The second enzyme of Histidine Biosynthesis from *Arabidopsis thaliana*

A thesis submitted to the
University of Glasgow

for the Degree of
DOCTOR OF PHILOSOPHY

IN
BIOCHEMISTRY AND MOLECULAR BIOLOGY

BY
SAMANTHA ALLISON CAMPBELL

Division of Biochemistry and Molecular Biology
University of Glasgow
Glasgow, G12 8QQ
SCOTLAND

NOVEMBER, 1999

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

For My Parents

Summary

The original objectives of the project were to clone the cDNAs for the first three enzymes of histidine biosynthesis from *Arabidopsis* and subsequently overexpress and characterise at least one of the enzymes. The cDNA for the second enzyme of the pathway, the bifunctional HisIE was the only cDNA for which viable clones were obtained. The cDNA for the *Arabidopsis* hisIE clone contained a chloroplastic targeting sequence which was removed in order to allow expression of the active domains of the HisIE protein. A construct which allowed overexpression of the HisIE protein to approximately 30% total cell protein was generated and the protein was subsequently purified to homogeneity using three chromatographic steps (ion exchange on DEAE, Phenyl Sepharose and gel filtration). The substrate PR-ATP was generated by biotransformation from ATP and PRPP using a purified HisG extract from *E. coli* and was purified by a new method using ion exchange chromatography and a volatile buffer system. This gave the first homogeneous sample of PR-ATP and allowed the kinetic characterisation of the enzyme. Optimal assay conditions for the *Arabidopsis* HisIE enzyme were established and basic kinetic parameters including the K_m and k_{cat} of the purified enzyme were determined. Chemical modification experiments with the histidine modifying reagent DEPC resulted in the rapid loss of enzyme activity even at very low concentrations. This suggested that a histidine residue might be involved in enzyme activity. To test this hypothesis site directed mutagenesis of the two conserved histidine residues within the protein was carried out and the mutant proteins overexpressed and purified. Subsequent characterisation of the mutant proteins and comparison of the K_m and k_{cat} with the wild type values indicated that neither of these histidine residues was catalytically important but suggested that one of the residues (H146) might be important for the quaternary structure of the HisIE enzyme and the other (H245) might sterically hinder the active site.

Acknowledgements

I would like to thank Professor J.R Coggins for his support and guidance during the course of this work and for the opportunity to be a part of his group. I would also like to thank Dr D. Cole, Rhone Poulenc Agriculture Limited and the BBSRC who sponsored this project as a CASE award.

Thanks are also due to Dr S. Ali, Dr B. McGinn, Dr A. Lapthorn , Ms R. MacFie, Mr B. Lokhamp, and all other members of the Molecular Enzymology Group for their help and advice during the course of this project.

Thanks to Dr D. Rycroft, University of Glasgow for carrying out the NMR analysis and to Dr K. Lilley, University of Leicester for carrying out Mass Spectrometry analysis.

Special thanks to Mr J. Greene for his technical support and good humour, both of which were essential to this project, and also to Mrs J. Greenwood for secretarial assistance.

Finally I would like to thank my family and friends for their love and support over the past three years.

Table of Contents

CHAPTER 2 MATERIALS AND METHODS

CHAPTER 4 PURIFICATION OF THE HISIE ENZYME

CHAPTER 6 CHARACTERISATION OF THE HISIE ENZYME

List of Figures

List of Tables

Abbreviations

A	absorbance
AICAR	5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside
ADP	adenosine diphosphate
ALS	acetolactate synthase
amp	ampicillin
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BSA	bovine serum albumin
CIAP	calf intestinal alkaline phosphatase
(k)Da	(kilo) Dalton
dNTP	deoxyribonucleotide triphosphate
DEAE	diethylaminoethyl
DEPC	diethylpyrocarbonate
DNAse	deoxyribonuclease
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetate
$[E_T]$	enzyme concentration
EPSPS	5 enol pyruvyl shikimic 3-phosphate synthase
FPLC	fast protein liquid chromatography
hisX	histidine mutant or gene
hisX	cDNA or clone for a histidine enzyme
HisX	enzyme or protein
HP	histidinol phosphate
HOL	L-histidinol
HPLC	high pressure liquid chromatography
IAP	imidazole acetole phosphate
IGP	imidazole glycerol phosphate
IGPD	imidazole glycerol phosphate dehydratase
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilo base pairs

k_{cat}	turnover number
K_i	dissociation constant of enzyme inhibitor complex
K_m	Michaelis constant
k_{obs}	first order rate constant
K_s	dissociation constant of enzyme substrate complex
LB	Luria Bertani medium
NMR	nuclear magnetic resonance
ODS	octadecylsilane
PAGE	polyacrylamide gel-electrophoresis
PCR	polymerase chain reaction
PHD	profile fed neural network system from Heidelberg
PR-AMP	phosphoribosyl AMP
PR-ATP	phosphoribosyl ATP
PRFAR	N'-[(5'-phosphoribulosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide
PRPP	phosphoribosyl pyrophosphate
Pro-FAR	N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide
R-5-P	ribose 5 phosphate
RNAse	ribonuclease
rpm	revolution per minute
S-200	Sephacel S-200
SDS	sodium dodecyl sulphate
$t_{1/2}$	half life
TAE	Tris-acetate/ EDTA
TBE	Tris-borate/ EDTA buffer
TE	Tris-acetate/ EDTA buffer
TEMED	N, N, N, N'-tetramethyl-ethylenediamine
TFB1	transformation buffer 1
TFB2	transformation buffer 2
TNBS	trinitrobenzene sulphonic acid
Tris	Tris (hydroxymethyl) aminomethane

tRNA	transfer RNA
U	units of enzyme activity
UV	ultra violet
V	volts
V_{max}	maximum velocity
λYES	Yeast *E. coli* shuttle vector

