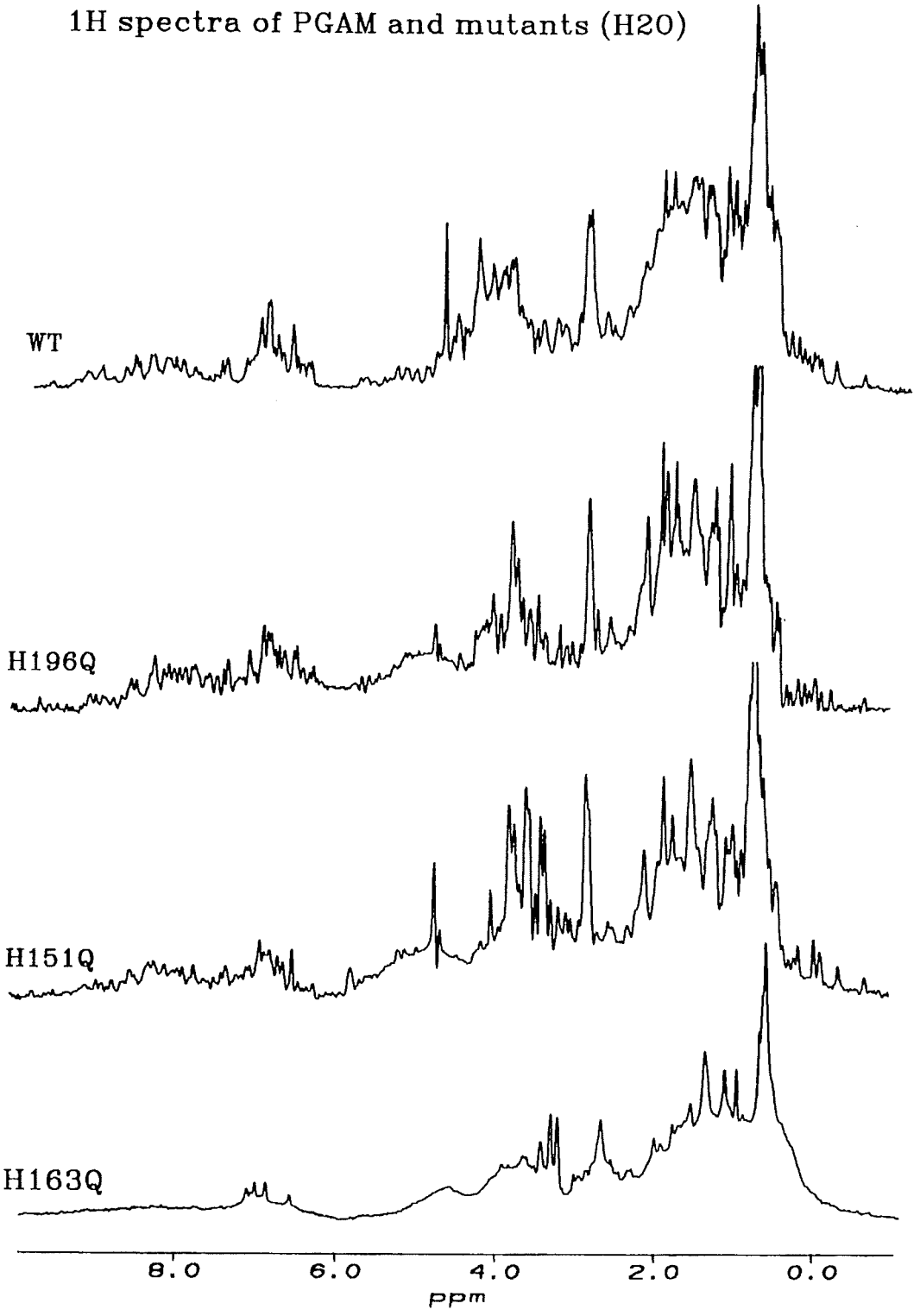

20000
15000
10000
5000
0
-5000
-10000
-15000
[Θ] deg.cm 2 dmol -1
200
220
240
260
Wavelength (nm)

Figure 4-7 600 MHz proton NMR spectra of wild-type and mutant PGAMs.

Spectra were recorded at 25°C in 90% H_2O as described in the text.


1H spectra of PGAM and mutants (H2O)
WT
H196Q
H151Q
H163Q
8.0
6.0
4.0
2.0
0.0
ppm

4.2.5 Discussion

The results described in this paper show that the gene encoding *S. pombe* PGAM can be successfully expressed in strain S150-gpm::HIS3 of *S. cerevisiae*. Good yields of both wild-type and mutant enzymes were obtained and easily purified to homogeneity. The mass spectra of the enzymes (Fig. 4-4) confirm directly the substitution of glutamine for histidine, and exemplify the exquisite resolving power of electrospray mass spectrometry in this type of application.

In earlier work [24], mass spectrometry has been used to examine the formation and decay of the phosphorylated form of PGAMs from *S. cerevisiae* and *S. pombe*. As previously reported the phosphorylated form of the *S. pombe* enzyme has a half life of less than 1 min at pH 8.0; this was confirmed in the present work and also found to be the case for the H151Q and H196Q mutants. The similar behaviour of these mutants to the wild-type enzyme is consistent with their comparable kinetic properties (Table 4-1).

As shown in Fig. 4-7, good quality 1D proton NMR spectra can be obtained for wild-type *S. pombe* PGAM and the H151Q and H196Q mutant enzymes. Taken together with some preliminary 2D NMR spectral information on the wild-type enzyme (data not shown), it would appear to be feasible to undertake a full assignment and structural characterisation of the enzyme using multiple isotope labelling techniques.

One aim of the experiments described in this paper was to examine the role of His-163 in *S. pombe* PGAM. In the case of the *S. cerevisiae* enzyme (where the corresponding residue is His-181), it was found that the mutase activity of the H181A mutant was reduced by some 10^4-fold compared with wild-type enzyme [27], consistent with the proposal that His-181 is involved in proton transfer reactions which are associated with the transfer of the phospho group to a suitable nucleophile, either water or 2- or 3-PGA. The decrease in the mutase and phosphatase activities (Table 4-1) and the increase in the stability of the phosphorylated form (Fig. 4-5) of the H163Q mutant of *S. pombe* PGAM would at first sight appear to indicate that His-163 plays an analogous role in this enzyme. However, the results of the spectral studies indicate that the structure of this mutant (H163Q) is different from wild-type enzyme, and it is therefore possible that the effects on the kinetic properties arise from these structural changes rather than from the alteration of the histidine *per se*.

These data provide further cautionary evidence that conclusions regarding the functional role of particular side chains reached by site-directed mutagenesis studies should always be regarded as provisional until it has been shown the mutation causes no detectable structural

alteration of the protein. The effects of substitution of amino acid residues could be on the pathway of folding and/or on the final structure itself. It should, however, be noted that despite the alteration in structure, His-15 of the H163Q mutant can be phosphorylated on addition of BPG (Fig. 4-5).

Inspection of the X-ray structure of *S. cerevisiae* PGAM [3] indicates a possible reason for the structural changes observed in the H163Q mutant of the *S. pombe* enzyme. The conserved Glu-86 in the *S. cerevisiae* enzyme (Glu-93 in *S. pombe* sequence) is much less exposed to solvent than would be expected for a charged side chain and may well be stabilised by an ionic interaction with the adjacent His-181 (His-163 in the *S. pombe* enzyme). It is possible therefore that if the histidine is replaced by glutamine, disruption of this interaction would destabilise the local structure leading to an altered structure. It is of interest that the H181A mutant of *S. cerevisiae* PGAM has a similar structure to wild-type enzyme as detected by CD or fluorescence; however the H181A mutant is less stable towards denaturation by GdnHCl. Addition of BPG to the H181A mutant (i.e. formation of the phosphorylated enzyme) is accompanied by very pronounced changes in both near and far UV CD spectra and by dissociation of the tetramer into dimers [27]. Further indications of the importance of interactions involving residues at or near the active site come from the observation that a number of patients with hereditary PGAM-M (muscle type enzyme) deficiencies have mutations in this region (E89A and R90W) [28] (the corresponding residues in the *S. cerevisiae* enzyme are Glu-86 and Arg-87). In the related enzyme, bisphosphoglycerate mutase, mutation of a corresponding residue (Arg-89) to cysteine, serine, glycine or lysine leads to derivatives which are much less active catalytically and also much less thermostable. Mutation of His-187 in this enzyme (corresponding to His-163 in *S. pombe* PGAM) to asparagine, tyrosine or aspartic acid also led to derivatives with similar losses of activity and stability [29,30]. Taken together, these findings clearly show the sensitivity of the overall structure of PGAM to perturbations at the active site.

4.2.6 Acknowledgements

We wish to thank the Biotechnology and Biological Sciences Research Council for provision of CD and mass spectrometer facilities, the Engineering and Physical Sciences Research Council for provision of the NMR facility, the Welcome Trust and European Commission for support, Drs. Paul Barlow and Stanislava Uhrinova for obtaining and interpreting the NMR spectra, Dr. S.M. Kingsman for supplying the plasmid pMA91, and Prof. John Coggins and Dr. Andrew Pitt for helpful discussions.

4.2.7 References

1. Fothergill-Gilmore, L.A. and Watson, H.C. (1989) *Adv. Enzymol. Relat. Areas Mol. Biol.* **62**, 227-313.
2. White, M.F. and Fothergill-Gilmore, L.A. (1988) *FEBS Lett.* **229**, 383-387.
3. Watson, H.C. (1982) Protein Data Bank Entry: Phosphoglycerate Mutase (Yeast) (3PGM), Brookhaven, New York.
4. Huang, Y. and Dennis, D.T. (1995) *Eur. J. Biochem.* **229**, 395-402.
5. Nairn, J., Price, N.C., Fothergill-Gilmore, L.A., Walker, G.E., Fothergill, J.E. and Dunbar, B. (1994) *Biochem. J.* **297**, 603-608.
6. Johnson, C.M. and Price, N.C. (1987) *Biochem. J.* **252**, 111-117.
7. Sedmak, J.J. and Grossberg, S.E. (1977) *Anal. Biochem.* **79**, 544-552.
8. Gill, S.C. and von Hippel, P.H. (1989) *Anal. Biochem.* **182**, 319-326.
9. Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983) *Gene* **24**, 1-14.
10. White, M.F. and Fothergill-Gilmore, L.A. (1992) *Eur. J. Biochem.* **207**, 709-714.
11. Chen, B. and Przybyla, A.E. (1994) *Biotechniques* **17**, 657-659.
12. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci.* **74**, 5463-5467.
13. Sali, A., Overington, J.P., Johnson, M.S. and Blundell, T.L. (1990) *Trends Biochem. Sci.* **15**, 235-240.
14. Brünger, A.T. (1990) XPLOR Version 2.1 Manual.
15. Geourjon, C. and Deleage, G. (1994) *Protein Engineering* **7**, 157-164.
16. Rost, B. and Sander, C. (1994) *Proteins* **19**, 55-72.
17. Summers, N.L., Carlson, W.D. and Karplus, M. (1987) *J. Mol. Biol.* **196**, 175-198.
18. Sutcliffe, M.J., Hayes, F.R.F. and Blundell, T.L. (1987) *Protein Engineering* **1**, 385-392.
19. MacArthur, M.W. and Thornton, J.M. (1991) *J. Mol. Biol.* **218**, 397-412.
20. Jones, T.H. and Thirup, S. (1986) *EMBO J.* **5**, 819-822.
21. Laskowski, R.A. (1993) *J. App. Crystallography* **26**, 283-291.
22. Lüthy, R., Bowie, J.U. and Eisenberg, D. (1992) *Nature* **356**, 83-85.
23. PROSA 1.0 Manual (1994) Centre of Applied Molecular Engineering, University of Salzburg.
24. Nairn, J., Krell, T., Coggins, J.R., Pitt, A.R., Fothergill-Gilmore, L.A., Walter, R. and Price, N.C. (1995) *FEBS Lett.* **359**, 192-194.

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

26. Morris, A.L., MacArthur, M.W., Hutchinson, E.G. and Thornton, J.M. (1992) *Proteins* **12**, 345-364.

27. White, M.F., Fothergill-Gilmore, L.A., Kelly, S.M. and Price, N.C. (1993) *Biochem. J.* **291**, 479-483.

28. Tsujino, S., Shanske, S., Sakoda, S., Toscano, A. and DiMauro, S. (1995) *Muscle and Nerve* **Suppl. 3**, S50-S53.

29. Rosa, R., Blouquit, Y., Calvin, M.-C., Prome, D., Prome, J.-C. and Rosa, J. (1989) *J. Biol. Chem.* **264**, 7837-7843.

30. Garel, M.-C., Lemarchandel, V., Calvin, M.-C., Arous, N., Craescu, C.T., Prehu, M.-O., Rosa, J. and Rosa, R. (1993) *Eur. J. Biochem.* **213**, 493-500.

5. Overexpression, characterisation and preliminary X-ray crystallographic analysis of Shikimate kinase from *Erwinia chrysanthemi*. Hyper-reactivity of the P-loop lysine at the ATP binding site.

Tino Krell, Adrian J. Lapthorn, Malcolm J. Horsburgh, Joseph E. Coyle and John R. Coggins

Overexpression, characterisation and preliminary X-ray crystallographic analysis of Shikimate kinase from Erwinia chrysanthemi. Hyper-reactivity of the P-loop lysine at the ATP binding site.

5.1.1 Summary

Shikimate kinase from *Erwinia chrysanthemi* has been overexpressed in *Escherichia coli*, purified and chemically and kinetically characterised. The lysine residue in the ATP binding site (the P-loop lysine) was found to be hyper-reactive to modification with trinitrobenzenesulfonic acid (TNBS). Light scattering was used to assess the suitability of different complexes of SK for crystallisation. The complex of SK with ADP and shikimate was the only monodisperse solution at 298K, and this combination was the only complex that gave crystals. The crystals, grown by the vapour-diffusion method using sodium chloride as a precipitant, were tetragonal, space group $P4_12_12$ or enantiomorph with cell dimensions a=b=108.5 Å and c=92.8Å (at 100K). Native crystals diffract to better than 2.6 Å on a Synchrotron X-ray source.

5.1.2 Introduction

Shikimate kinase (SK) (EC 2.7.1.71) catalyses the fifth reaction of the aromatic biosynthetic pathway (the shikimate pathway) converting shikimic acid into shikimate 3-phosphate using ATP as a co-substrate. Unusually for enzymes in the middle of metabolic pathways SK occurs in two different isoforms, SKI and SKII, and it has been proposed that shikimate may therefore be a branch-point intermediate for two different pathways (Weiss and Edwards, 1980). Both the *aroL* gene, which encodes SKII (De Feyter and Pittard, 1986a; Millar *et al.*, 1986), and the *aroK* gene, which encodes SKI, have been cloned and sequenced from *E. coli* (Whipp and Pittard, 1995; Løbner-Olsen and Marinus, 1992). SKII appears to play a major role in the shikimate pathway, its expression is controlled by the *tyrR* regulator gene and it is repressed by tyrosine and tryptophan (Ely and Pittard, 1979; DeFeyter *et al.*, 1986). The role of SKI is much less clear, its expression is constitutive and the enzyme has a very much lower affinity for shikimate (K_m for shikimate of 20 mM compared with 200 μM for SKII)(DeFeyter and Pittard, 1986b) and it has been suggested that in *E. coli* SKI has been displaced by the catalytically more efficient and better regulated SKII (Whipp and Pittard, 1995). Another possible explanation is that SKI normally carries out other functions in the cell and that it

phosphorylates shikimate only fortuitously (DeFeyter and Pittard, 1986b). This is consistent with the observation that SKI is associated with sensitivity to the antibiotic mecillinam which clearly implies an alternative biological role for SKI (Vinella *et al.*, 1996).

Shikimate kinases contain the Walker *et al.* (1981) type A-motif (GXXGXGK[T/S]), a common structural feature of many ATP/GTP-proteins. It was proposed by Milner-White *et al.*(1991) that this motif is part of a core structure consisting of four strands and four helices which is found in a very large number ATP/GTP-binding proteins. The GK[T/S] motif occurs in the P-loop which is the site of ATP/GTP binding (Mueller and Schulz, 1992). The lysine residue in the P-loop has been extensively studied by mutagenesis (Tian *et al* 1990; Korangy and Julin, 1992; Reinstein *et al.* 1990) and is thought to have a catalytic role (Reinstein *et al.*, 1990). To provide further mechanistic and structural information about this important family of proteins we have chemically characterised the conserved lysine residue in the ATP-binding site and commenced a crystallographic study of SKII. Our initial attempts to crystallise SKII from *E. coli* failed, but the SK of *Erwinia chrysanthemi,* which appears to be an SKII type enzyme (Minton *et al.*, 1989), yielded crystals suitable for X-ray analysis. Since the only lysine residue in SK of *E. chrysanthemi* is the P-loop residue the chemical characterisation was particularly straightforward.

5.1.3 Materials and Methods

Expression in *E. coli.* The coding region of the *aroL* gene from *E. chrysanthemi* was amplified from the plasmid pASN32 (Minton *et al.*, 1989) using VENT polymerase (New England Biolabs) and the two synthetic oligonucleotides SKin1 and SKin2. SKin1 (TCGTGGG**CATATG**ACAGAACCCATTTTTATG) contained an *Nde*I site (bold) to facilitate cloning and the 5' end of the coding sequence. SKinII (GG**AGATCT***TTA*GGCCGCAGGCAGACGCAT) contained a *Bgl*II site (bold) and had the TGA stop codon of the gene changed to TAA, which corresponds to TTA (italics) in 5'-3' direction and which is a preferred codon for highly expressed genes in *E. coli.* Amplification was performed with 2.5 U of polymerase, 100 pmol of primers, 250 nmols of each deoxy-nucleosidetriphosphate and 10 ng of plasmid DNA. The amplified gene was cloned into the *Nde*I-*Bgl*II sites of the T7 expression vector pTB361 (Brockbank and Barth, 1993). The construct (pTB361SK) was sequenced on both strands using synthetic oligonucleotide primers on denatured double-stranded DNA by the method of Sanger *et al.* (1980).

E. coli BL21(DE3) was transformed with pTB361SK using a standard $CaCl_2$ method (Sambrock *et al.*, 1989) Afterwards a positive colony was transformed with pLysS (Studier and Moffatt, 1986) using the same methodology.

For expression studies 20 ml of an overnight culture of *E. coli* BL21(DE3)pTB361SKpLysS was grown in Luria Bertrami medium (LB) containing tetracycline (12.5 μg/ml) and chloramphenicol (17 μg/ml) at 37 °C. 50 ml of LB plus antibiotics were then inoculated with 5 ml of overnight culture and grown to an $A_{600\ nm}$ of 0.6. The culture was then made 0.4 mM with respect to isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for a further 5-6 hours removing 1 ml samples at time-points. The $A_{600\ nm}$ of the samples were measured and then they were centrifuged at full speed for 1 min in a microfuge. The cells were resuspended in 10 μl of SDS gel sample buffer for every 0.1 $A_{600\ nm}$ unit and boiled for 5 min before loading 20 μl per well onto a SDS gel (Laemmli, 1970).

Large scale growth of bacteria for subsequent enzyme purification was accomplished by growing 500 ml cultures in 2 l flasks in a flat-box shaker at 37°C. Typically 25 ml of overnight culture was inoculated into 500 ml of LB containing appropriate antibiotics. Growth for 4-5 hours at 37°C was usually necessary before induction with IPTG and growth for a further 4 hours at 37°C for expression. A 6-liter growth typically yields 12 g (wet weight) of cells.

Protein Purification. All steps after cell breakage were performed at 4°C. *E .coli* BL21(DE3)pTB361SKplysS cells (14 g) suspended in 20 ml 50 mM Tris/HCl, pH 7.5 containing 20 mM KCl, 5 mM $MgCl_2$ and 0.4 mM DTT (buffer A) were broken by two passes through a French pressure cell. This material was then diluted with 80 ml buffer A and centrifuged at 100 000 x g for 1 hour. SK was purified from the resulting cell-free extract.

The supernatant was applied to a DEAE-Sephacel anion-exchange column (25 cm x 2.1 cm diam., flow rate of 20 ml/h) equilibrated in buffer A. The column was then washed with the same buffer until the $A_{280\ nm}$ of the eluate was less than 0.3. SK was eluted with a linear gradient of 0-300 mM KCl in 800 ml of buffer A (flow rate 40 ml/h) and fractions (14 ml) collected and assayed as described.

Pooled fractions were made up to 1.2 M $(NH_4)_2SO_4$ by addition of solid $(NH_4)_2SO_4$ and applied to a Phenyl-Sepharose hydrophobic interaction column (10 cm x 1.1 cm diam., flow rate of 10 ml/h) equilibrated in 100 mM Tris/HCl, pH 7.5 containing 1.2 M $(NH_4)_2SO_4$ and 0.4 mM DTT (buffer B). The column was then washed with 50 ml of buffer B, and the SK was eluted with a linear gradient (200 ml) of 1.2 M-0.0 M $(NH_4)_2SO_4$ in buffer B. The flow rate was 10 ml/h and 4 ml fractions were collected and assayed.

Fractions containing SK activity were pooled and dialysed against 55 mM Tris/HCl, pH 7.5 containing 550 mM KCl, 5.5 mM $MgCl_2$, 2.75 mM shikimic acid, 1.375 mM ADP and 0.45 mM DTT. The dialysed material was concentrated by vacuum dialysis and glycerol was added to a final concentration of 10% (v/v). The enzyme solution was applied to a Sephacryl S200 (superfine grade) column (85 cm x 2.5 cm diam., flow rate of 8 ml/h) that had been equilibrated in 50 mM Tris/HCl, pH 7.5 containing 500 mM KCl, 5 mM $MgCl_2$, 2.5 mM shikimic acid, 1.25 mM ADP and 0.4 mM DTT and the enzyme was eluted with the same buffer (the flow rate 8 ml/h, 4 ml fractions). Fractions containing SK activity were dialysed against buffer A containing 50% (v/v) glycerol before long-term storage at -20°C.

Protein concentrations were determined by the method of Bradford (1976) using bovine serum albumin as a standard, or spectrophotometrically at 280 nm in a Phillips Model PU8700 spectrophotometer using quartz cuvettes. The extinction coefficients were calculated according to Gill and von Hippel (1989). The concentration values determined by the two methods agreed within 5%.

Enzyme Assay. SK was assayed at 25°C by coupling the release of ADP to pyruvate kinase (EC 2.7.1.40) and lactate dehydrogenase (EC 1.1.1.27). Shikimate-dependent oxidation of NADH was monitored at 340 nm (ε=6180 $M^{-1}cm^{-1}$). The assay mixture contained 50 mM triethanolamine hydrochloride/KOH buffer, pH 7.0, 50 mM KCl, 5 mM $MgCl_2$, 1.6 mM shikimic acid, 5 mM ATP, 1 mM phosphoenolpyruvate, 0.1 mM NADH, 3 units of pyruvate kinase/ml and 2.5 units of lactate dehydrogenase/ml.

Mass Spectrometry. Positive ion electrospray mass spectrometry was performed on a VG Platform quadrupole mass spectrometer (2-3000 amu range) fitted with a pneumatically assisted electrospray (ionspray) source and controlled via the VG Mass-Lynx software (VG Biotech. Ltd, Altricham, Cheshire, UK). Carrier solvent (1:1 (v/v) acetonitrile/water) infusion was controlled at 10 µl/min using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, USA). Protein samples were dissolved in carrier solvent at a concentration of 20 pmol/µl, centrifuged at 5000 x g for 2 min and then 10-20 µl samples injected into the carrier stream. The instrument was calibrated immediately before use with horse heart myoglobin.

Enzyme Modification. Enzyme modification were carried out in a volume of 2 ml, at 25°C with constant stirring. During the inactivation enzyme aliquots were taken for enzyme assay. The percentage of inactivation was calculated as the ratio of enzyme activity after a certain time of treatment to the enzyme activity at time 0.

Modification of SK with phenylglyoxal (PGO)-Aliquots of SK (5 μg/ml) were incubated in 100 mM-sodium bicarbonate buffer, pH 9.1 for 5 minutes and then PGO was added (freshly made up 50 mM stock solution in water) to a final concentration of 0.5-4.0 mM.

Modification of SK with trinitrobenzenesulfonic acid (TNBS)-The inactivation with TNBS was carried out in the dark. The enzyme (4 μg/ml) was pre-incubated for 5 min in 50 mM borate buffer, pH 9.2. A 500 μM TNBS solution in the same buffer was prepared and aliquots added to a final TNBS concentration of 0.5-2.5 μM.

Substrate Protection against Modification with PGO and TNBS-For substrate protection experiments with SK 50 mM stock solutions of shikimate, ATP with the molar amount of $MgCl_2$ and shikimate with ADP and molar amounts of $MgCl_2$ were made up in the inactivation buffer and the pH readjusted using 4 M KOH. Defined aliquots of the stock solution was added to the pre-incubation mix and then the group specific reagent added.

Dynamic Light Scattering, Crystallisation and X-ray Analysis. Experiments were carried out using a DYNA-PRO. 801 dynamic light scattering/molecular sizing instrument (Protein Solutions, Buckinghamshire, UK) and recorded data were processed using AUTOPRO software. Protein solutions (1 mg/ml) were in 50 mM Tris/HCl, pH 7.6.

Enzyme samples were dialysed exhaustively into 20 mM Tris/HCl, pH 7.6. Afterwards shikimic acid and ADP (freshly made up solution in water, pH adjusted to 7.6) were added to a final concentration of 5 mM and $MgCl_2$ was added to a final concentration of 10 mM. Enzyme was concentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, UK) to a concentration of 16 mg/ml. Crystallisation was achieved at 293K by the sitting-drop vapour-diffusion technique. In the final condition for crystallisation the reservoir solution contained 2.16 M sodium chloride, 100 mM Hepes buffer, pH 6.9. Protein samples (6 μl) were mixed with equal amounts of reservoir solution and allowed to equilibrate.

The X-ray diffraction data were collected on the 9.5 beam line at the CLRC Daresbury Laboratory at 0.9199 Å wavelength using a MAR Research imaging-plate scanner. The crystal was soaked in crystallisation buffer containing 12.5% (v/v) glycerol for one minute and mounted in a loop. The crystal was immediately cryo-cooled to 100K using an Oxford Cryosystems cryostream cooler. A native dataset to 2.6 Å resolution was collected as 1 degree oscillation frames. Data frames were processed with DENZO and scaled with SCALEPACK (Otwinowski, 1993).

5.1.4 Results and Discussion

Overexpression and purification. SK of *E. chrysanthemi* is expressed as a soluble protein in *E. coli* where it accumulates to 20-30 % of the total soluble protein. The vector pLysS (Studier and Moffatt, 1986) was used to reduce the basal level of expression of SK prior to induction with IPTG. The pLysS plasmid was essential for overexpression since, in its absence cell growth was very poor. This suggests that SK is either toxic or inhibits cell growth.

The purification protocol (see Materials and Methods, Table 5-1) is based on the purification of the *E. coli* enzyme (Millar *et al.*, 1986). The main problem was the dramatic decrease of enzyme solubility after separation on Phenylsepharose. The enzyme was highly soluble in buffer A during cell-breakage and DEAE-Sephacel chromatography. However, dialysis of the SK containing eluate of the phenylsepharose column into buffer A resulted in major losses of the enzyme by irreversible precipitation. This problem could be circumvented by a dialysis into buffer A containing a substrate/product mixture of shikimic acid and ADP which increased enzyme solubility to at least 70 mg/ml. The substrate/product mixture needed to be present in the running buffer of the final Sephacryl S200 column. The adenylate kinases, a family of enzymes with high sequence homology to the shikimate kinases, have been shown to undergo major structural changes upon substrate binding (Vonrhein *et al.*, 1995). The increase in solubility of SK accompanying substrate binding suggests that such changes also occur with SK.

Characterisation. The purified protein was shown to be of high purity using electrospray mass spectrometry (Fig. 5-1). Two different enzyme species were detected, one corresponded to full-length enzyme (60%) and the other to enzyme lacking the N-terminal methionine (mass difference of 131 Da) (40%). This heterogeneity presumably results from incomplete processing of the enzyme by the *E. coli* cells due to the very high level of over-expression.

Table 5-1 Purification of shikimate kinase from *E. coli* BL21(DE3)pTB361SKpLysS.

The results presented are for a purification of 14 g of cells. Full details are given in the Material and Methods section.

	Vol.	Activity	Total activity	Yield	Protein	Total protein	Specific activity	Purification
	ml	*U/ml*	*U*	*%*	*mg/ml*	*mg*	*U/mg*	*fold*
crude extract	120	700	84 000	100.0	13.8	1662	50.5	1
DEAE-Sephacel	245	350	85 781	102.1	4.9	1210	70.8	1.4
Phenyl-Sepharose	128	495	63 360	75.4	4.8	615	103.0	2.0
Sephacryl S200	25	2192	54 812	65.3	12.3	319	178.3	3.5

Table 5-2 Kinetic properties of shikimate kinases from *E. chrysanthemi* and *E. coli.*

K_m values were calculated using Lineweaver-burk plots. The assay mix to measure K_m(shikimate) contained 5 mM ATP and 3 mM shikimate for K_m (ATP). [a] averaged value taken from Millar *et al.* (1986) and De Feyter and Pittard (1986b); [b] values taken from De Feyter and Pittard (1986b).

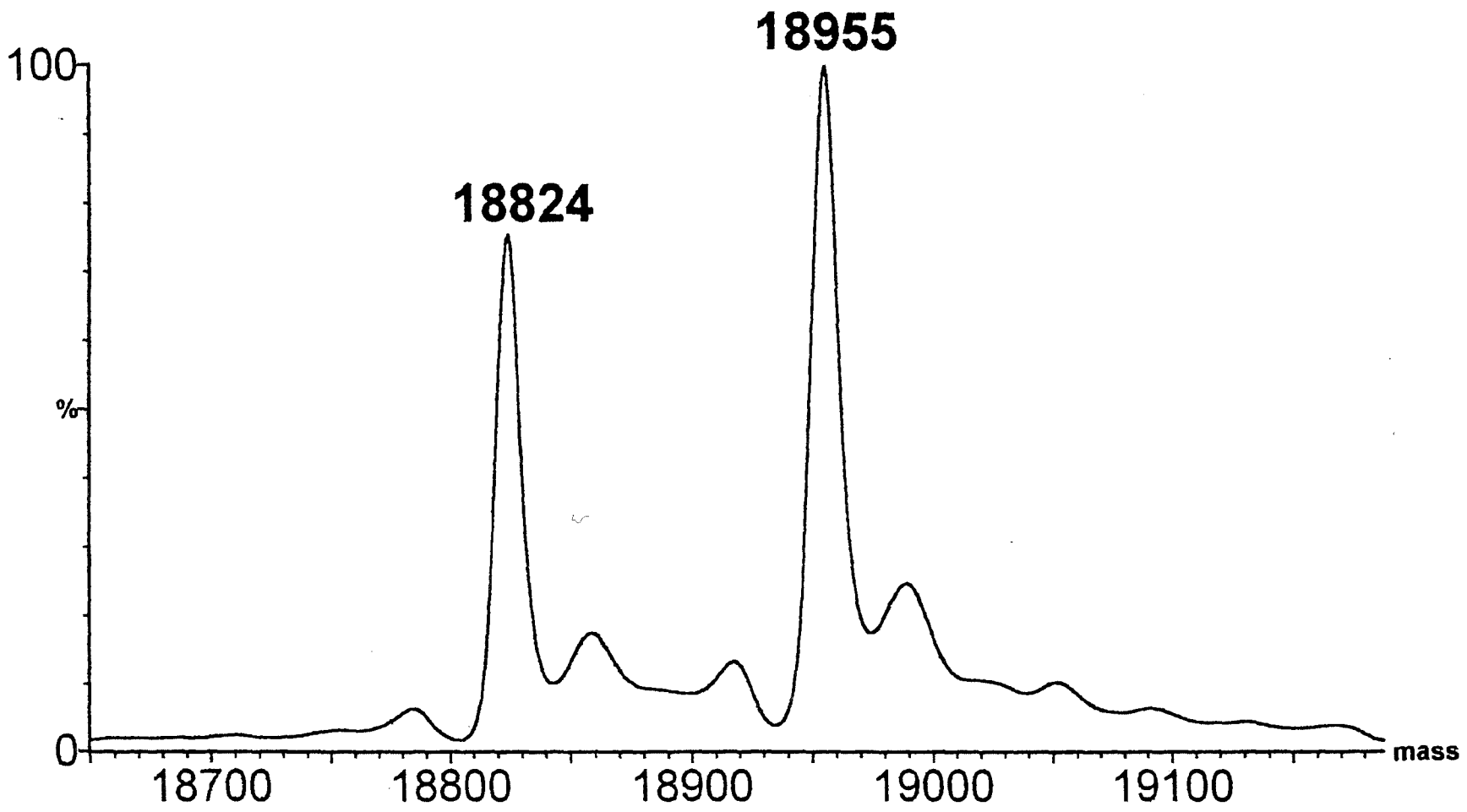
	K_m (shikimate)	K_m (ATP)	k_{cat}	k_{cat}/K_m (shikimate)
	μM	*μM*	s^{-1}	$M^{-1} s^{-1}$
E. chrysanth.	330	700	57.1	1.72×10^5
E. coli (SK II)	200[b]	160[b]	27.1[a]	1.36×10^5
E. coli (SK I)	20 000[b]	-	-	-

The major difference between SK I and SK II of *E. coli* is their K_m for shikimic acid (Table 5-2). The K_m and k_{cat}/K_m for the *E. chrysanthemi* enzyme of 330 μM and 1.72×10^5 $M^{-1} s^{-1}$ (Table 5-2) confirms the sequence based prediction that the *E. chrysanthemi* enzyme is a type II enzyme.

TNBS is a lysine specific reagent which can be used to measure the basicity of amino groups (Means and Feeney, 1971; Lundblad and Noyes, 1984). *E. chrysanthemi* SK has only one lysine residue, the P-loop K15. Incubation of SK with TNBS results in an extremely rapid decrease in enzyme activity (Fig. 5-2A) due to the reaction of the single, essential lysine residue. The inactivation shows pseudo-first order kinetics and the second order rate constant, as calculated from the plot of the apparent inactivation rate constants (k_{obs}) obtained from Fig. 5-2A against the concentration of TNBS, was 61 500 M^{-1} min^{-1}. The presence of ATP in the inactivation mixture affords protection from inactivation, whereas the presence of shikimic acid resulted in no significant change in the rate of inactivation (Fig. 5-2B). Evidence from mass spectrometry (data not shown), indicated that there was a single site of modification and together with the protection data suggested that inactivation was due to the reaction of the active centre K15.

Figure 5-1 MaxEnt deconvolution mass spectra of recombinant shikimate kinase from *E. chrysanthemi.*

MaxEnt spectra can be used for a quantitative analysis by comparing integrated peak-areas (Ferrige *et al.*, 1992). The crystallised protein was heterogeneous: the peak with the M_r of 18 955 confirms the theoretical M_r of 18 955.85 as deduced from the nucleotide sequence (Minton *et al.*, 1989). The peak with M_r of 18 824 corresponds to enzyme species with the N-terminal methionine residue cleaved off (mass difference of 131 Da). The small peaks to the right of the major peaks are potassium adducts.

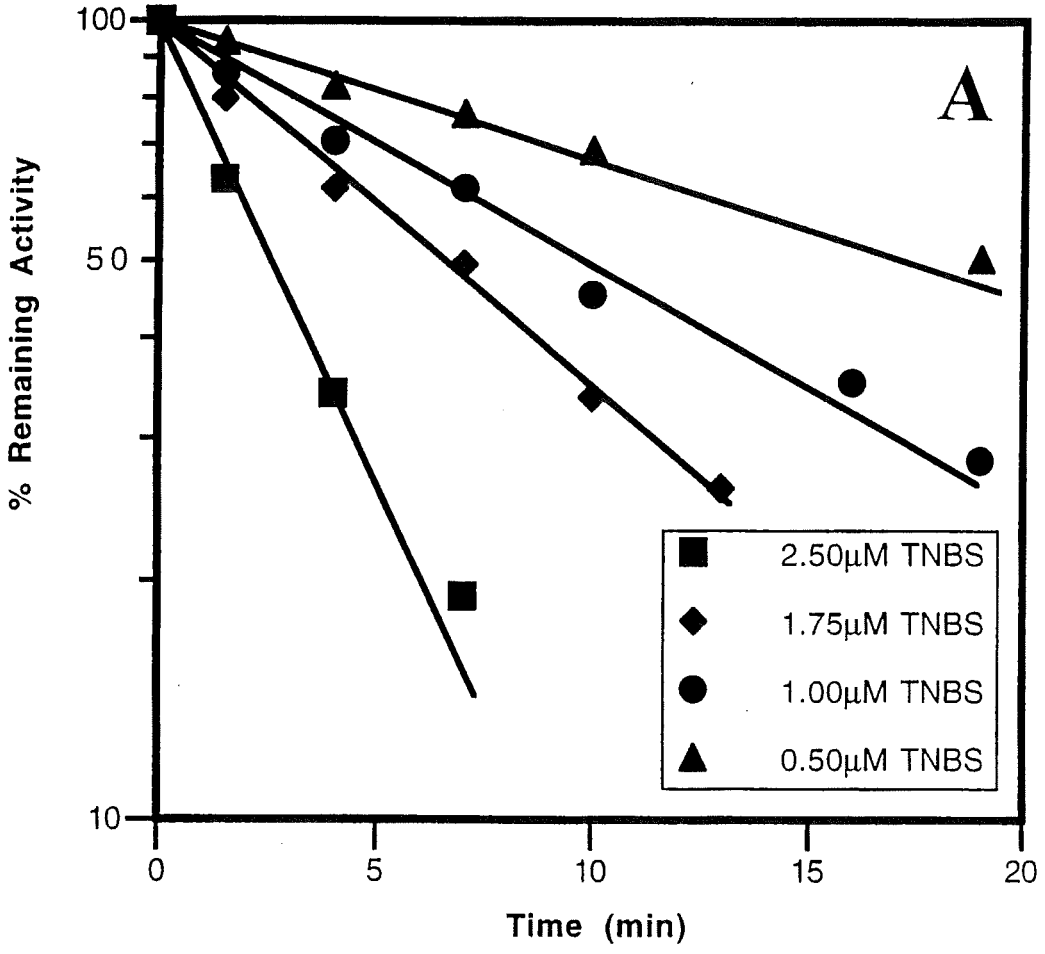

18955
18824
100
%
0
18700
18800
18900
19000
19100
mass

Another enzyme of the shikimate pathway, shikimate dehydrogenase, has been shown to have a lysine residue which is essential for enzyme activity (Chackrewarthy, 1995). Shikimate dehydrogenase was inactivated with TNBS using exactly the same procedure and buffer system as described in this work (Chackrewarthy, 1995). However, the second order rate constant was found to be only 400 M^{-1} min^{-1} (Chackrewarthy, 1995). This value is in agreement with TNBS inactivation studies of essential lysine residues of other non ATP utilising proteins (Goldin and Frieden, 1971; Suzuki *et al.*, 1995). Our data show that the P-loop lysine of SK is hyper-reactive and strongly suggest that the pK_a of this lysine residue is very much lower than the pK_a of typical protein lysine residues. This is consistent with the observation of Komatsutakaki (1995) which shows that the P-loop lysine of the β-subunit of ATP synthase is hyper-reactive towards pyridoxal phosphate. Experiments involving the TNBS modification of pyruvate kinase have also identified an essential hyper-reactive lysine (Johnson *et al.*, 1979). Taken together these data suggests a general role for hyper-reactive lysine residues in the active site of ATP utilising enzymes.

SK can be inactivated by the arginine specific reagent phenylglyoxal (Fig. 5-3, A). The reaction followed pseudo-first order kinetics (data not shown) and both shikimate alone and ADP alone provide significant protection against inactivation (Fig. 5-3, B) which implies a role for arginine residues both in shikimate and ADP binding. The second order rate constant for PGO inactivation of SK was 114 $M^{-1}min^{-1}$. This value is very similar to the rate of PGO inactivation for the type I dehydroquinase (*E. coli*) which has a second order rate constant of 107 $M^{-1}min^{-1}$ (T. Krell, unpublished data) and suggests a common role for arginine residues in substrate recognition by these two shikimate pathway enzymes. In the case of the type I dehydroquinase this arginine has been identified by mass spectrometry as R213 (Krell *et al.*, 1996).

Figure 5-2 Kinetics of inactivation of *E. chrysanthemi* shikimate kinase with different concentrations of TNBS.

A, Time course of reaction; B, Inactivation with TNBS (1.5 μM) in the presence of shikimate and ATP. Full experimental details in the Material and Methods section.