Chapter 1

Introduction

1.1 General Introduction

1.1.1 Why study amino acid biosynthesis?

As modern farming techniques concentrate on gaining the maximum yield from the smallest possible area, the use of herbicides is constantly on the increase. The herbicide industry is now a multi-million pound industry and many more millions are being pumped into herbicide research in the quest for a more effective, more efficient and cheaper alternative. An ideal herbicide should be effective at killing the weed but not the crop plant, should have low application rates, be environmentally friendly and be relatively non toxic to wildlife etc. (Cole, 1994). It should also be relatively inexpensive and there should also be a very low tendency for resistance to emerge.

In recent years the amino acid biosynthetic pathways which are present in plants, bacteria and fungi, have been investigated not only because of the intrinsic scientific interest in how these key biological building blocks are synthesised but also with a view to developing inhibitors of these pathways as potential antibacterial or herbicidal agents (Mousdale and Coggins, 1991). The pathway intermediates and the functional enzyme domains are highly conserved between bacteria and plants. Two amino acid biosynthetic pathways, the shikimate pathway and the pathway leading to the branched chain amino acids, have already proved to be excellent targets for herbicides (La Rossa and Falco, 1984, Kishore and Shah, 1988).

The shikimate pathway which is used by both plants and bacteria to generate the aromatic compounds including the amino acids tyrosine, phenylalanine and tryptophan has proved useful as a target for herbicidal action. The herbicide Glyphosate™ (N-phosphonomethyl glycine) is one such inhibitor of this pathway (Mousdale and Coggins, 1991). It has been available for over twenty years but it was only in 1980 that the site of action was determined as 5-enol-pyruvyl shikimic acid 3-

phosphate synthase (EPSPS; EC 2.5.1.19), which is the sixth enzyme of this pathway (Steinrucken and Amrhein, 1980).

Acetolactate synthase (ALS EC 4.1.3.18) a key enzyme in the synthesis of the branched chain amino acids, leucine, isoleucine and valine, is the target for the sulphonylurea (LaRossa and Schloss, 1984), imidazolinone (Shanner *et al.*,1984) and triazolopyrimidine sulfonamides (Kleschick and Gerwick, 1989). The sulphonylureas and the imidazolinones appear to interact rather differently with ALS which has proved to be a very useful target for herbicide action (Pillmoor *et al.*,1995).

The underlying rationale behind the work described in this thesis was to investigate whether the early enzymes of histidine biosynthesis might also be a good target for herbicide development. There were indications that this pathway might be a good target since the compound Amitrole™, which inhibits Imidazoleglycerol phosphate dehydratase (IGPD), the sixth enzyme of the histidine biosynthetic pathway, (Hilton *et al.*, 1965) has herbicidal effects. More recently some triazole phosphonates have also been shown to inhibit IGPD; these compounds, which have a similar potency to Glyphosate™, may be useful as herbicides (Cox *et al.*, 1997).

Certain *Brassica* species such as the metal hyperaccumulator *Alyssum* have been found to accumulate heavy metals like nickel and zinc and appear to be tolerant of such metals. These plants have been found to contain increased levels of free histidine within the cytoplasm (Krämer *et al.*, 1996). It may be possible to use plants which have the potential to produce more free histidine in order to remove heavy metals from contaminated land.

Because of its potential as a target for herbicides and because of the interest in overproducing histidine as a means of developing metal tolerant plants, a detailed study

of the enzymes of the pathway and their regulation in plants was a worthwhile objective.

1.1.2 Histidine Biosynthesis- General introduction

To date the histidine pathway has been studied well in *Escherichia coli* (Winkler, 1988) and in several other bacteria. In both *E. coli* and *Salmonella typhimurium* the pathway consists of a series of ten reactions which are carried out by eight enzymes (three are bifunctional), all of which, at least in the case of *E. coli*, have been cloned and sequenced (Table 1.1). In bacterial systems the genes of the pathway are organised in an operon and precise details of the ordering of the genes has been worked out for *E. coli* and also *S. typhimurium* (Carlomagno *et al.*, 1988). In contrast to this very little work has been done on the pathway in plants. Until 1990 the only evidence of the existence of a histidine biosynthetic pathway in plants was limited to the work of Wiater (1971). He demonstrated the presence of activities for phosphoribosyl-ATP transferase, imidazoleglycerol phosphate dehydratase and histidinol phosphate phosphatase in shoot extracts of barley, oats and peas. In plants the pathway enzymes occur in the chloroplast although they are nuclear encoded and possess a targeting sequence to direct them to the chloroplast (Nagai *et al.*, 1992[a and b], Tada *et al.*, 1995). From studies on the published sequence in the databases, the *Arabidopsis* genes for histidine biosynthesis do not appear to be organised in an operon and are encoded by individual genes on different chromosomes. When I began to investigate the early enzymes of histidine biosynthesis only two plant genes from the histidine pathway had been cloned. The cloned genes encoded two of the later enzymes, namely