A
100
50
10
% Remaining Activity
0
5
10
15
20
Time (min)
2.50µM TNBS
1.75µM TNBS
1.00µM TNBS
0.50µM TNBS

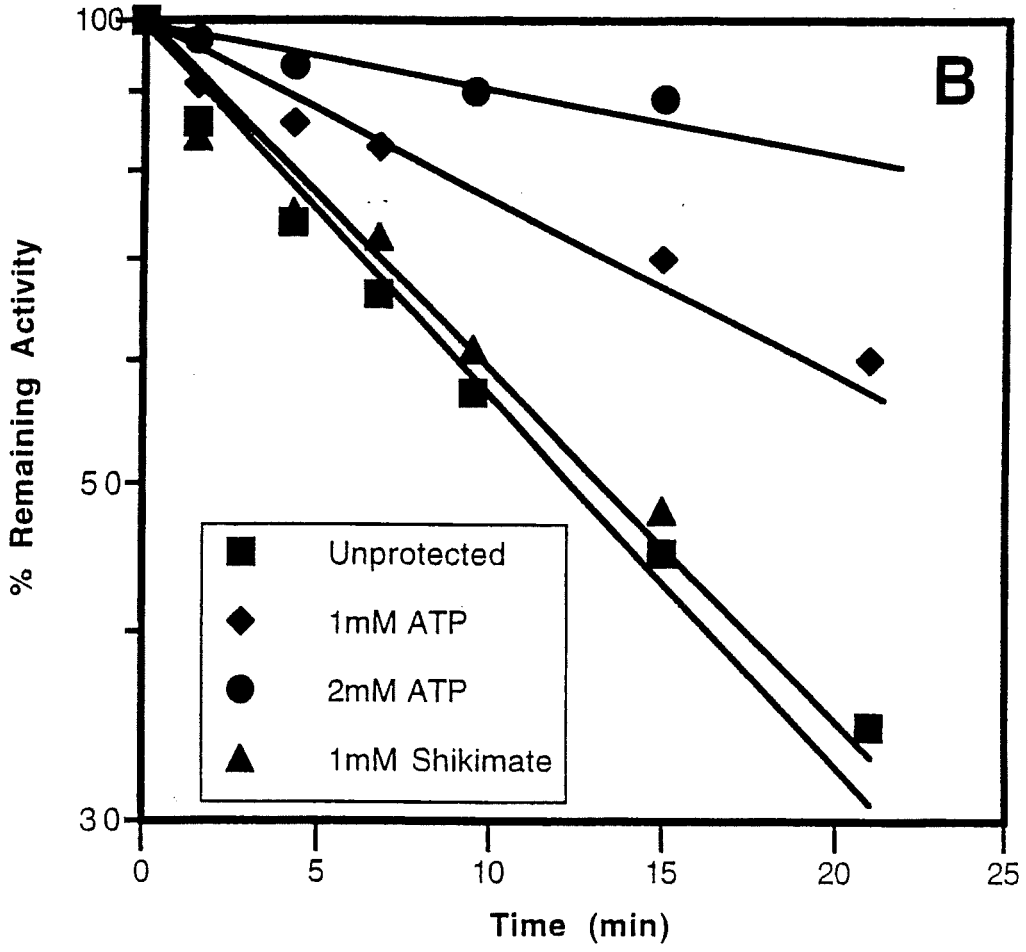
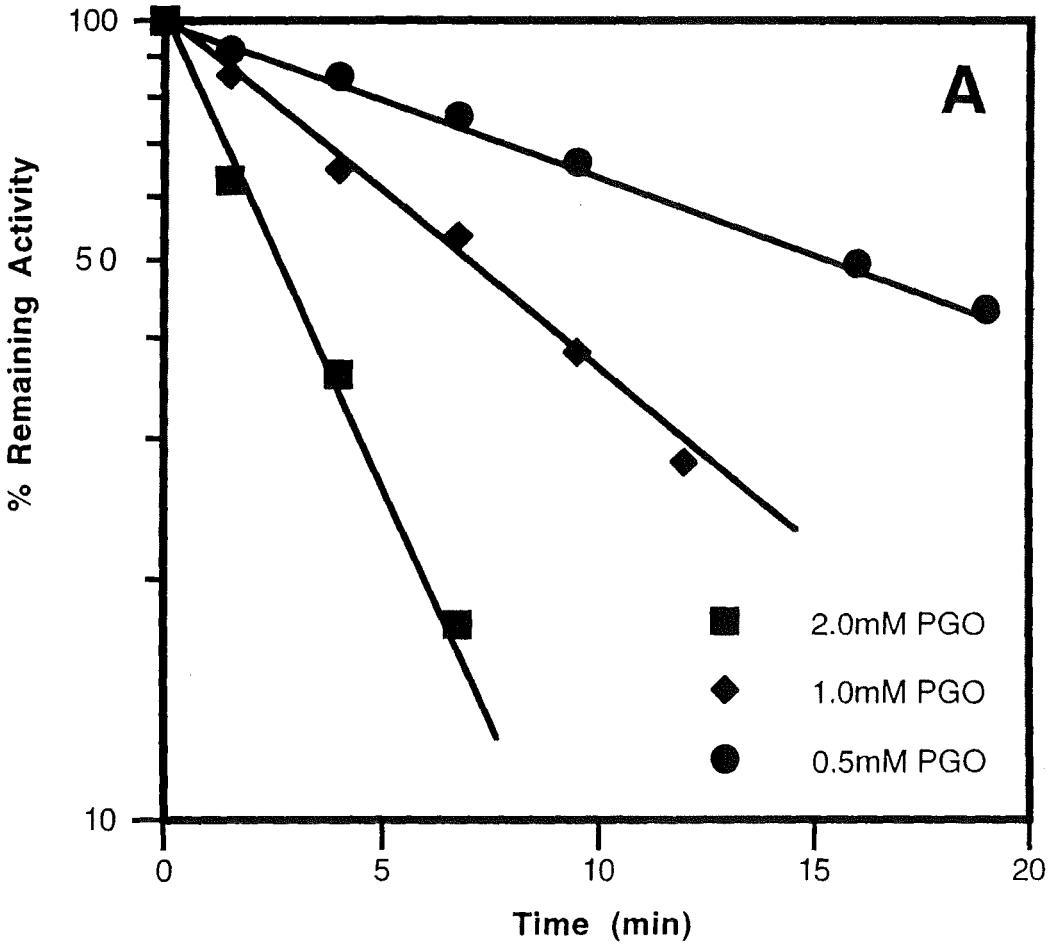

B
100
50
30
% Remaining Activity
0
5
10
15
20
25
Time (min)
Unprotected
1mM ATP
2mM ATP
1mM Shikimate

Figure 5-8 Kinetics of inactivation of *E. chrysanthemi* shikimate kinase with different concentrations of PGO.

A, Time course of reaction; B, Inactivation with PGO in the presence of shikimate, ATP and ADP. Full experimental details in the Material and Methods section. The second-order rate constant, as calculated from the plot of the apparent rate constants (k_{obs}) obtained from (A) against the concentration of PGO, was 114 M^{-1} min^{-1}.


100
50
10
% Remaining Activity
A
2.0mM PGO
1.0mM PGO
0.5mM PGO
0
5
10
15
20
Time (min)

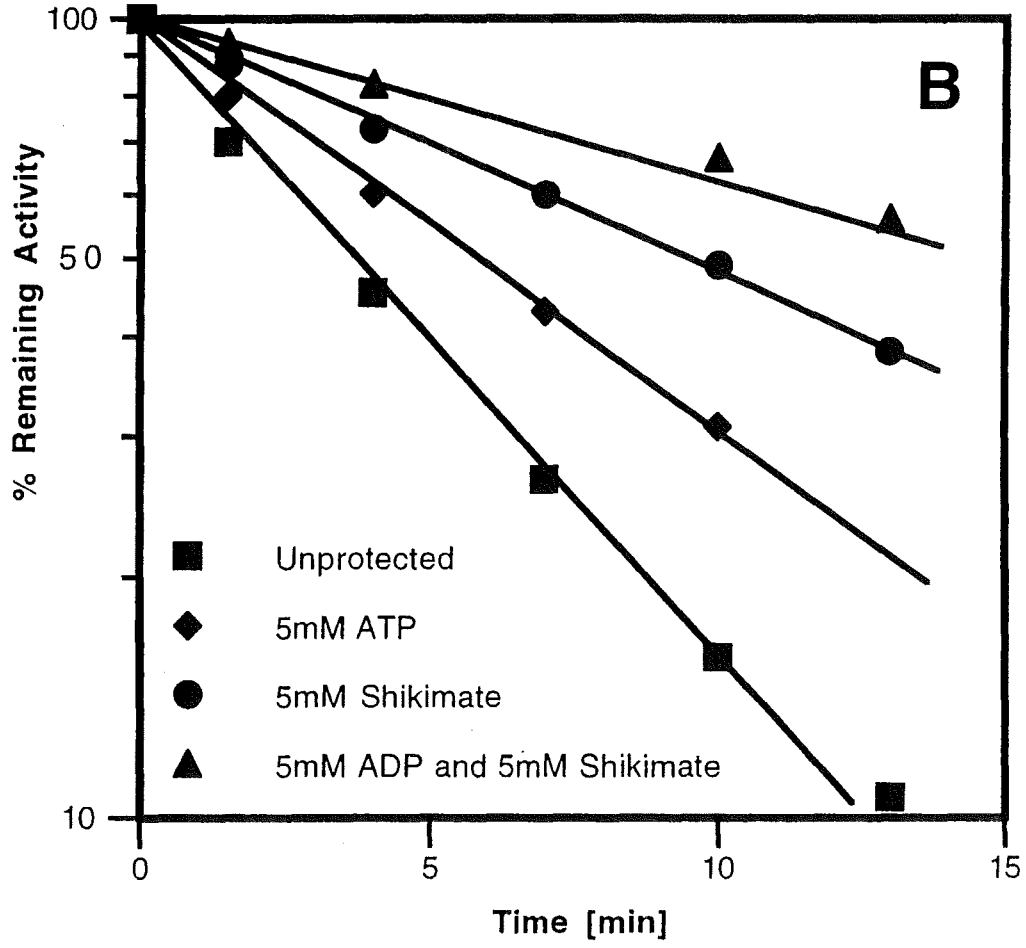

B
100
50
10
% Remaining Activity
0
5
10
15
Time [min]
Unprotected
5mM ATP
5mM Shikimate
5mM ADP and 5mM Shikimate

Table 5-3 Dynamic light scattering measurements of shikimate kinase (*E. chrysanthemi*) in the presence of different additives.

The hydrodynamic radius represents the median particle size present in the sample cell. The polydispersity value indicates the standard deviation of the spread of particle sizes about the reported average radius. A rise in polydispersity in relation to the average radius represents greater spread in the size distribution. The estimated M_W is calculated from the hydrodynamic radius (R_H) using an empirically derived relationship between the R_H and M_W values for a number of well-characterised globular proteins in buffered solution.

additives	hydrodynamic radius	polydispersity	estimated M_W	conclusion
	nm	*nm*	*kDa*	
shikimate (5 mM) ADP (5 mM) $MgCl_2$ (10 mM)	2.1±0.0	0.477±0.02	18	monodisperse
ADP (5 mM) $MgCl_2$ (10 mM)	3.0±0.0	1.698±0.01	43	polydisperse
shikimate (5 mM)	2.7±0.1	1.193±0.20	30	polydisperse

Dynamic Light Scattering, Crystallisation and X-ray Analysis. Dynamic light scattering experiments were carried out in parallel with crystallisation trials to assess conditions for the monodispersity of the enzyme. Several conditions were investigated and it was shown that monodispersity could only be achieved in the presence of the substrate shikimate with ADP and $MgCl_2$ (Table 5-3). Under those conditions the molecular weight obtained was 18 000 Da which confirms the monomeric state of the enzyme. Furthermore, the solubility of the enzyme was found to be greatly increased in the presence of shikimate, ADP and $MgCl_2$ (to at least 70 mg/ml).

Crystallisation conditions were screened using the sparse-matrix method (Jancarik and Kim, 1991) at 293K using the sitting-drop vapour-diffusion technique and 3 mg/ml protein. Crystallisation trials only yielded crystals in the presence of shikimate, ADP and $MgCl_2$ which confirms the value of dynamic light scattering for assessing the suitability of conditions for protein crystallisation. The optimised crystallisation

conditions are shown in materials and methods. Crystals appear after 10-12 days and continue to grow as tetragonal bipyramids up to a maximum size of 0.7 x 0.2 x 0.2 mm (Fig. 5-4).

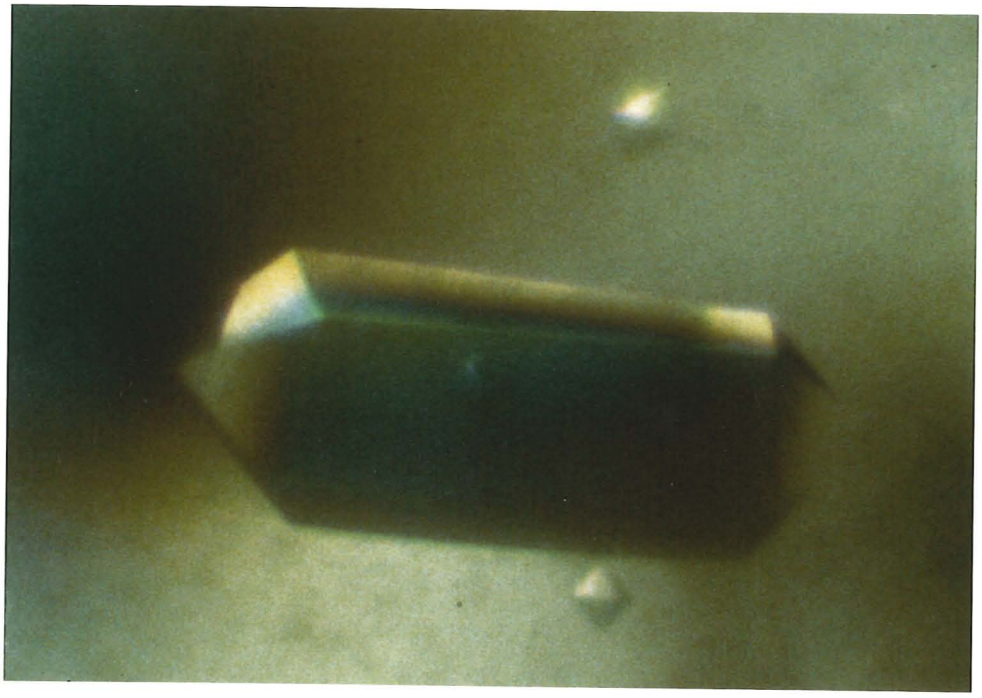


Figure 5-4 A crystal of shikimate kinase from *E. chrysanthemi*.

The size is approximately 0.6 x 0.175 x 0.175 mm.

Native crystals diffract to better than 2.6 Å on a Synchrotron X-ray source at 100K. Cryo-cooling was essential since crystals were radiation sensitive. A complete dataset was collected at the Daresbury SRS station 9.5. The data were processed with the programme DENZO (Otwinowski, 1993) and the crystals were found to belong to the tetragonal crystal system, with unit-cell dimensions of a=b=108.5 Å and c=92.8 Å. Analysis of the systematic absences in the data revealed absences at h=2 and l=4n along the (h00) and (00l) axes, respectively, which are consistent with the space group of $P4_12_12$ or enantiomorph. An assumption of two molecules per asymmetric unit leads to an acceptable packing density, V_m of 3.6 $Å^3$ Da^{-1}, corresponding to a solvent content of 66% (Mathews, 1968).

5.1.5 Acknowledgements

This research was supported by BBSRC and the University of Glasgow. We would like to thank Dr. Jim Naithsmith at the University of St. Andrews for use of his dynamic light scattering equipment, Dr. I. Anton and Ms. Bryony Melville for assistance with molecular biology, Mr. A. Elwell for computer support, Mr. J. Greene for technical assistance, Mr. J. Maclean and Dr. J. Beauchamp for help with data collection.

5.1.6 References

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

Brockbank, S.M.V. and Barth, P.T. (1993) *J. Bacteiol.* **305**, 707-710.

Chackrewarthy, S., Ph.D. thesis, University of Glasgow.

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

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

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

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

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

Gill, C.S. and von Hippel, P.H. (1989) *Anal. Biochem.* **182**, 319-326.

Goldin, B.R. and Frieden, C. (1971) *Biochemistry* **10**, 3527-3534.

Jancarik, J. and Kim, S.H. (1991) *J. Appl. Cryst.* **24**, 299-329.

Johnson, S.C., Bailey, T., Becker, R.R., Cardenas, J.M. (1979) *Biochem. Biophys. Res.Comm.* **90**, 525-530.

Komatsutakaki, M. (1995) *Eur. J. Biochem.* **228**, 265-270.

Korangy, F. and Julin, D. A. (1992) *J. Biol. Chem.* **267**, 1727-1732.

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

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

Lundblad, R.L. and Noyes, C. M. (1984) *Chemical Reagents for Protein Modification*, CRC Press, Boca Ratón, Florida.

Mathews, B. W. (1968) *J. Mol. Biol.* **33**, 491-497.

Means, G.E. and Feeney, R.E. (1971) *Chemical modification of proteins*, Holden-Day, Inc, San Francisco, Cambridge, London, Amsterdam.

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

Milner-White, J.E., Coggins, J.R. and Anton, I.A. (1991) *J. Mol. Biol.* **221**, 751-754.

Minton, N.P., Whitehead, P.J., Atkinson, T. and Gilbert, H.J. (1989) *Nucleic Acids Res. 17*, 1769.

Mueller, C.W. and Schulz G.E. (1992) *J. Mol. Biol.* **224**, 159-177.

Otwinowski, Z. (1993) *Data Collection and Processing. Proceedings of the CCP4 Study Weekend*, pp. 56-62. Daresbury Laboratory, Warrington, England.

Reinstein, J., Schlichting, I. and Wittinghofer, A. (1990) *Biochemistry* **29**, 7451-7459.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989*) Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.H. and Roe, B.A. (1980) *J. Mol. Biol.* **143**, 161-178.

Studier, F.W. and Moffatt, B.A. (1986) *J. Mol. Biol.* **189**, 113-130.

Suzuki, K., Mizuguchi, M., Gomi, T. and Itagaki, E.J. (1995) *J. Biochem.* **117**, 579-585.

Tian, G., Yan, H., Jiang, R.-T., Kishi, F., Nakazawa, A. and Tsai, M.-D. (1990) *Biochemistry* **29**, 4296-4304.

Vinella, D., Gagny, B., Joseleau-Petit, D., D'Ari, R. and Cashel, M. (1996), *J. Bacteriol.* **178**, 3818-3828.

Vonrhein, C., Schlauderer, G.J. and Schulz, G. (1995) *Structure* **3**, 483-490.

Weiss, U. and Edwards, J.M. (1980) *The biosynthesis of aromatic compounds*, John Wiley and Sons, New York.

Whipp, M.J. and Pittard, A.J. (1995) *J. Bacteriol.* **177**, 1627-1629.

6. Identification of active site residues of *Escherichia coli* Shikimate Dehydrogenase.

Sureka Chackrewarthy, Tino Krell and John R. Coggins

Identification of active site residues of *Escherichia coli* Shikimate Dehydrogenase.

6.1.1 Abstract

The group specific reagents trinitrobenzene sulphonate and phenylglyoxal have been used to identify 3 active site residues (K15, K65 and R154) in *Escherichia coli* shikimate dehydrogenase. Electrospray mass spectrometry was used to characterise both the modified enzyme and to identify the modified peptides following enzymatic digestion. The involvement of these residues in the catalytic mechanism of *E. coli* shikimate dehydrogenase and the advantages of electrospray mass spectrometry in identifying active site residues of proteins are discussed.

6.1.2 Introduction

The shikimate pathway is an important target for the development of herbicides and antimicrobial agents (Davies *et al.*, 1994; Jude *et al.*, 1996). We are using a combined approach involving chemical modification and X-ray crystallography to characterise the active sites of the shikimate pathway enzymes to facilitate novel inhibitor design. Earlier work has indicated that lysine and arginine residues play key roles in substrate binding and catalysis for the dehydroquinases (Krell *et al.*, 1995; Krell *et al.* 1996a), EPSP synthase (Huynh *et al.*, 1988; Huynh *et al.*, 1988a; Padgette *et al.*, 1988) and for shikimate kinase (Krell *et al.*, 1996b). Here we report an investigation of the role of lysine and arginine residues in shikimate dehydrogenase (SDH).

SDH (E.C.1.1.1.25), which catalyses the fourth step in the shikimate pathway, uses NADPH to reduce 3-dehydroshikimate to shikimate (Yaniv and Gilvarg, 1955). The hydride transfer is stereospecific (Danzette and Azerad, 1974), the enzyme is monomeric (Chaudhuri *et al.*, 1987) with a chain length of 272 residues and an M_r of 29 414 (Anton and Coggins, 1988).

The lysine specific reagent trinitrobenzene sulfonate (TNBS) (Kotaki *et al.*, 1964; Okuyama and Satake, 1960) and the arginine specific reagent phenylglyoxal (PGO) (Takahashi, 1968; Lundblad and Noyes, 1984) have been used to demonstrate the involvement of two lysine residues and one arginine residue at the active site of

shikimate dehydrogenase. Electrospray mass spectrometry was used to monitor the reactions, to characterise singly and multiply modified enzyme species, and to localise the sites of modification.

6.1.3 Experimental

Materials-TNBS was from BDH; NADP from Boehringer (Mannheim, Germany); shikimate, PGO, trypsin and chymotrypsin from Sigma (Poole, Dorset, England), Guanidinium chloride from GibcoBRL (Paisley, Scotland). HPLC grade water, acetonitrile and trifluoroacetic acid were from Rathburn (Walkerburn, Scotland). Methyl shikimate was a kind gift from Dr. C. Abell (Department of Chemistry, University of Cambridge).

Enzyme purification and assay-SDH was purified to homogeneity from an overproducing strain of *E. coli* according to the procedure described by Chaudhuri *et al.*(1987). Enzyme was assayed in the reverse direction by monitoring the reduction of NADP at 340nm ($e=6.2x10^3M^{-1}cm^{-1}$) at 25°C. The standard assay mixture (1 ml) contained 2 mM NADP and 4 mM shikimate in 100 mM Na_2CO_3, pH 10.6.

Enzyme modification-Enzyme modification were carried out in a volume of 2 ml, at 25°C with constant stirring. During the inactivation enzyme aliquots were taken for enzyme assay. The percentage of inactivation was calculated as the ratio of enzyme activity after a certain time of treatment to the enzyme activity at time zero.

The modification of SDH with TNBS was carried out in the dark. The enzyme (30 μg/ml) was pre-incubated for 5 min in 50 mM borate buffer, pH 9.2. A 10 mM TNBS solution in the same buffer was prepared and aliquots added to a final TNBS concentration of 30-175 μM.

To modify SDH with PGO aliquots of enzyme (30 μg/ml) were incubated in 100 mM sodium bicarbonate buffer, pH 9.4 for 5 minutes and then PGO was added (freshly made up 50 mM stock solution in water) to a final concentration of 0.5-3.0 mM.

Substrate protection against modification with PGO and TNBS-For substrate protection experiments 50 mM stock solutions of shikimate, methyl shikimate and NADP were made up in the inactivation buffer and the pH readjusted using 4 M KOH. Defined aliquots of the stock solution were added to the pre-incubation mix and then the group specific reagent added.

Preparation of inactivated enzyme samples for ES-MS-Enzyme inactivated to different extents was prepared for mass spectrometry by stopping the reaction by gel filtration on a Sephadex G50 column (200 mm x 17 mm, flow rate 15 ml/h) eluted with 0.5% (w/v) ammonium bicarbonate, pH 8.0. To remove low molecular mass contaminants prior to mass spectrometry the enzyme samples were washed twice by diluting 50-fold with HPLC grade water and reconcentrated using Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, UK).

Proteolytic digests-Modified and native SDH were digested in a total volume of 2 ml with continuous stirring.

PGO modified enzyme samples were denatured in 0.5% (w/v) ammonium bicarbonate buffer, pH 9.0 containing 8 M urea for 1 hour at 37°C. 0.5% (w/v) ammonium bicarbonate buffer, pH 9.0 (37°C) was added to lower the urea concentration to 2 M and the mixture was incubated with 5% (w/w) trypsin (1 g/l stock solution in 1 mM HCl) for 4 hours at 37°C. Afterwards, another 2% (w/w) trypsin was added and incubation continued for further two hours. The digestion was stopped by freezing the samples at -80°C.

TNBS modified enzyme samples were denatured in 0.5% (w/v) ammonium bicarbonate buffer, pH 8.0 containing 4 M GdnHCl for 15 min at 30°C. 0.5% (w/v) ammonium bicarbonate buffer, pH 8.0 (30°C) was added to lower the GdnHCL concentration to 1 M and the mixture was incubated with 5% (w/w of SDH) chymotrypsin (fresh solution in water). Proteolysis was allowed to continue for 1 h and the digestion was stopped by freezing the samples at -80 °C.

Electrospray mass spectrometry-ES-MS was performed on a VG platform quadrupole mass spectrometer (2-3000 amu range) filled with a pneumatically assisted electrospray source and controlled via the VG MassLynx software (VG Biotech Ltd., Altrincham, Cheshire, U.K.). Carrier solvent [1:1(v/v) acetonitrile: water, 0.2% formic acid] infusion was controlled at 10µl/minute using a Harvard syringe pump (Harvard Apparatus, South Natic, MA, U.S.A.). Protein samples were dissolved in carrier solvent at a concentration of 20 pmol/ml, centrifuged at 5000 x g for 2 minutes and then 10-20 µl samples were injected. The MaxEnt deconvolution (Ferrige *et al.*, 1992) was applied for quantitative analysis of raw data using 1.0 Da peak width and 1.0 Da/channel resolution.

Liquid chromatography mass spectrometry (LC/MS)-The protein digests (typically 75 µg) were separated by HPLC on a C-4 reverse phase column

(2.0 x 150 mm; Delta-pac^{tm}, Waters, Waterford, Hertfordshire, UK) using 2% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid as the initial solvent (flow rate 0.25 ml/min); after an 8 min wash to remove GdnHCl a linear gradient of 2-70% acetonitrile (v/v) in 0.1% (v/v) trifluoracetic acid was applied to elute the peptides. The column eluate was introduced directly into the mass spectrometer with a drying gas flow of 400 l/h and the source temperature was set at 100°C. The absorption profile of the eluted peptides was recorded at 214 nm and centroid mass spectra in the range 400-1800 Da/z were recorded at 4 s intervals.

6.1.4 Results and Discussion

Identification of active site lysine residues-SDH can be inactivated by treatment with the lysine specific reagent TNBS (Fig. 6-1A). The inactivation kinetics followed pseudo-first order kinetics (data not shown) and the second-order rate constant, as calculated from the plot of pseudo-first order rate constants against the concentrations of TNBS, was 405 $M^{-1}min^{-1}$. Shikimate and NADP afforded protection against inactivation, however maximum protection was observed in the presence of both shikimate and NADP (Fig 6-1, B) indicating that the inactivation was due to modification of residues in or near the active site.

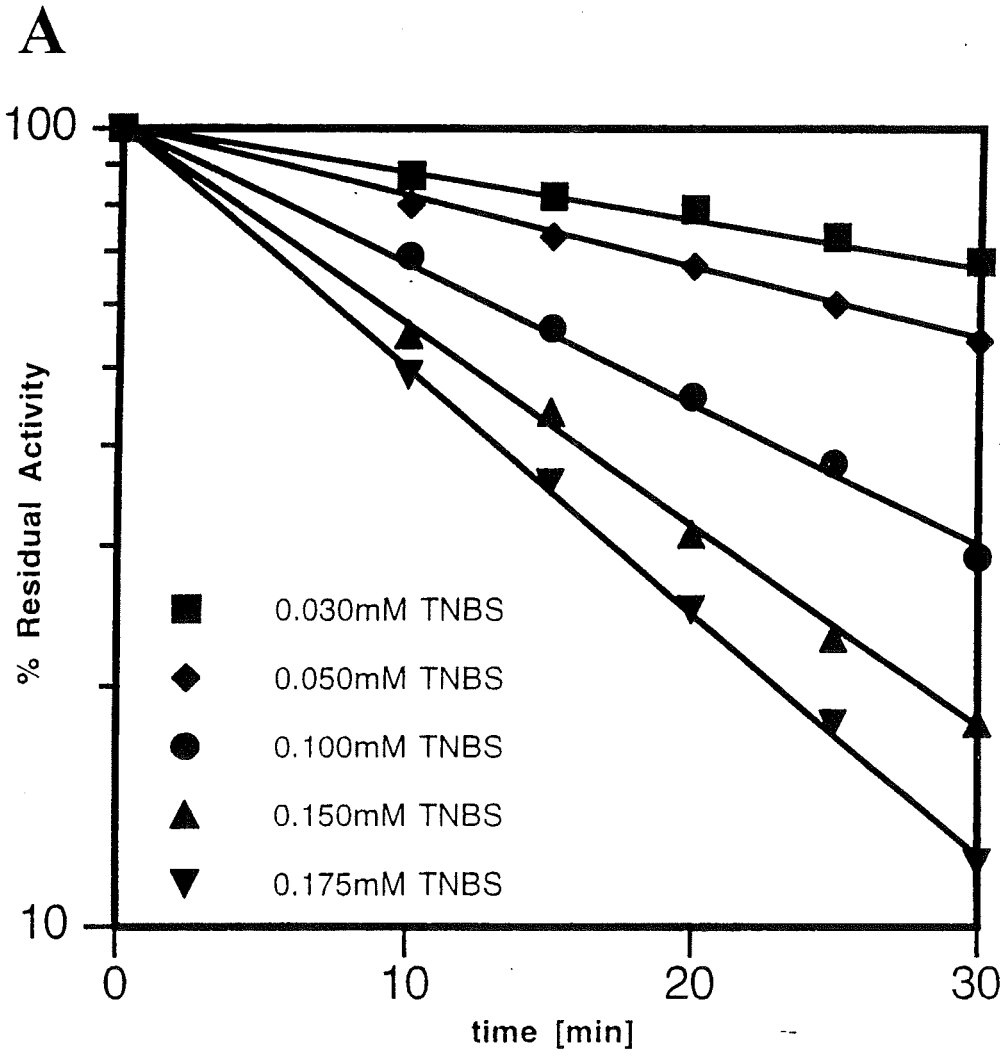
The stoichiometry of incorporation of TNBS into SDH was determined using ES-MS (Fig. 6-2). The 80% inactivated (unprotected) sample revealed modified enzyme species with three different masses (Fig. 6-2, A), corresponding to one, or two or three lysines modified per peptide chain. The sample protected by both NADP and shikimate was 86% active and ES-MS showed that it contained only one modified species with a mass of 29 625 Da (29 414 Da + 211 Da) (Fig. 6-2B) indicating that only one lysine was modified. Therefore, in the working enzyme, the substrates NADP and shikimate protect two lysine residues from reaction with TNBS.

To identify the sites of modification, unmodified and modified (protected and unprotected) enzyme samples were digested with chymotrypsin, the peptides separated using HPLC and directly injected into the electrospray mass spectrometer. Mass spectra were recorded every 4 seconds. Afterwards, all recorded spectra from the digest of the unmodified enzyme were scanned for the theoretical masses of chymotryptic peptides. All recorded spectra for digests of the modified enzyme (protected and unprotected) were scanned for theoretical masses of the chymotryptic peptides and for chymotryptic

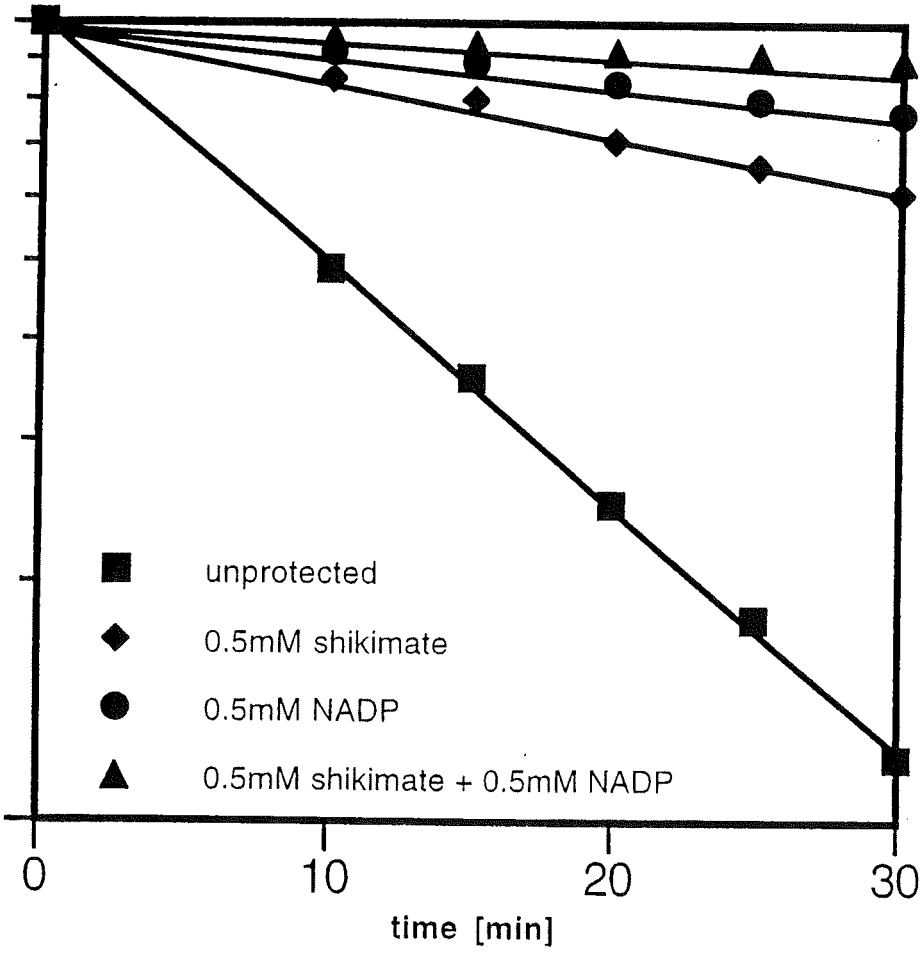
Figure 6-1 Kinetics of inactivation of *E. coli* SDH by TNBS.

Enzyme (30 μg/ml) was treated with TNBS in 50 mM borate buffer, pH 9.2 and aliquots were withdrawn at intervals for determination of activity.

(A) Inactivation of SDH by TNBS; the different concentrations of TNBS are shown in the box. (B) Inactivation of SDH by TNBS (0.175 mM) in the presence and absence of shikimate and NADP. (C) Inactivation of SDH by TNBS (0.20 mM) in the presence and absence of shikimate and methyl shikimate.


A
100
10
% Residual Activity
0.030mM TNBS
0.050mM TNBS
0.100mM TNBS
0.150mM TNBS
0.175mM TNBS
0
10
20
30
time [min]

B


unprotected
0.5mM shikimate
0.5mM NADP
0.5mM shikimate + 0.5mM NADP
0
10
20
30
time [min]

C

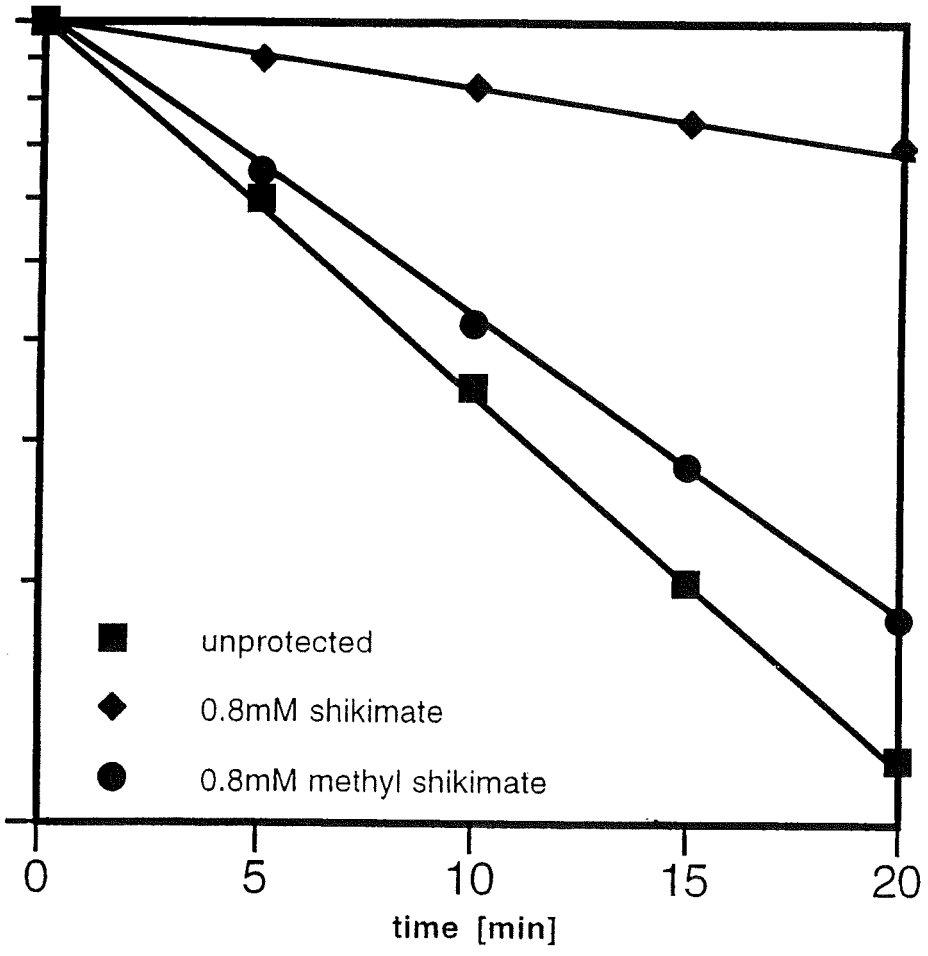
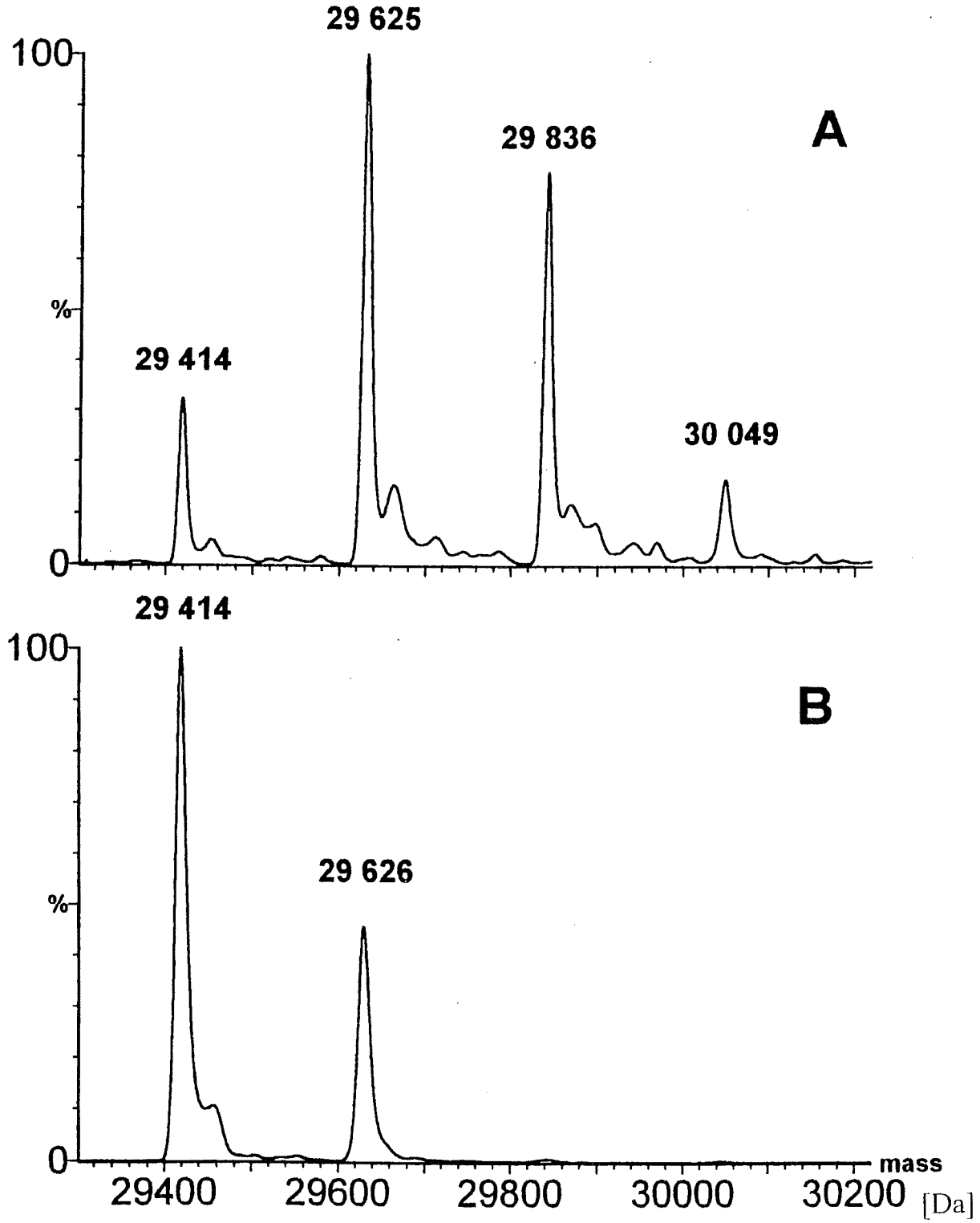

unprotected
0.8mM shikimate
0.8mM methyl shikimate
0
5
10
15
20
time [min]

Figure 6-2 MaxEnt deconvolution electrospray mass spectra of TNBS modified SDH.

In the absence (A) and presence (B) of shikimate (2 mM) and NADP (2 mM). Both enzyme samples were incubated with a 30 fold molar excess of TNBS for 1 hour. Afterwards, the residual enzyme activities were measured as 20% (A) and 86% (B) and the inactivation reactions were stopped by loading the samples onto a gel filtration column. Samples for mass spectrometry were prepared as described in Experimental. The native M_r of SDH is 29 414 Da and the reaction of one molecule TNBS results in a mass increase of 211 Da.


29 625
100
29 836
A
%
29 414
30 049
0
29 414
100
B
29 626
%
0
mass
29400
29600
29800
30000
30200
[Da]

peptides with an additional mass of 211 Da which corresponds to the incorporation of one molecule of TNBS.

In the digest of the unmodified enzyme all the lysine containing chymotryptic peptides were identified. The scan for TNBS modified peptides in the unprotected modified enzyme resulted in the detection of three modified chymotryptic peptides, all containing a single TNBS group corresponding to an additional mass of 211 Da: peptides 8-18 (1365.28 Da), 52-69 (2020.02 Da) and 216-222 (1015.94 Da) (Fig. 6-3, B). No such masses were detected in the digest of the unmodified enzyme. Peptide 52-69 contains two lysine residues one of which (K65) is next to a potential chymotrypsin cleavage site. The failure of chymotrypsin to cleave the peptide chain of the modified protein at position 64 suggests that K65 is the site of TNBS modification. Peptide 52-64 was observed as a minor component of the chymotryptic digest confirming the ability of chymotrypsin to cut the unmodified enzyme at this position (the modified enzyme sample contained 20% unmodified enzyme, see Fig. 6-2, A) and no modified peptide 52-64 was detected. The three lysine residues modified in the unprotected sample are thus K65, K15 and K217 or K219 (peptide 216-222 contains two lysine residues).

In the protected modified sample only peptide 216-222 showed an added mass of 211 Da (Fig. 6-3, A). No signals were obtained for a mass corresponding to peptide 8-18 +211 Da nor for mass corresponding to peptide 52-69 +211 Da indicating complete protection. From these data it appears that K65 and K15 are completely protected by shikimate and NADP from TNBS modification.

A plot of log k_{obs} against log [TNBS] yielded a straight line with a slope of 1.07 (data not shown) which suggests that only one essential lysine is responsible for the TNBS mediated inactivation of SDH. Comparison of SDH sequences from different organisms (Fig. 6-5) revealed that only K65 is completely conserved in an alignment of SDH sequences. There are no other conserved lysine residues, which provides further support for the proposal that K65 is the essential lysine residue in SDH.

Lysine residues play an important role in binding anionic substrates in proteins (Wilderstein *et al.*, 1992; Mildhausen and Levy, 1975). These experiments suggest that the essential lysine of SDH may be involved in binding the carboxylate group of shikimate. To test this hypothesis substrate protection experiments were carried out with the methyl ester of shikimate. As shown in Fig. 6-1C methyl shikimate, in comparison to

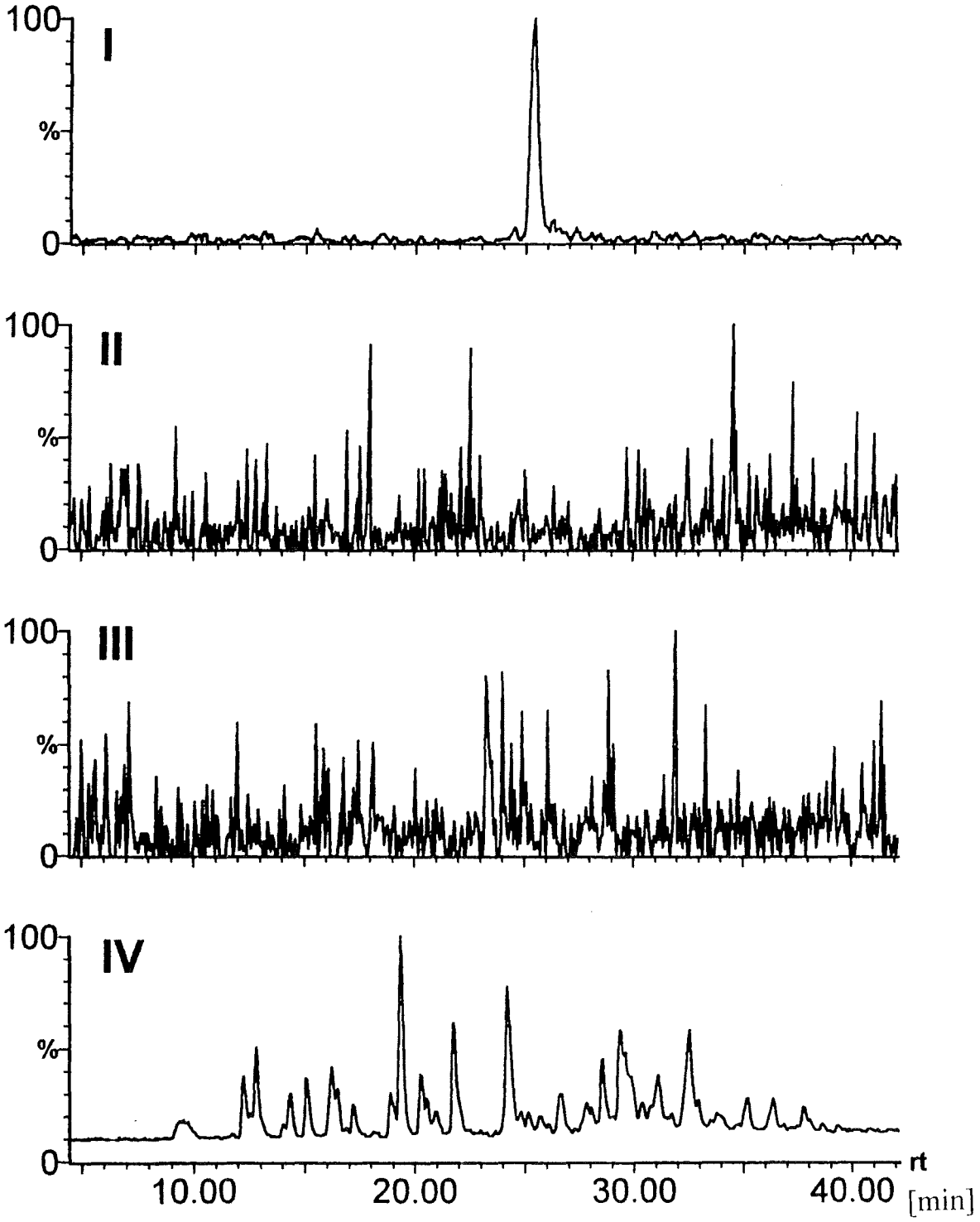
Figure 6-3 LC/MS data of chymotryptic digests of *E. coli* SDH inactivated after TNBS treatment.

Centroid spectra were collected every 4 sec.; traces I, II and II show the relative abundance of mass species in all the recorded spectra (around 650) against retention time; I, scan within all the recorded spectra for a mass of modified peptide 216-222 + 211 Da (1015.94 Da, corresponds to 1 TNBS attached); II, scan for a mass of modified peptide 8-18 + 211 Da (1365.28 Da); III, scan for the mass of modified peptide 52-69 + 211 Da (2020.02 Da); IV, simple HPLC trace at 215 nm against retention time (min).

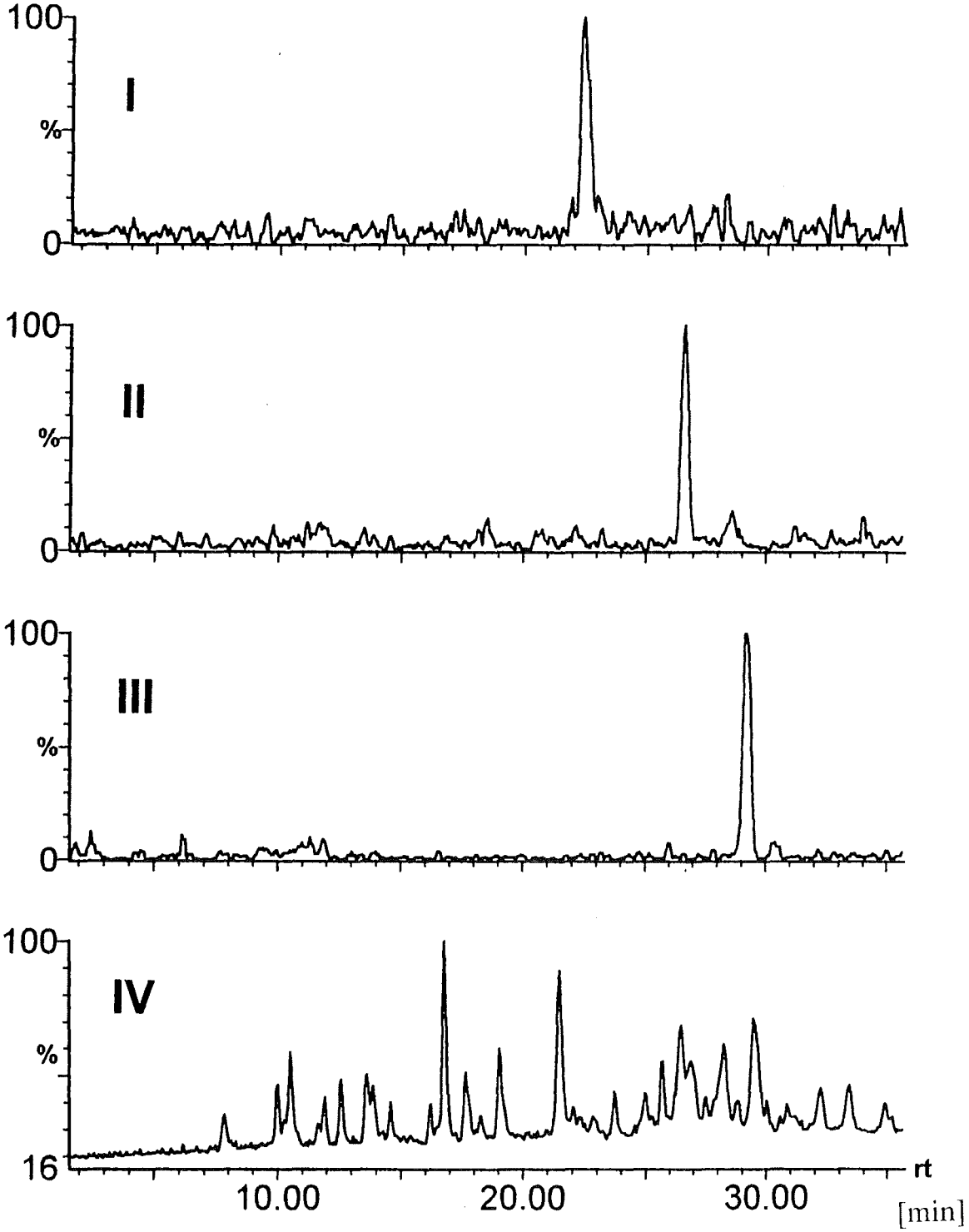
(A) Inactivation in the presence of NADP (2 mM) and shikimate (2 mM). The enzyme sample is the chymotryptic digest of modified SDH as shown in Figure 6-2B (86% residual activity and one site of modification) Traces II and III are typical for background noise, no peptides of the searched masses were detected. In contrast trace I, with a distinct signal above background noise.

(B) Inactivation in the absence of NADP and shikimate. The enzyme sample is the chymotryptic digest of modified SDH as shown in Figure 6-2A (20% residual activity and three sites of modification).

A



B



shikimate, affords only very little protection against TNBS inactivation, which is consistent with a role for the lysine in binding the shikimate carboxyl group.

Identification of an essential arginine residue-Inactivation of SDH with the arginine specific reagent PGO followed pseudo-first order kinetics (data not shown). NADP afforded strong protection against inactivation, whereas neither shikimate alone nor shikimate in combination with NADP afforded significant protection (data not shown).

Arginine residues react with PGO in a 1:1 or 1:2 stoichiometry (Krell *et al.*, 1995). Such adducts can be detected in the mass spectrometer according to their shifts in M_r of +116 Da (1:1 stoichiometry) or +250 Da (1:2 stoichiometry). The formation of the 1:1 adduct proceeds more rapidly than the subsequent reaction of a second molecule PGO with the 1:1 adduct (Krell *et al.*, 1995, Krell *et al.* 1996a).

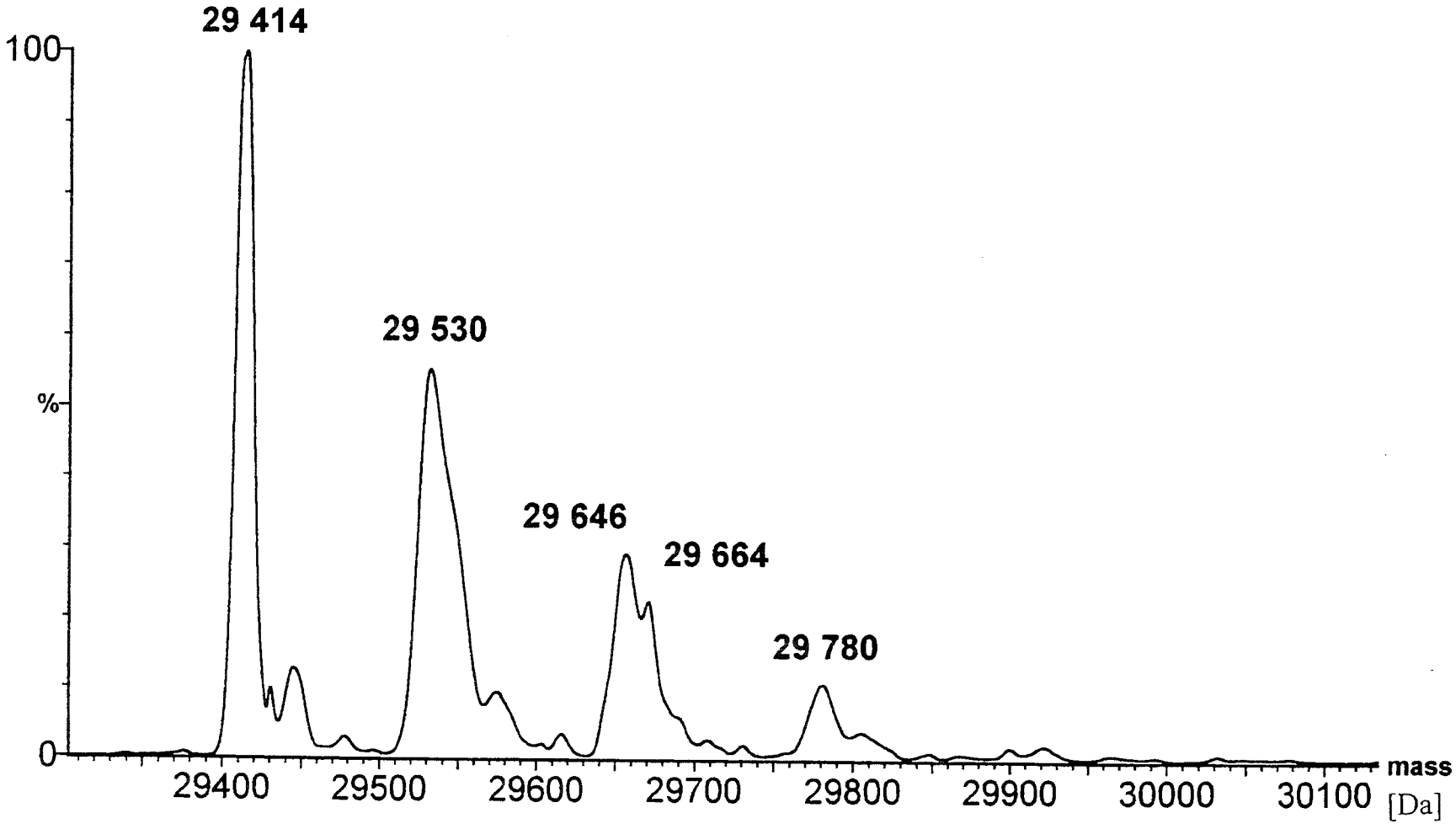
In the early stage of the modification of SDH by PGO the inactivation is due to the reaction at two sites. Figure 6-4 shows a mass spectrum of a 60% active sample. The two major modified species (29 530 Da and 29 646 Da) correspond to protein modified in 1:1 stoichiometry at one or two sites. The less abundant peak at 29 664 Da is due to the formation of a 1:2 adduct at a single site and the peak at 29 780 Da to protein modified at two sites, one containing a 1:2 adduct and the second a 1:1 adduct.

Proteolytic digestion in combination with LC/MS was again employed to localise the sites of reaction. Trypsin cleaves at unmodified arginine and lysine residues; no cleavage is expected after a modified arginine residue. Recorded spectra were therefore scanned for the theoretical tryptic peptides containing a single internal PGO-modified arginine residue (+116 Da). This tryptic digestion allowed the precise location of the site of reaction. Two modified peptides were detected: peptide 151-160 + 116 Da (1219.26 Da) and peptide 79-95 +116 Da (1874.04 Da). No such masses were found in the tryptic digest of the unmodified enzyme. This indicates that R154 and R90 are the major sites of modification.

The inactivation reaction with PGO was repeated in the presence of NADP. After the same time of modification the residual enzyme activity was 95% (compared with 60% from the unprotected sample). However, the only modified peptide found in the mass spectrometer corresponded to residues 79-95 +116 Da (1874.04 Da). No signal was obtained for peptide 151-160 +116 Da (1219.26 Da), indicating that R154 is completely protected from PGO modification by NADP.

Figure 6-4 MaxEnt deconvolution electrospray mass spectra of PGO modified SDH.

SDH was incubated in 1.5 mM PGO for 30 min. The residual activity was 60%.


29 414
29 530
29 646
29 664
29 780
100
%
0
29400
29500
29600
29700
29800
29900
30000
30100
mass
[Da]

```
                  1               15                                  45
E. coli...         MET..YA VFGNPIAHSK SPFIHQQFAQ QLNIEHPYGR VLAPINDFIN
H. influenzae..   .MDL..YA VWGNPIAQSK SPLIQNKLAA QTHQTMEYIA KLGDLDAFEQ
P. aeroginosa..   .MDR..YC VFGNPIGHSK SPLIHRLFAE QTGEALVYDA QLAPLDDFPG
B. aphidicola  MCKLEKFNYA LFGNPIDHSQ SPKIHNFFAT QTGILHIYKA INIPLDQFSS

                  46                  65                              95
E. coli        TLNAFFSAGG KGANVTVPFK EEAFARADEL TERAALAGAV NTLMRLEDGR
H. influenzae  QLLAFFEEGA KGCNITSPFK ERAYQLADEY SQRAKLAEAC NTLKKLDDGK
P. aeroginosa  FARRFFEQG. KGANVTVPFK EEAYRLVDEL SERATRAGAV NTLIRLADGR
B. aphidicola  VVSDFFKKNI KGANVTAPFK KEAYFFSDKL TERAKIAQSV NTLKKISDKC

                  96                                                 143
E. coli        LLGDNTDGVG LLSDL.ERLS FIRPGLRILL IGAGGASRGV LLPLLSL.DC
H. influenzae  LYADNTDGIG LVTDL.QRLN WLRPNQHVLI LGAGGATKGV LLPLLQA.QQ
P. aeroginosa  LRGDNTDGAG LLRDLTANAG VDVRGKRVLL LGAGGAVRGV LEPFLGECPA
B. aphidicola  ILGDNTDGIG LLSDLV.RLN FIKKNFSILI LGAGGAVKGV LLPLLSL.GC

                  144       154                                      193
E. coli        AVTITNRTVS RAEELAKLFA HTGSIQALSM DELEGHEFDL IINATSSGIS
H. influenzae  NIVLANRTFS KTKELAERFQ PYGNIQAVSM DSIPLQTYDL VINATSAGLS
P. aeroginosa  ELLIANRTAR KAVDLAERFA DLGAVHGCGF AEVEGP.FDL IVNGTSASLA
B. aphidicola  SVYILNRTIL NAKILVKQFN KYGKIFVFDR QNFKQQNFDL VINAMSRNTE

                  184                                                241
E. coli        GDIPAIPSSL IHPG.IYCYD MFYQKG.KTP FLAWCEQRGS KRNADGLGML
H. influenzae  GGTASVDAEI LKLG.SAFYD MQYAKGTDTP FIALCKSLGL TNVSDGFGML
P. aeroginosa  GDVPPLAQSV IEPGRTVCYD MMYAKEP.TA FNRWAAERGA ARTLDGLGML
B. aphidicola  KKNFTL...I LITSKRFFYD MNYS.TRNTP FINWCSKAGG SFISNGIGML

                  232                                 272
E. coli        VAQAAHAFLL WHGVLPDVEP VIKQLQEELS A.
H. influenzae  VAQAAHSFHL WRGVMPDFVS VYEQLKKAML ..
P. aeroginosa  VEQAAEAFFL WRGVRPASAP VLETLRRQLA TV
B. aphidicola  VFQAAYSFLE WHNVLPEINY IINILNIK.. ..
```

Figure 6-5 Sequence alignment of monofunctional SDH sequences.

The numbering refers to the *E. coli* enzyme, identified essential lysine and arginine residues are highlighted in bold.

The significant protection afforded by NADP against PGO inactivation indicates the presence of at least one arginine residue at the coenzyme binding site of SDH. Chemical modification (Yang and Schwert, 1972; Foster and Harrison, 1974) and crystallography (Scrutton *et al.*, 1990; Mittl *et al.*, 1994) have demonstrated the presence of arginine residues in the coenzyme binding site of several NADP linked dehydrogenases. R154 identified here as an essential residue in the NADP binding site of SDH is substituted by lysine or glutamine in some SDH sequences (Fig. 6-5). Since NAD does not provide any protection against PGO inactivation of SDH it seems likely that R154 hydrogen bonds to the 2' phosphate group of NADP.

6.1.5 Conclusion

This study demonstrates the simplicity of using electrospray mass spectrometry to monitor the stoichiometry of chemical modification reactions and for the rapid analysis of peptide maps for the location of modification sites. A particular advantage of this method for analysis chemically modified proteins is the lack of a need for radio-labelled reagents.

Using this approach three residues in the active site of shikimate dehydrogenase have been identified. K65, which is conserved in all known SDH sequences is proposed to have a role in binding the carboxylate group of shikimic acid, R154 is proposed to have a role in NADP binding. Protection studies with shikimate and NADP also suggest that K15, which is not conserved in other SDH's, is also present in the active site of the enzyme.

6.1.6 References

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

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

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

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, R.G. (1994) *Antimicrob. Agents Chemother.* **38**, 403-406.

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

Foster, M. and Harrison, J. H. (1974) *Biochem. Biophys. Res. Commun.* **58**, 263-267.

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

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

Jude, D. A., Ewart, C. D. C., Thain, J. L., Davies, G. M. and Nichols, W. W. (1996) *Biochim. Biophys. Acta* **1279**, 125-129.

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

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

Krell, T., Horsburgh, M. J., Cooper, A., Kelly, S. M. and Coggins, J. R. (1996a) *J. Biol. Chem.* **271**, 24492-24497.

Krell, T., Lapthorn, A. J., Horsburgh, M. J., Coyle, J. E. and Coggins, J. R. (1996b) manuscript submitted to Eur. J. Biochem.

Lundblad, R. L. and Noyes, C. M. (1984) *Chemical Reagents for Protein Modification,* Volume II, CRC Press, Boca Ratón.

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

Mittl, P. R. E., Berry, A., Scrutton, N. S., Perham, R. N. and Schulz, G. E. (1994) *Protein Science* **3**, 1504-1514.

Okuyama, T. and Satake, K. (1960) *J. Biochem. Tokyo* **47**, 454-461.

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

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

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

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

Yang, P. C. and Schwert, G. W. (1972) *Biochemistry* **11**, 2218-2224.

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

7. General Discussion

Arginine residue 23 in *S. coelicolor* DHQ (corresponding to R19 in the *A. nidulans* DHQ) has been shown in both species to be essential for enzyme function (chapter 3.1). The residue was clearly hyper-reactive to the arginine specific reagent PGO and mass spectrometry was used to demonstrate that the loss of enzyme activity was caused by the modification of a single residue. According to Patthy and Thesz (1979) such hyper-reactive arginines are likely to be involved in binding the carboxylate group of substrates.

R23 of *S. coelicolor* DHQ was replaced by lysine, glutamine and alanine residues using site-directed mutagenesis (chapter 3.2) and the mutant enzymes were found to have a much reduced k_{cat}. Interestingly, the K_m values for all three mutants were lower than for the native enzyme, indicating tighter substrate binding, which was confirmed by direct measurements of substrate binding using isothermal titration calorimetry (ITC). Those findings make a role for R23 in substrate binding unlikely. R23A mutant is 10 fold less active than R23K and R23Q (Table 3-3) which indicates the importance of a positive charge at that position. Therefore, R23 is unlikely to be involved in substrate binding but it may have a role in stabilising a negatively charged transition state.

Further support for this hypothesis comes from Bottomley *et al.* (1996a) who observed that the addition of a substrate/product mixture to the enzyme prior to PGO modification accelerates the rate of inactivation of *A. nidulans* DHQ. This is consistent with my observation that the presence of a substrate/product mixture does not protect R23 (*S. coelicolor*) from PGO modification. Those data indicate conformational changes upon substrate binding which might bring R23 in vicinity to the substrate, making it thereby even more reactive towards PGO.

The sequences of *S. coelicolor* and *Mycobacterium tuberculosis* DHQ contain only a single tryptophane residue (W66 for *S. coelicolor*). Using an approach of c.d. spectroscopy and fluorescence measurements the tryptophan residue was found to be close to the active site of both enzymes (Boam *et al.*). Furthermore, tyrosine residue 28 (*S. coelicolor*) has been identified as a residue in or close to the active site (Fig. 3-4).

In the last year the three-dimensional structure of *M. tuberculosis* DHQ (type II) has been solved (Gourley, 1996). Although the preliminary model of the enzyme is not

refined yet there is certainty about the protein topology, which is similar to p21 ras (Pai *et al.*, 1989) and other ATP/GTP binding proteins. All three identified active site residues (R23, Y28, W66) appear to be in a pocket which is likely to be the active site pocket of the enzyme (D. Gourley, personal communication).

The three-dimensional structure of the type I DHQ from *S. typhi,* complexed with borohydride reduced product, has also been solved (Polikarpov *et al.*). The overall structure is a parallel α/β-barrel with the active site located within the centre of the barrel. This general topology is similar to another hydrolyase, the glycolytic enzyme enolase (2-phospho-D-glycerate hydrolyase, EC 4.2.1.11) (Lebioda *et al.*, 1989). However, the catalytic mechanism is totally different; a "catalytic" metal ion plays the major role in the function of enolase (Faller *et al.*, 1977). Interestingly, the α/β-barrel topology has also been observed in structures of other enzymes employing Schiff-base formation in their mechanisms; examples are the class I aldolases (Sygusch *et al.*, 1987) and N-acetylneuraminate lyase (Izard *et al.*, 1994). Arginine residue 213 of the *E.coli* enzyme was identified as hyper-reactive and essential for enzyme function (chapter 3.2). The X-ray structure of the type I enzyme shows that R213 is clearly involved in substrate binding by interacting with the carboxylate group of the product dehydroshikimate. This is in contrast to the type II enzyme where the hyper-reactive arginine is unlikely to be involved in substrate binding.

Recombinant DHQ from *S. coelicolor* was expressed in *E. coli*, purified and characterised (chapter 3.2). However, the recombinant enzyme was found to be almost 20 times more active and with an K_m value almost twice as high as the previously isolated enzyme (White *et al.,* 1990). When the kinetic characterisation of the recombinant enzyme was repeated with less pure substrate similar results to these of White *et al.* (1990) were obtained. This suggests that the use of less pure substrate was the cause for the discrepancy in the kinetic data between the wild-type and the recombinant DHQ and underlines the necessity to check the purity of the synthesised substrate as shown in 2.4. Furthermore, those findings could be exploited to identify an enzyme inhibitor which is possibly present in the cruder substrate.

ES-MS was found to be a very useful tool to follow site-specific chemical modification of proteins. Clear information can be obtained about the number of modification sites as well as the stoichiometry of reaction. MaxEnt transformation allowed the relative quantification of modified species. The sites of modification were detected by measuring the masses of modified peptides after their separation on an on-line HPLC. The identification of essential amino acids was especially straightforward when initial inactivation of the enzyme was due to the reaction of a single amino acid residue.

The main limitation of this technique turned out to be the specificity of the chemical reagent as well as the stability of the adduct. The identification of the potentially essential histidine residues was attempted for both type II DHQ and SDH (data are not shown). Both enzymes appeared to be sensitive to treatment with the histidine specific reagent diethylpyrocarbonate (DEPC) (Miles, 1977; Dumas and Raushel, 1990). However, reaction of DEPC with both enzymes was very rapid. Unlike the PGO modifications no correlation between the relative amount of modified enzyme species (as observed in the mass spectrometer) to the residual enzyme activity (as determined by the enzyme assay) could be established. For example, a 20% DEPC-inactivated sample of SDH contained 5 modified species and very little native enzyme.

Furthermore DEPC was not found to be very specific to histidine residues, as for example a 70% DEPC-inactivated sample of DHQ (*S. coelicolor*) contained more modified enzyme species than histidine residues. Two amino acid residues (probably histidine residues) of SDH were protected by its substrates from inactivation with DEPC. However, the localisation of the sites of reaction failed since no DEPC adducts were detected by the mass spectrometer. SDH was found to be very difficult to digest (Chackrewarthy, 1995); to achieve complete tryptic digestion the enzyme needed to be denatured using 6 M urea prior to the addition of the protease, which resulted in an increase in pH. DEPC adducts are known to have only a limited half life at pH>8 (Miles, 1977). Therefore, the DEPC modifications which were detected mass spectrometrically for the enzyme were reversed during denaturation and tryptic digestion and this way escaped detection by LC/ES-MS.

No major difficulties were encountered in the detection of adducts caused by the modification by PGO, TNM and TNBS. Under the described conditions PGO was found to be very specific for arginine residues, not a single non-arginine adduct was ever identified in type I and type II DHQ as well as SDH samples modified with PGO. The reaction stoichiometry was complex; in contrast to Takahashi (1968) the rate-limiting step appeared to be the second step, the formation of the 2 PGO : 1 arginine adduct (Fig. 1-5). In the early part of the reaction (up to 70% of initial enzyme activity remaining) only 1 : 1 adducts were detected and only in the later stages of the inactivation reaction did 2 PGO : 1 arginine adducts appear. This reaction scheme was found to be the same in all PGO modification experiments which were carried out with enzymes from 5 different sources. Therefore, assuming a two phase reaction scheme, it is not possible to calculate the number of modified arginine residues in a protein from the amount of incorporated radiolabelled PGO using a 2 : 1 reaction stoichiometry.

However, a strict 1 : 1 reaction stoichiometry was observed for the reaction of p-Iodo-phenylglyoxal (data not shown) which was synthesised from p-hydroxy-acetphenone (Fodor and Kovacs, 1949; D. Gourley, unpublished data). Stereochemical hindrance caused by the bulky iodo-substitution is likely to be the reason for the fixed stoichiometry. p-Iodo-phenylglyoxal was synthesised to introduce a heavy atom into DHQ using its hyper-reactivity towards PGO. Introduction of heavy atoms into proteins is necessary if the crystallographic phase problem is to be solved by single or multiple isomorphous replacement. However, the p-iodo-phenylglyoxal adduct was not stable enough to survive the several weeks which were needed for crystal growth (D. Gourley, personal communication).

TNM appeared to undergo a secondary reaction (chapter 3.2, Fig. 3-4). Besides the nitration of the tyrosine residue an oxidation of the nitrated enzyme also occurred. The localisation and identification of the secondary site of reaction failed which suggests that the modification was reversed during denaturation and proteolytic digestion. TNBS was selective to lysine residues and no secondary reactions were observed.

Chapter 4 contains measurements of the rates of dephosphorylation of phosphorylated forms of phosphoglycerate mutases using ES-MS. Britton *et al.* (1972) have estimated the half life of the phosphorylated form of *Saccharomyces cerevisiae* PGAM to be 1-2 min. To employ ES-MS for measurements is was therefore essential to

find a very quick way to separate the phosphorylated enzyme from its phosphorylating agent, 2,3-bisphosphoglycerate, and to transfer the enzyme quickly into either water or a low ionic strength volatile buffer. In standard mass spectrometric measurements the separation and transfer into a volatile solvent was usually accomplished by gel-filtration followed by several rounds of diluting and reconcentrating the enzyme using Amicon concentrators. This procedure, which takes usually one day, is obviously not suitable for measuring fast rates. Instead, phosphorylated enzyme was applied to a NAP 5 column (Pharmacia, Uppsala, Sweden), equilibrated with 10 mM ammonium bicarbonate. This procedure allowed samples to be analysed by ES-MS within 3 min of mixing the enzyme with 2,3-BPG. Only negligible traces of 2,3-bisphosphoglycerate as well as the monophosphoglycerates were detected in the mass spectrometer.

In comparison to PGAM from *Schizosaccharomyces pombe* (Nairn *et al.*, 1994) the *Saccharomyces cerevisiae* enzyme (White and Fothergill-Gilmore, 1988) has been shown to have a C-terminal segment of 14 amino acids. This segment is thought to prevent water from entering the active site (Fothergill-Gilmore and Watson, 1989) which is reflected by the a lower phosphatase activity of the *S. cerevisiae* PGAM in comparison to the *S. pombe* enzyme (chapter 4.1). Interestingly, this was consistent with the measurement of the half life ($t_{1/2}$) of the phosphorylated forms of the enzymes; 35 min for the *S. cerevisiae* PGAM and less than 1 min for the *S. pombe* enzyme. Addition of the substrate analogue 2-phosphoglycollate to the phosphorylated form of the *S. cerevisiae* PGAM resulted in an approximately 100-fold acceleration of the dephosphorylation reaction. A mutant of *S. cerevisiae* PGAM was prepared lacking the 14 C-terminal amino acids (J. Nairn, unpublished data) and its kinetic parameters are currently being determined. This should give further evidence for the importance of the C-terminal segment.

Three site-directed mutants were prepared from the *S. pombe* enzyme, replacing histidine residues 151, 163 and 196 with glutamine residues. H151 and H196 are not conserved in multiple alignment of currently available PGAM sequences. However, H163 is a conserved residue and is thought to be involved in proton transfer reactions which are associated with the transfer of the phospho group to a nucleophile (White *et al.*, 1993). In comparison to the phosphorylated native enzyme the phosphorylated form of H163Q was found to be extremely stable, with a $t_{1/2}$ of at least 400 min which is a 400 fold increase in comparison to the native enzyme. Furthermore, a

230 fold decrease in the mutase activity of H163Q was observed. But it is unfortunately not possible to correlate those data since the amino acid substitution caused considerable changes in the secondary structure of the mutant enzyme as seen in both the near and far u.v.-c.d. spectra (Fig. 4-6). However, these experiments show the necessity of verifying the conformation of enzyme species altered by site-directed mutagenesis. The kinetic parameter of mutants H151Q and H196Q were very similar to the wild-type enzyme.

In previous years attempts to crystallise shikimate kinase II from *E. coli* have failed. It was therefore attempted to obtain an SK II from another source. The cloning of a DNA fragment of *Erwinia chrysanthemi* with 53% sequence identity to the *E. coli* SK II was reported in 1989 (Minton). This fragment, which was very likely to be a SK II enzyme, was cloned into a T7 based expression system and overexpression of the enzyme was achieved to an amount of up to 30% of the total cellular protein (chapter 5).

A purification protocol, based on the purification of the *E.coli* SKII (Millar *et al.*, 1986), was developed. The main problem to overcome was the drastic decrease in enzyme solubility during dialysis into 50 mM Tris/HCl after separation on Phenyl-Sepharose. This might indicate the presence of a compound which forms a complex (regulatory protein?) with the enzyme during the initial steps of the purification and which is separated from SK by hydrophobic interaction chromatography. However, addition of ADP and shikimate greatly increased the solubility of SK to at least 70 mg/ml. The increase in solubility suggests major structural changes accompanying substrate binding. This is in analogy with other kinases such as adenylate kinase (Schulz *et al.*, 1990).

The purified enzyme was found to be heterogeneous, consisting of 60% full length enzyme and 40% with the N-terminal methionine cleaved off (Fig. 5-1). This incomplete processing of the enzyme by the *E. coli* cells was probably due to the very high level of overexpression. The presence of the N-terminal methionine was found to be very useful for crystallographic studies since it turned out to be a platinum binding site. Kinetic characterisation clearly classify the *E. chrysanthemi* enzyme as a type II enzyme (Table 5-2). The role of SK I is still unclear, it must be somehow related to the biosynthesis of aromatic compounds since it is expressed in an operon together with another shikimate pathway enzyme. It has definitely a dual biological role (Vinella *et al.*, 1996) but maybe its primary function has shifted during evolution from

being involved in the biosynthesis of aromatic compounds to a still unknown second biological role which is reflected by the un-physiologically low affinity for shikimic acid (De Feyter and Pittard, 1986b).

E. chrysanthemi SK contains only one lysine residue, the one in the P-loop, which simplifies the characterisation of its chemical reactivity using the lysine specific reagent TNBS. The reactivity with TNBS measures the basicity of the lysine residue (Means *et al.*, 1972). Experiments were repeated several times before the measured second order rate constant of 61 500 M^{-1} min^{-1} was believed to be true. This value is around two orders of magnitude higher than that observed for lysine residues in non-ATP utilising enzymes (Chackrewarthy, 1995, Goldin and Frieden, 1971). However, there are reports about hyper-reactive lysine residues in the ATP binding sites of the β-subunit of ATP synthase as well as pyruvate kinase (Komatsutakaki, 1995; Johnson *et al.*, 1979). It is tempting to suggest a general role for hyper-reactive lysine residues in the active sites of ATP utilising enzymes.

SK of *E. chrysanthemi* can be inactivated by treatment with PGO. ADP as well as shikimic acid protect the protein from inactivation, suggesting roles for arginine residues in the binding of both substrates.

The enzyme complexed with ADP and shikimic acid was crystallised using the sitting-drop vapor-diffusion method and NaCl as the precipitant. To specify suitable crystallisation conditions dynamic light scattering experiments were carried out. Several combinations of additives were tried; the only monodisperse solution was achieved in the presence of both ADP and shikimate, which confirms the value of dynamic light scattering for assessing conditions for protein crystallisation (Table 5-3).

The structure of the enzyme complexed with ADP and shikimic acid was solved (Krell, unpublished data). Attempts to solve the phase problem using molecular replacement with adenylate kinase as a model failed. Instead phases were generated using multiple isomorphous replacement using a mercury and a platinum derivative. The enzyme showed an overall topology similar to adenylate kinase (Dreusicke *et al.*, 1988). There are 5 central, parallel β-strands with several helices either side. In analogy to adenylate kinase (Vonrhein *et al.*, 1995) SK contains a so called lid-domain which is folded over the bound ADP molecule. The C4 hydroxyl group of shikimic acid is co-ordinated by D34 which is part of the VDL motif (Bugg *et al.*, 1991) which confirms

the hypothesis of the authors that the conserved aspartate co-ordinates the "conserved" C4 hydroxyl group of the substrate. The structure clearly shows the core-motif as suggested by Milner-White *et al.* (1991). The conserved R139 is involved in binding shikimic acid by forming a hydrogen bond with the carboxyl group of shikimic acid, R110 lines up parallel to the adenine ring system contributing to its positioning. Further structural work will include the solution of unliganded SK structures as well as structures with substrate analogues which mimic transition states during enzyme action.

SDH was inactivated by the lysine specific reagent TNBS; ES-MS showed that clearly two lysine residues were protected from inactivation by the presence of shikimic acid and NADP (Fig. 6-1). The two lysine residues were identified by LC/ES-MS but from the analysis of the inactivation kinetics only one residue (K65) was found to be essential for enzyme function. This residue has been proposed to be involved in binding the carboxylate moiety of shikimic acid since methyl shikimate, in contrast to shikimate, did not provide protection from inactivation with TNBS (Chackrewarthy, 1995).

In general essential lysine residues have been shown to be involved in substrate binding (Valentini *et al.*, 1996), but seem to have more often a catalytic role (Kamps and Sefton, 1986; Au *et al.*, 1989; Monnaie *et al.*, 1994). In another dehydrogenase, glucose-6-phosphate dehydrogenase, an essential lysine residue has been identified using site specific chemical modification (Camardella *et al.*, 1988) for which the authors proposed a role in substrate binding. However, site-directed mutants of the enzyme in which this lysine is replaced (Bautista *et al.*, 1995) and the three dimensional structure of the enzyme (Rowland *et al.*, 1994) have identified this lysine as a catalytic residue. Although there is strong evidence that K65 of SDH is involved in substrate binding, site-directed mutagenesis and the three-dimensional structure will be needed to clarify the situation.

In three consecutive enzymes of the shikimate pathway, type I DHQ. SDH and SK lysine residues are essential for enzyme action. Interestingly, the three lysines have completely different roles: formation of a Schiff-base which functions as a electron sink stabilising a negatively charged intermediate (type I DHQ), substrate binding (SDH), and participating in the catalytic mechanism for the P-loop lysine of SK which is not thought to be primarily involved in binding nucleotide phosphates (Reinstein *et al.*, 1990).

The identification of active site residues by chemical modification was shown to be consistent with the X-ray crystallographic studies. Mass spectrometry was found to be a very powerful method to follow chemical inactivation reactions. Many of the covalent adducts generated by site-specific chemical modification are stable enough for a localisation using a combined approach of proteolytic digestion and liquid chromatography mass spectrometry. Therefore, specific chemical modification in combination with mass spectrometric techniques appears to be a very useful tool to obtain information about the chemistry of enzyme active sites.

8. References

Adams, M.J. (1987) In *Enzyme Mechanisms*, eds. Page, M.I. and Williams, A., pp 477-505, The Royal Society of Chemistry, London.

Anton, I.A. and Coggins, J.R. (1988) Sequencing and overexpression of the *Escherichia coli aroE* gene encoding shikimate dehydrogenase. *Biochem. J.* **249**, 319-326.

Aronson, B.D., Somerville, R.L., Epperly, B.R. and Dekker, E.E. (1989) The primary structure of *Escherichia coli* L-threonine dehydrogenase. *J. Biol. Chem.* **264**, 5226-5232.

Au, D.C., Masure, R. and Storm, D.R. (1989) Site-directed mutagenesis of lysine 58 in a putative ATP-binding domain of the calmodulin-sensitive adenylate cyclase from *Bordella pertussis* abolishes catalytic activity. *Biochemistry* **28**, 2772-2776.

Bachmann, B. (1983) Linkage map of *Escherichia coli* K-12, Edition 7. *Microbiol. Rev.* **47**, 180-230.

Balinsky, D. and Davies, D.D. (1961) Aromatic biosynthesis in higher plants. 1. Preparation and properties of dehydroshikimic reductase. *Biochem. J.* **80**, 292-296.

Balinsky, D. and Dennis, A.W. (1970) Shikimate dehydrogenase. *Methods Enzymol.* **17**, 354-359.

Balinsky, D., Dennis, A.W. and Cleland, W.W. (1971) Kinetic and isotope-exchange studies on shikimate dehydrogenase from *Pisum sativum*. *Biochemistry* **10**, 1947-1952.

Banerji, S., Wakefield, A.E., Allen, A.G., Maskell, D.J., Peters, S.E. and Hopkin, J.M. (1993) The cloning and characterisation of the *arom* gene of *Pneumocystis carinii*. *J. Gen. Microbiol.* **139**, 2901-2914.

Bautista, J.M., Mason, P.J. and Luzzatto, L. (1995) Human glucose-6-phosphate dehydrogenase lysine 205 is dispensable for substrate binding but essential for catalysis. *FEBS Lett.* **366**, 61-64.

Bennett, W.S. and Steitz, T.A. (1980) Structure of a complex between yeast hexokinase A and glucose. *J. Mol. Biol.* **140**, 211-230.

Beri, R.K., Grant, S., Roberts, C.F., Smith M. and Hawkins, A.R. (1990) Selective overexpression of the *QUT*E gene encoding catabolic 3-dehydroquinase in multicopy transformants of *Aspergillus nidulans*. *Biochem. J.* **265**, 337-342.

Berlyn, M.B. and Giles, N.H. (1969) Organisation of enzymes in the polyaromatic synthetic pathway: separability in bacteria. *J. Bacteriol.* **99**, 222-230.

Boam, D.J., Kelly, S.M., Price, N.C., Krell, T., Gourley, D.G., Coggins, J.R., Virden, R. and Hawkins, A.R., Spectroscopic probes reveal contrasting active site architectures in two type II dehydroquinases. Manuscript in preparation.

Bonner, C.A. and Jensen, R.A. (1994) Cloning of the bifunctional dehydroquinase-shikimate dehydrogenase of aromatic amino acid biosynthesis in *Nicotina tabacum*. *Biochem. J.* **302**, 11-14.

Borders, C.L., Jr. and Riordan, J.F. (1975) An essential arginyl residue at the nucleotide binding site of creatine kinase. *Biochemistry* **14**, 4699-4704.

Bottomley, J.R., Clayton, C.L., Chalk, P.A. and Kleanthous, C. Cloning, sequencing, expression, purification and preliminary characterisation of a type II dehydrogenase from *Helicobacter pylori*. (1996b) *Biochem. J.* **318**, in press.

Bottomley, J.R., Hawkins, A.R. and Kleanthous, C. (1996a) Conformational changes and the role of metals in the mechanism of type II dehydroquinase from *Asergillus nidulans*. *Biochem. J.* **318**, in press

Boudet, A.M., Graziana, A. and Ranjeva, R. (1985) In *The Biochemistry of Plant Phenolics,* eds. Van Sumere, C.F. and Lea, P.J., pp 135-159, Annual Proceedings of the Phytochemical Society of Europe, Clarendon Press, Oxford.

Boys, C.W.G., Bury, S.M., Sawyer, L., Charles, I.G., Hawkins, A.R., Deka, R., Kleanthous, C. and Coggins, J.R. (1992) The crystallisation of a type I 3-dehydroquinase from *Salmonella typhi*. *J. Mol. Biol.* **227**, 352-355.

Bradford, M.M. (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein dye binding *Anal. Biochem.* **72**, 248-254.

Bray, M.R. and Clarke, A.J. (1995) Identification of an essential tyrosyl residue in the binding site of *Schizophyllum commune* Xylanase A. *Biochemistry* **34**, 2006-2014.

Britton, H.G., Carreras, J. and Grisolía, S. (1972) Mechanism of yeast phosphoglycerate mutase. *Biochemistry* **11**, 3008-3014.

Bugg, T.D.H., Abell, C. and Coggins, J.R. (1988) Specificity of E. coli shikimate dehydrogenase towards analogues of 3-dehydroshikimic acid. *Tetrahedron* **29**, 6779-6782.

Bugg, T.D.H., Alefounder, P.R. and Abell, C. (1991) An amino acid sequence motif observed amongst enzymes of the shikimate pathway. *Biochem. J.* **276**, 841-845.

Butler, J.R., Alworth, W.L. and Nugent, M.J. (1974) Mechanism of dehydroquinase catalysed dehydration. I. Formation of a Schiff-base intermediate. *J.Am. Chem. Soc.* **96**, 1617-1618.

Camardella, L., Caruso, C., Rutigliano, B., Romano, M.D., DiPrisco, G. and Descalzi-Cancedda, F. (1988) Human erythrocyte glucose-6-phosphate dehydrogenase - Identification of a reactive lysysl residue labelled with pyridoxal 5'-phosphate. *Eur. J. Biochem.* **124**, 485-489.

Case, M.E. and Giles, N.H. (1971) Partial enzyme aggregates formed by pleiotropic mutants in the *arom* gene cluster of *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* **68**, 58-62.

Chackrewarthy, S. (1995) Ph.D. thesis, Studies on the active site of *E. coli* shikimate dehydrogenase, University of Glasgow.

Charles, I.G., Keyte, J.W., Brammer, W.J., Smith, M. and Hawkins, A.R. (1986) Nucleotide sequence encoding the biosynthetic dehydroquinase function of the pentafunctional *arom* locus of *Aspergillus nidulans*. *Nucleic Acids Res.* **14**, 2201-2213.

Chaudhuri, S. and Coggins, J.R.(1985) The purification of shikimate dehydrogenase from *Escherichia coli*. *Biochem. J.* **226**, 217-223.

Chaudhuri, S., Duncan, K., Graham, L.D. and Coggins, J.R. (1991) Identification of the active-site lysine residues of two biosynthetic 3-dehydroquinases. *Biochem. J* **275**, 1-6.

Chaudhuri, S., Lambert, J.M., McColl, L.A. and Coggins, J.R. (1986) Purification and characterisation of 3-dehydroquinase from *Escherichia coli*. *Biochem. J.* **239**, 699-704.

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

Coggins, J.R. and Boocock, M.R. (1986) In *Multidomain proteins*, eds. Hardie, D.R. and Coggins, J.R., pp 259-281, Elsvier, Amsterdam.

Coggins, J.R., Boocock, M.R., Chaudhuri, S., Lambert, J.M., Lumsden, J., Nimmo, G.A. and Smith, D.D.S. (1987) The AROM multifunctional enzyme from *Neurospora crassa*. *Methods Enzymol.* **142**, 325-341.

Cooper, A. and Johnson, C.M. (1994) In *Methods in Molecular Biology*, eds. Jones, C., Mulloy, B. and Thomas, A.H., pp 137-150, Hunana Press, Toronto, N.J.

Cornish, E.C., Argyropoulos, V.P., Pittard, J. and Davidson, B.E. (1986) Structure of the *Escherichia coli* K-12 regulatory gene *tyrR*- nucleotide sequence and sites of initiation of transcription and translation. *J. Biol. Chem.* **261**, 403-410.

Czupryn, M.J., McCoy, J.M. and Scoble, H.A. (1995) Structure-function relationships in human interleukin-11- Identification of regions involved in activity by chemical modification and site-directed mutagenesis. *J. Biol. Chem.* **270**, 978-985.

Da Silva, A.J.F., Whittington, H.W., Clements, J., Roberts, C. and Hawkins, A.R. (1986) Sequence analysis and transformation by the catabolic 3-dehydroquinase (*QUT*E) gene from *Aspergillus nidulans*. *Biochem. J.* **240**, 481-488.

Davies, B.D. (1955) Intermediates in amino acid biosynthesis. *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, R.G. (1994) (6S)-6-Fluoroshikimic acid, an antimicrobial agent acting on the shikimate pathway. *Antimicrob. Agents Chemoth.* **38**, 403-406.

De Feyter, R.C. and Pittard, J. (1986a) Genetic and molecular analysis of *aroL*, the gene for shikimate kinase II in *Escherichia coli* K-12. *J. Bacteriol.* 165, 226-232.

De Feyter, R.C. and Pittard, J. (1986b) Purification and properties of shikimate kinase II from *Escherichia coli* K-12. *J. Bacteriol.* 165, 331-333.

De Feyter, R.C., Davidson, B.E. and Pittard, J. (1986) Nucleotide sequences of the transcription unit containing the *aroL* and *aroM* genes from *Escherichia coli* K-12. *J. Bacteriol.* **165**,233-239.

Deka, R.K., Anton, I.A., Dunbar, B. and Coggins, J.R. (1994) The characterization of the shikimate pathway enzyme dehydroquinase from *Pisum sativum*. *FEBS Lett.* **349**, 397-402.

Deka, R.K., Kleanthous, C. and Coggins, J.R. (1992) Identification of the essential histidine residue at the active site of *Escherichia coli* dehydroquinase. *J. Biol. Chem.* **267**, 22237-22242.

Dente, L., Cesareni, G., Cortese, R. (1983) pEMBL-A new family of single stranded plasmids. *Nucl. Acid Res.* **11**, 1645-1654.

Dever, T.E., Glynias, M.J. and Merrick, W.C. (1987) GTP-binding protein - 3 consensus sequence elements with distinct spacing. *Proc. Natl. Acad. Sci. U.S.A* **84**, 1814-1818.

Dreusicke, D., Karplus, A. and Schulz, G.E. (1988) Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution. *J. Mol. Biol.* **199**, 359-371.

Dumas, D.P. and Raushel, F.M. (1990) Chemical and kinetic evidence for an essential histidine in the phophodiesterase from *Pseudomonas diminuta*. *J. Biol. Chem.* **265**, 21498-21503.

Duncan, K., Edwards, R.M. and Coggins, J.R. (1987) The pentafunctional AROM enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. *Biochem. J.* **246**, 375-386.

Ely, R. and Pittard, J. (1979) Aromatic amino acid biosynthesis: Regulation of shikimate kinase in *Escherichia coli* K-12. *J. Bacteriol.* **138**, 933-943.

Eum, H.M. and Miles, E.W. (1984) Reaction of phenylglyoxal and (p-Hydroxyphenyl) glyoxal with arginines and cysteines in the α subunit of tryptophan synthase. *Biochemistry* **23**, 6484-6491.

Euverink, G.J.W., Hessels, G.I., Vrijbloed, J.W., Coggins, J.R. and Dijkhuizien, L. (1992) Purification and characterization of a dual function 3-dehydroquinate dehydratase from *Amycolatopsis methanolica*. *J. Gen. Microbiol.* **138**, 2449-2457.

Faller, L.D., Baroudy, B.M., Johnson, A.M. and Ewall, R.X. (1977) Magnesium ion requirements for yeast enolase activity. *Biochemistry* **16**, 3864-3869.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1989) Electrospray ionization for mass-spectrometry of large biomolecules. *Science* **246**, 64-71.

Ferrige, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A. and Skilling, J. (1992) The application of MaxEnt to high-resolution mass-spectrometry *Rapid Commun. Mass Spectrom.* **6**, 707-711.

Fodor, G. and Kovacs, Ö. (1949) A synthesis of adrenalin-like compounds. *J. Amer. Chem. Soc.* **71**, 1045-1048.

Fothergill-Gilmore, L.A. and Watson, H.C. (1989) The phosphoglycerate mutases. *Adv. Enzymol. Relat. Areas Mol. Biol.* **62**, 227-313.

Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1986) ATP-binding site of adenylate kinase: Mechanistic implications of its homologue with *ras*-encoded p21, F_1-ATPase, and other nucleotide-binding proteins. *Proc. Natl. Acad. Sci. USA* **83**, 907-911.

Gadda, G., Negri, A. and Pilone, M.S. (1994) Reaction of phenylglyoxal with arginine groups in D-amino-acid oxidase from *Rhodotorula gracilis*. *J. Biol. Chem.* **269**, 17809-17814.

Gaertner, F.H. and Cole, K.W. (1977) A cluster gene: Evidence for one gene, one polypeptide, five enzymes. *Biochem. Biophys. Res. Commun.* **75**, 259-264.

Ganem, B., Li, Y.T. and Henion, J.D. (1991) Detection of noncovalent receptor-ligand complex by mass spectrometry. *J. Am. Chem. Soc.* **113**, 6294-6296.

Garbe, T., Servos, S., Hawkins, A.R., Dimitriadis, G., Young, D., Dougan, G. and Charles, I.G. (1991) The *Mycobacterium tuberculosis* shikimate pathway genes: Evolutionary relationship between biosynthetic and catabolic 3-dehydroquinases. *Mol. Gen. Genet.* **228**, 385-392.

Geever, R.F., Huit, L., Baum, J.A., Tyler, B.M., Patel, V.B., Rutledge, B.J., Case, M.E. and Giles, N.H. (1989) DNA-sequence, organization and regulation of the QA gene-cluster of *Neurospora crassa*. *J. Mol. Biol.* **207**, 15-34.

Gilbert, W. (1991), unpublished , obtained from VecBase 3.0, copyright IBI.

Giles, N.H., Case, M.E., Baum, J., Geever, R., Huiet, L., Patel, V. and Tyler, B. (1985) Gene organisation and regulation in the QA (quinic acid) cluster of *Neurospora crassa*. *Microbiol. Reviews* **49**, 338-358.

Giles, N.H., Case, M.E., Partridge, C.W.H. and Ahmed, S.I. (1967) A gene cluster in *Neurospora crassa* coding for an aggregate of five aromatic synthetic enzymes. *Proc. Natl. Acad. Sci. USA* **58**, 1453-1460.

Gill, C.S. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Gillies, F.M. (1994) Studies on the AROM pentafunctional enzyme in *Saccharomyces cerevisiae* and *Aspergillus nidulans*. PhD thesis, University of Glasgow.

Goldfarb, A.R. (1966) A kinetic study of the reactions of amino acids and peptides with trinitrobenzenesulfonic acid. *Biochemistry* **5**, 2570-2574.

Goldin, B.R. and Frieden, C. (1971) Effect of trinitrophenylation of specific lysyl residues on the catalytic, regulatory, and molecular properties of bovine liver glutamate dehydrogenase. *Biochemistry* **10**, 3527-3534.

Gourley, D.G. (1996) Ph.D. thesis, Crystallographic studies on *Mycobacterium tyberculosis* type II dehydroquinase. University of Glasgow.

Gourley, D.G., Coggins, J.R., Isaacs, N.W., Moore, J.D., Charles, I.G. and Hawkins, A.R. (1994) Crystallization of a type-II dehydroquinase from *Mycobacterium-tuberculosis*. *J. Mol. Biol.* **241**, 488-491.

Graham, L.D., Edwards, R.M. and Coggins, J.R. (1993) Overexpression of the yeast multifunctional AROM protein. *Biochim. Biophys. Acta* **1216**, 417-424.

Grazi, E., Cheng, T. and Horecker, B.L. (1962) The formation of a stable aldolase-dihydroxyacetone phosphate complex. *Biochem. Biophys. Res. Commun.* **7**, 250-253.

Grewe, R. and Haendler, H. (1966) 5-Dehydroquinic acid. *Biochem. Preparations* **2**, 21-26.

Hanson, K.R. and Rose, I.A. (1963) The absolute stereochemical course of citric acid biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **50**, 981-988.

Harris, J., Kleanthous, C., Coggins, J.R., Hawkins, A.R. and Abell, C. (1993) Different mechanistic and stereochemical courses for the reactions catalyzed by type I and type II dehydroquinases. J. *Chem. Soc., Chem. Commun.* 1080-1081.

Harris, J.M., González-Bello, C., Kleanthous, C., Hawkins, A.R., Coggins, J.R. and Abell, C. (1996) Evidence from kinetic isotope studies for an enole intermediate in the mechanism of type II dehydroquinases. *Biochem. J.* (in press)

Haslam, E. (1974) *The Shikimate pathway*, Butterwoth, London.

Haslam, E. (1993) *Shikimic acid: Metabolism and Metabolites*. John Wiley & sons, Chichester, England.

Hautala, J. A., Jacobsen, J.W., Case, M.E. and Giles, N.H. (1975) Purification and characterization of catabolic dehydroquinase an enzyme of the inducible quinic acid catabolic pathway of *Neurospora crassa*. *J. Biol. Chem.* **250**, 6008-6014.

Hawkins, A.R. (1987) The complex *arom* locus of *Aspergillus nidulans*. Evidence for multiple gene fusions and convergent evolution *Curr Genet Curr. Genet.* **11**, 491-498.

Hawkins, A.R. and Lamb, H.K. (1995) The molecular biology of multidomain proteins - selected examples. *Eur. J. Biochem.* **232**, 7-18.

Hawkins, A.R. and Smith, M. (1991) Domain structure and interaction within the pentafunctional AROM polypeptide. *Eur. J. Biochem.* **196**, 717-724.

Hawkins, A.R., Giles, N.H. and Kinghorn, J.R. (1982) Genetical and biochemical aspects of quinate breakdown in the filamentous fungus *Aspergillus nidulans*. *Biochem. Genet.* **20**, 271-286.

Hawkins, A.R., Moore, J.D. and Adeokun, A.M. (1993) Characterization of the 3-dehydroquinase domain of the pentafunctional AROM protein, and the quinate dehydrogenase from *Aspergillus nidulans*, and the overproduction of the type II 3-dehydroquinase from *Neurospora crassa*. *Biochem. J.* **296**, 451-457.

Heatwole, V.M. and Sommerville, R.L. (1992) Synergism between the *trp* repressor and *tyr* repressor in repression of the *aroL* promoter of *Escherichia coli* K-12. *J. Bacteriol.* **174**, 331-335.

Hupe, D.J. (1987) In *Enzyme Mechanisms*, eds. Page, M.I. and Williams, A., pp 317-344, The Royal Society of Chemistry, London.

Izard, T., Lawrence, M.C., Malby, R.L., Lilley, G.G. and Colman, P.M. (1994) The 3-dimensional structure of N-acetylneuraminate lyase from *Escherichia coli*. *Structure* **2**, 361-369.

Jensen, R.A. (1986) The shikimate arogenate pathway - link between carbohydrate metabolism and secondary metabolism. *Physiol. Plant.* **66**, 164-168.

Johnson, S.C., Bailey, T., Becker, R.R. and Cardenas, J.M. (1979) Isolation and Sequence of a peptide in or near the active site of bovine muscle pyruvate kinase, *Biochem. Biophys. Res.Comm.* **90**, 525-530.

Jude, D.A., Ewart, C.D.C., Thain, J.L., Davies, G.M. and Nichols, W.W. (1996) Transport of the antibacterial agent (6S)-6-fluoroshikimate and other shikimate analogues by the shikimate transport system of *Escherichia coli*. *Biochim. Biophys. Acta* **1279**, 125-129.

Kadonaga, J.T. and Knowles, J.R. (1983) Role of mono-valent and divalent metal cations in the catalysis by yeast aldolase. *Biochemistry* **22**, 130-136.

Kamps, M.P. and Sefton, B.M. (1986) Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of $p60^{src}$. *Mol. Cell. Biol.* **6**, 751-757.

Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10 000 Daltons. *Anal. Chem.* **60**, 2299-2301.

Katta, V. and Chait, B.T. (1991) Observation of the heme-globin complex in native myoglobin by electrospray ionization mass spectrometry. *J. Am. Chem. Soc.* **113**, 8534-8535.

Katta, V. and Chait, B.T. (1993) Hydrogen/deuterium exchange electrospray ionization mass spectrometry: A method for probing protein conformational changes in solution. *J. Am. Chem. Soc.* **115**, 6317-6321.

Kleanthous, C. and Coggins, J.R. (1990) Reversible Alkylation of active site methionine residues in dehydroquinase. *J. Biol. Chem.* **265**, 10935-10939.

Kleanthous, C., Deka, R., Davis, K., Kelly, S.M., Cooper, A., Harding, S.E., Price, N.C., Hawkins, A.R. and Coggins, J.R. (1992) A comparison of the enzymological and biophysical properties of two distinct classes of dehydroquinase enzymes. *Biochem. J.* **282**, 687-695.

Komatsutakaki, M. (1995) Effects of energization and substrates on the reactivities of lysine residues of the chloroplast ATP synthase beta-subunit, *Eur. J. Biochem.* **228**, 265-270.

Koshiba, T. (1979) Organization of enzymes in the shikimate pathway of *Phaseolous mungo* seedlings. *Plant Cell Physiol.* **20**, 667-670.

Kubiseski, T.J., Green, N.C., Borhani, D.W. and Flynn, T.G. (1994) Studies on pig aldolase reductase - Identification of an essential arginine in the primary and tertiary structure of the enzyme. *J. Biol. Chem.* **269**, 2183-2188.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of bacteriophage T4. *Nature* **227**, 680-685.

Lai, C.Y and Horecker, B.L. (1972) Aldolase: A model for enzyme structure-function relationships. *Essays in Biochemistry* **8**, 149-178.

Lamb, H.K., Moore, J.D., Lakey, J.H., Levett, L. Wheeler, K.A., Lago, H., Coggins, J.R. and Hawkins, A.R. (1996) Comparative analysis of the *QUT*R transcription repressor protein and the 3 C-terminal domains of the pentafunctional AROM enzyme. *Biochem. J.* **313**, 941-950.

Lebioda, L., Stec, B. and Brewer, J.M. (1989) The structure of yeast enolase at 2.25 Å resolution - an 8-fold beta+alpha-barrel with a novel beta-beta-alpha-alpha (beta-alpha)$_6$ topology. *J. Biol. Chem.* **264**, 3685-3693.

Leech, A.P., James, R., Coggins, J.R. and Kleanthous, C. (1995) Mutagenesis of active-site residues in type I dehydroquinase from *Escherichia coli* - stalled catalysis in a histidine to alanine mutant. *J. Biol. Chem.* **270**, 25827-25836.

Li, Y.T., Hsieh, Y.L., Henion, J.D., Senko, M.W., McLafferty, F.W. and Ganem, B. (1993) Mass spectrometric studies on noncovalent dimers of leucine zipper peptides. *J. Am. Chem. Soc.* **115**, 8409-8413.

Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J. and Slonimski, P.P. (1989) Birth of the D-E-A-D box. *Nature* **337**, 121-122.

Løbner-Olsen, A. and Marinus, M.G. (1992) Identification of the gene (*aroK*) encoding shikimic acid kinase I of *Escherichia coli*. *J. Bacteriol.* **174**, 525-529.

Lumsden, J. and Coggins, J.R. (1977) The subunit structure of the AROM multienzyme complex of *Neurospora crassa*: a possible pentafunctional polypeptide chain. *Biochem. J.* **161**, 599-607.

Lundblad, R.L. and Noyes, C.M. (1984) *Chemical Reagents for Protein Modification*, CRC Press, Boca Ratón, Florida.

Mann, M. and Wilm, M. (1995) Electrospray mass spectrometry for protein characterisation. *TIBS* **20**, 219-224.

Matsuo, Y. and Nishikawa, K. (1994) Protein structural similarities predicted by a sequence structure compatibility method. *Protein Science* **3**, 2055-2063.

Means, G.E., Congdon, W.I. and Bender, M.L. (1972) Reactions of 2,4,6-trinitrobenzenesulfonate ion with amines and hydroxide ions. *Biochemistry* **11**, 3564-3571.

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) Strategies for nonradioactive methods in the localisation of phosphorylated amino acids in proteins. *FASEB* 7, 776-782.

Miles, E.W. (1977) Modification of histidyl residues in proteins by diethylpyrocarbonate. *Methods Enzymol.* **47**, 431-442.

Millar, G., Lewendon, A., Hunter, M.G. and Coggins, J.R. (1986) The cloning and expression of the *aroL* gene from *Escherichia coli* K-12. *Biochem. J.* **237**, 427-437.

Milner-White, J.E., Coggins, J.R. and Anton, I.A. (1991) Evidence for an ancestral core structure in nucleotide-binding proteins with the type A motif. *J. Mol. Biol.* **221**, 751-754.

Minton, N.P., Whitehead, P.J., Atkinson, T. and Gilbert, H.J.(1989) Nucleotide sequence of an *Erwinia chrysanthemi* gene encoding shikimate kinase. *Nucleic Acids Res.* **17**, 1769.

Moffatt, B.A. and Studier, F.W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**, 221-227.

Mohan, P.M., Basu, A., Basu, S., Abraham, K.I., and Modak, M.J. (1988) DNA binding domain of *Escherichia coli* DNA polymerase I: Identification of arginine-841 as an essential residue. *Biochemistry* **27**, 226-233.

Moller, W. and Amons, R. (1985) Phosphate-binding sequences in nucleotide-binding proteins. *FEBS Lett.* **186**, 1-7.

Monnaie, D., Dubus, A., Cooke, D., Marchand-Brynaert, J., Normark, S. and Frere, J.M. (1994) Role of residue Lys 315 in the mechanism of action of the *Enterobacter cloacae* 908R β-Lactamase. *Biochemistry* **33**, 5193-5201.

Moore, J.D. and Hawkins, A.R. (1993) Overproduction of, and interaction within, bifunctional domains from the amino-termini and carboxy-termini of the pentafunctional AROM protein of *Aspergillus nidulans*. *Mol. Gen. Genet.* **240**, 92-102.

Moore, J.D., Coggins, J.R., Virden, R. and Hawkins, A.R. (1994) Efficient independent activity of a monomeric, monofunctional dehydroquinate synthase derived from the N-terminus of the pentafunctional AROM protein of *Aspergillus-nidulans*. *Biochem. J.* **301**, 297-304.

Moore, J.D., Hawkins, A.R., Charles, I.G., Deka, R., Coggins, JR., Cooper, A., Kelly, S.M. and Price, N.C. (1993) Characterization of the type I dehydroquinase from *Salmonella typhi*. *Biochem. J.* **295**,277-285.

Mousdale, D.M. and Coggins, J.R. (1985) Subcellular localisation of the common shikimate pathway enzymes in *Pisum sativum L. Planta* **163**, 241-249.

Mousdale, D.M., Campbell, M.S. and Coggins, J.R. (1987) Purification and characterisation of bifunctional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings. *Phytochemistry* **26**, 2665-2670.

Mueller, C. W. and Schulz G. E. (1992) Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap5A refined at 1.9 Å resolution. *J. Mol. Biol.* **224**, 159-177.

Mueller, M.J., Samuelsson, B. and Haeggström, J.Z. (1995) Chemical modification of Leukotriene A(4) hydroxylase - Indications for essential tyrosyl and arginyl residues at the active site. *Biochemistry* **34**, 3536-3543.

Nairn, J., Price, N.C., Fothergill-Gilmore, L.A., Walker, G.E., Fothergill, J.E. and Dunbar, B. (1994) The amino acid sequence from the small monomeric phosphoglycerate mutase from the fission yeast *Schizosaccharomyces pombe*. *Biochem. J.* **297**, 603-608.

Neale, A.D., Scopes, R.K., Kelly, J.M. and Wettenhall, R.E.H. (1986) The 2 alcohol dehydrogenases of *Zymomonas mobilis* - purification by differential dye ligand chromatography, molecular characterisation and physiological roles. *Eur. J. Biochem.* **154**, 119-124.

Pai, E.F., Kabsch, W., Krengel, U., Holmes, K.C., John, J. and Wittinghofer, A. (1989) Structure of the guanine nucleotide binding domain of the Ha-Ras oncogene product p21 in the triphosphate confirmation. *Nature* **341**, 209-214.

Patel, V.B. and Giles, N.H. (1979) Purification of the *arom* multienzyme aggregate from *Euglena gracilis*. *Biochim. Biophys. Acta.* **567**, 24–34.

Patthy, L. and Thesz, J. (1979) Origin of the selectivity of α-dicarbonyl reagents for arginyl residues of anion binding sites. *Eur. J. Biochem.* **105**, 387-393.

Pitt, A. (1996) In *Proteins Labfax*, ed. Price, N.C., pp 174-183, Academic Press, San Diego and BIOS Scientific Publishers Ltd., Oxford.

Pittard, J. (1987) In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*. eds. Neidhart, F.C., Ingraham, J., Low, K.B., Magasanik, M., Schaechter, M. and Umbarger, H.E., pp 368-394, American Society for Microbiology, Washington, D.C.

Plapp, B.V., Zeppenzauer, E. and Brändén, C.I. (1978) Crystallization of liver alcohol dehydrogenase by the modification of amino groups. *J. Mol. Biol.* **119**, 451-453.

Polikarpov, I., Shrive, A.K., Coggins, J.R., Hawkins, A.R. and Sawyer, L. Three dimensional structure of type I dehydroquinase from *Salmonella typhi*. Manuscript submitted to Nature.

Polley, L.D. (1978) Purification and characterization of 3-dehydroquinate hydrolase and shikimate oxidoreductase - Evidence for a bifunctional enzyme. *Biochim. Biophys. Acta.* **526**, 259-266.

Provencher, S. W. and Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-37.

Qamar, S., Marsh, K. and Berry, A. (1996) Identification of arginine 331 as an important active site residue in the class II fructose-1,6- bisphosphate aldolase of *Escherichia coli*. *Protein Science* **5**, 154-161.

Ray, W.J., Jr. and Koshland, D.E., Jr. (1961) A method for characterising the type and numbers of groups involved in enzyme action. *J. Biol. Chem.* **236**, 1973-1979.

Ray, W.J., Jr. and Koshland, D.E., Jr. (1962) Identification of amino acids involved in phosphoglucomutase action. *J. Biol. Chem.* **237**, 2493-2505.

Reinstein, J., Schlichting, I. and Wittinghofer, A. (1990) Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of *Escherichia coli*. *Biochemistry* **29**, 7451-7459.

Riordan, J.F. (1979) Arginyl residues and anion binding sites in proteins. *Mol. and Cell. Biochem.* **26**, 71-92.

Riordan, J.F., McElvany, K.D. and Borders, C.L., Jr. (1977) Arginyl residues: Anion recognition sites in enzymes. *Science* **195**, 884-885.

Robertson, D.L., Hilton, S., Wong K.R., Koepke, A. and Buckley, J.T. (1994) Influence of active site and tyrosine modification on the secretion and activity of the *Aeromonas hydrophila* Lipase/acyltransferase. *J. Biol. Chem.* **269**, 2146-2150.

Robinson, C.V., Gross, M., Eyles, S.J., Ewbank, J.J., Mayhew, M., Hartl, F.U., Dobson, C.M. and Radford, S.E. (1994) Conformation of GroEL-bound α-lactalbumin probed by mass spectrometry. *Nature* **372**, 646-651.

Rose, Z.B. (1970) Evidence for a phosphohistidine protein intermediate in the phosphoglycerate mutase reaction. *Arch. Biochem. Biophys.* **140**, 508-513.

Rose, Z.B., Hamasaki, N. and Dube, S. J. (1975) The sequence of a peptide containing the active site phosphohistidine residue of phosphoglycerate mutase from chicken breast muscle. *J. Biol. Chem.* **250**, 7939-7942.

Rowland, P., Basak, A.K., Gover, S., Levy, H.R. and Adams, M.J. (1994) The 3-dimensional structure of glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* refined at 2.0 Angstrom resolution. *Structure* **2**, 1073-1087.

Sanderson, K.E. and Roth, J.R. (1983) Linkage map of *Salmonella typhimurium*. *Microbiol. Rev.* **47**, 410-553.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.H., and Roe, B.A. (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing *J. Mol. Biol.* **143**, 161-178

Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) The P-loop - a common motif in ATP- and GTP- binding proteins. *Trends Biochem. Sci.* **15**, 430-434.

Schindler, P.A., Settineri, C.A., Collet, X., Fielding, C.J. and Burlingame, A.L. (1995) Site-specific detection and structural characterisation of the glycosylation of human plasma proteins lecithin: cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion. *Protein Science* **4**, 791-803.

Schmid, J., Schaller, A., Leibinger, U., Boll, W. and Amrhein, N. (1992) The *in-vitro* synthesized tomato shikimate kinase precursor is enzymatically active and is imported and processed to the mature enzyme by chloroplasts. *The Plant Journal* **2**, 375-383.

Schmidt, C.L., Danneel, H.-J., Schultz, G. and Buchanan, B.B. (1990) Shikimate kinase from spinach chloroplasts - purification, characterisation, and regulatory function in aromatic amino acid biosynthesis. *Plant Physiol.* **93**, 759-766.

Schulz, G.E., Müller, C.W. and Diedrichs, K. (1990) Induced-fit movements in adenylate kinase. *J. Mol. Biol.* **213**, 627-630.

Schulze-Siebert, D. Heineke, D., Scharf, H. and Schultz, G. (1984) Pyruvate-derived amino acids in spinach chloroplasts - Synthesis and regulation during photosynthetic carbon metabolism. *Plant. Physiol.* **76**, 465-471.

Servos, S., Chatfield, S., Hone, D., Levine, M., Dimitriadis, G., Pickard, D., Dougan, G., Fairweather, N. and Charles, I. (1991) Molecular cloning and characterisation of the *aroD* gene encoding 3-dehydroquinase from *Salmonella typhi*. *J. Gen. Microbiol.* **137**, 147-152.

Singer, S.J. (1967) Covalent labelling of active sites. *Adv. Prot. Chem.* **22**, 1-54.

Siuzdak, G. (1996) *Mass Spectrometry for Biotechnology*, Academic Press, San Diego.

Siuzdak, G., Krebs, J.F., Benkovic, S.J. and Dyson, H.J. (1994) Binding of hapten to a single-chain catalytic antibody demonstrated by electrospray mass spectrometry. *J. Am. Chem. Soc.* **116**, 7937-7938.

Smith, D.D.S. and Coggins, J.R. (1983) Limited proteolysis of the AROM enzyme complex of *N. crassa*: evidence for the isolation of a bifunctional domain of the pentafunctional polypeptide. *Biochem. J.* **213**, 405-415.

Sokolovsky, M., Harell, D. and Riordan, J.F. (1969) Reaction of tetranitromethane with sulfhydryl groups in proteins. *Biochemistry* **8**, 4740-4745.

Sokolovsky, M., Riordan, J.F. and Vallee, B.L. (1966) Tetranitromethane: A reagent for the nitration of tyrosyl residues in proteins. *Biochemistry* **5**, 3582-3589.

Sprinson, D.B. (1961) The synthesis of aromatic compounds from D-glucose. *Adv. Carbohydrate Chem.* **15**, 235-252.

Steinruecken, H.C. and Amrhein, N. (1980) The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid 3-phosphate synthase. *Biochem. Biophys. Res. Commun.* **94**, 1207-1212.

Sygusch, J., Beaudry, D. and Allaire, M. (1987) Molecular architecture of rabbit skeletal muscle aldolase at 2.7 Å resolution. *Proc. Natl. Acad. Sci. USA* **84**, 7846-7850.

Takahashi, K. (1968) The reaction of phenylglyoxal with arginine residues in proteins. *J. Biol. Chem.* **243**, 6171-6179.

Taniguchi, H., Suzuki, M., Manenti, S. and Titani, K. (1994) A mass spectrometric study on the *in vivo* posttranslational modification of GAP-43. *J. Biol. Chem.* **269**, 22481-22484.

Tedeschi, G., Negri, A., Ceciliani, F., Biondi, P.A., Secchi, C. and Ronchi, S. (1992) Chemical modification of functional arginyl residues in beef kidney D-aspartate oxidase. *Eur. J. Biochem.* **205**, 127-132.

Turner, M.J., Smith, B.W. and Haslam, E. (1975) The shikimate pathway. Part IV. The stereochemistry of the 3-dehydroquinate dehydratase reaction and observations on 3-dehydroquinate synthetase. *J. Chem. Soc. Perkin Trans.* 52-55.

Urushibara, N., Kumazaki, T. and Ishii, S. (1992) Hemoglobin-binding sites on human haptoglobin - identification of lysyl residues participating in binding. *J Biol. Chem.* **267**, 13413-13417.

Valentini, G., De Gregorio, A., Di Salvo, C., Grimm, R., Bellocco, E., Cuzzocrea, G. and Iadarola, P. (1996) An essential lysine in the substrate-binding site of ornithine carbamoyltransferase. *Eur. J. Biochem.* **239**, 397-402.

Vinella, D., Gagny, B., Joseleau-Petit, D., D'Ardi, R. and Cashel, M. (1996) Mecillinam resistance in *Escherichia coli* is conferred by loss of a second activity of the *aroK* protein. *J. Bacteriol.* **178**, 3818-3828.

Vonrhein, C., Schlauderer, G.J. and Schulz, G.E. (1995) Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure* **3**, 483-490.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the α-and β- subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Weiss, U. and Edwards, J.M. (1980) *The biosynthesis of Aromatic compounds*, John Wiley and Sons, New York..

Weiss, U. and Mingioli, E.S. (1956) Aromatic biosynthesis XV. The isolation and identification of shikimic acid 5-phosphate. *J. Am. Chem. Soc.* **78**, 2894-2898.

Werber, M.M., Moldovan, M. and Sokolovsky, M. (1975) Modification of arginyl residues in porcine carboxypeptidase B. *Eur. J. Biochem.* **53**, 207-216.

Whipp, M.J. and Pittard, A.J. (1995) A reassessment of the relationship between *aroK*- and *aroL*- encoded shikimate kinase enzymes of *Escherichia coli*. *J. Bacteriol.* **177**, 1627-1629.

White, M.F. and Fothergill-Gilmore, L.A. (1988) Sequence of the gene encoding phosphoglycerate mutase from *Saccharomyces cerevisiae*. *FEBS Lett.* **229**, 383-387.

White, M.F., Fothergill-Gilmore, L.A., Kelly, S.M. and Price, N.C. (1993) Dissociation of the tetrameric phosphoglycerate mutase from yeast by a mutation in the subunit contact region. *Biochem. J.* **291**, 479-483.

White, P, Young, J., Hunter, I.S., Nimmo, H.G. and Coggins, J.R. (1990) The purification and characterisation of 3-dehydroquinase from *Streptomyces coelicolor*. *Biochem. J.* **265**, 735-738.

Wiseman, T., Williston, S., Brandts, J.F. and Lin, L-N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131-137.

GLASGOW UNIVERSITY LIBRARY