Enzyme name	***S. cerevisiae* gene name**	***E. coli* gene name**
phosphoribosyl- ATP transferase (EC 2.4.2.17)	*HIS1*	*hisG*
phosphoribosyl-ATP pyrophosphohydrolase phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) bifunctional	*HIS4*	*hisIE*
phosphoribosyl-formimino-5-aminoimidazole-4-carboxamide ribonucleotide isomerase (EC 5.3.1.16)	*HIS6*	*hisA*
imidazole glycerol phosphate synthase	*HIS7*	*hisH:hisF*
imidazole glycerol phosphate dehydratase (EC 4.2.1.19) bifunctional	*HIS3*	*hisB*
histidinol phosphate amino transferase (EC 2.6.1.9)	*HIS5*	*hisC*
histidinol phosphate phosphatase (EC 3.1.3.15)	*HIS2*	*hisB*
histidinol dehydrogenase (EC 1.1.1.23) bifunctional	*HIS4*	*hisD*

Table 1.1 Summary of enzymes involved in histidine biosynthesis and the corresponding genes from *E. coli* and *S. cerevisiae*

imidazoleglycerol phosphate dehydratase from *Arabidopsis thaliana* and wheat (Tada *et al.*, 1994, 1995) and histidinol dehydrogenase from cabbage, *Brassica oleracea* (Nagai *et al.*, 1992[b]). Since I began this work there have been a number of papers published detailing the cloning and genetic analysis of several of the early histidine biosynthetic genes from *Arabidopsis* by a group in Japan. This would indicate that there is substantial interest in gaining knowledge about the genes and enzymes involved in histidine biosynthesis.

In order to investigate the possibility of using this pathway as a herbicidal target the genes for the individual steps of the pathway need to be isolated and the enzymes overexpressed and characterised.

1.1.3 The Histidine Biosynthetic Pathway

The initial substrates of the histidine pathway are 5-phosphoribosyl pyrophosphate (PRPP) and ATP (Figure 1.1). In plants ribose-5-phosphate is produced via the Calvin cycle and converted to PRPP by the action of ribose-5-phosphate pyrophosphotransferase (PRPP synthetase) (Gross *et al.*, 1983). The ATP that is used in the pathway is derived either photosynthetically or from the glycolytic pathway. Histidine biosynthesis is a metabolically demanding process and requires large amounts of energy. It has been calculated that 41 molecules of ATP are utilised for each histidine molecule synthesised (Brenner and Ames, 1971). The histidine pathway branches after the fourth enzyme; the main branch carries on to produce histidine and the side branch leads to purine biosynthesis, thus the substrates for histidine biosynthesis are the same as those used for purine biosynthesis.

The intermediates of the histidine biosynthetic pathway appear to be conserved between all species, however the structure and organisation of the genes and the

corresponding proteins seems to vary (Fani *et al.*, 1995, Alifano *et al.*,1996). In bacteria such as *E. coli* all 8 genes are organised in an operon, whereas in *Saccharomyces cerevisiae* the seven genes encoding the histidine biosynthetic enzymes are distributed over 6 chromosomes (Fani *et al.*, 1995). The number of genes and hence the structure of the gene products varies between organisms (Table 1.1). In *E. coli* there are 3 bifunctional enzymes HisIE, HisB and HisD. In *S. cerevisiae* and *Neurospora crassa* several genes are linked together to form a multifunctional protein containing the HisIE and HisD activities (Fani *et al.*, 1995).

In some organisms an enzyme can be monofunctional but in other organisms it is bifunctional in nature. The best example of this is the *hisB* gene product IGPD. In *S. cerevisiae* the IGPD enzyme is a monofunctional protein (Mano *et al.*, 1993) but in *E. coli* it is a bifunctional enzyme which also encodes histidinol phosphate phosphatase (Chiarotti *et al.*, 1986).

Histidine biosynthesis plays an important role in cellular metabolism since it is connected with both *de novo* purine biosynthesis and nitrogen metabolism. It has been suggested that the histidine biosynthetic pathway was already part of the metabolic abilities of the common ancestor of all cellular organisms (Lazcano *et al.*, 1992). Histidine may be a molecular remnant of a catalytic ribonucleotide from an early biochemical stage in which RNA played a role in catalysis (White, 1976). Studies on the evolution of this pathway indicate that a gene duplication from a common ancestor gave rise to *hisA* and *hisF* (Fani *et al.*, 1994, 1997). Elongation, gene fusion and duplication events can be identified to show the evolution of the histidine biosynthetic pathway from an ancient pathway to the complex but well refined pathway observed in modern organisms (Fani *et al.*, 1995).

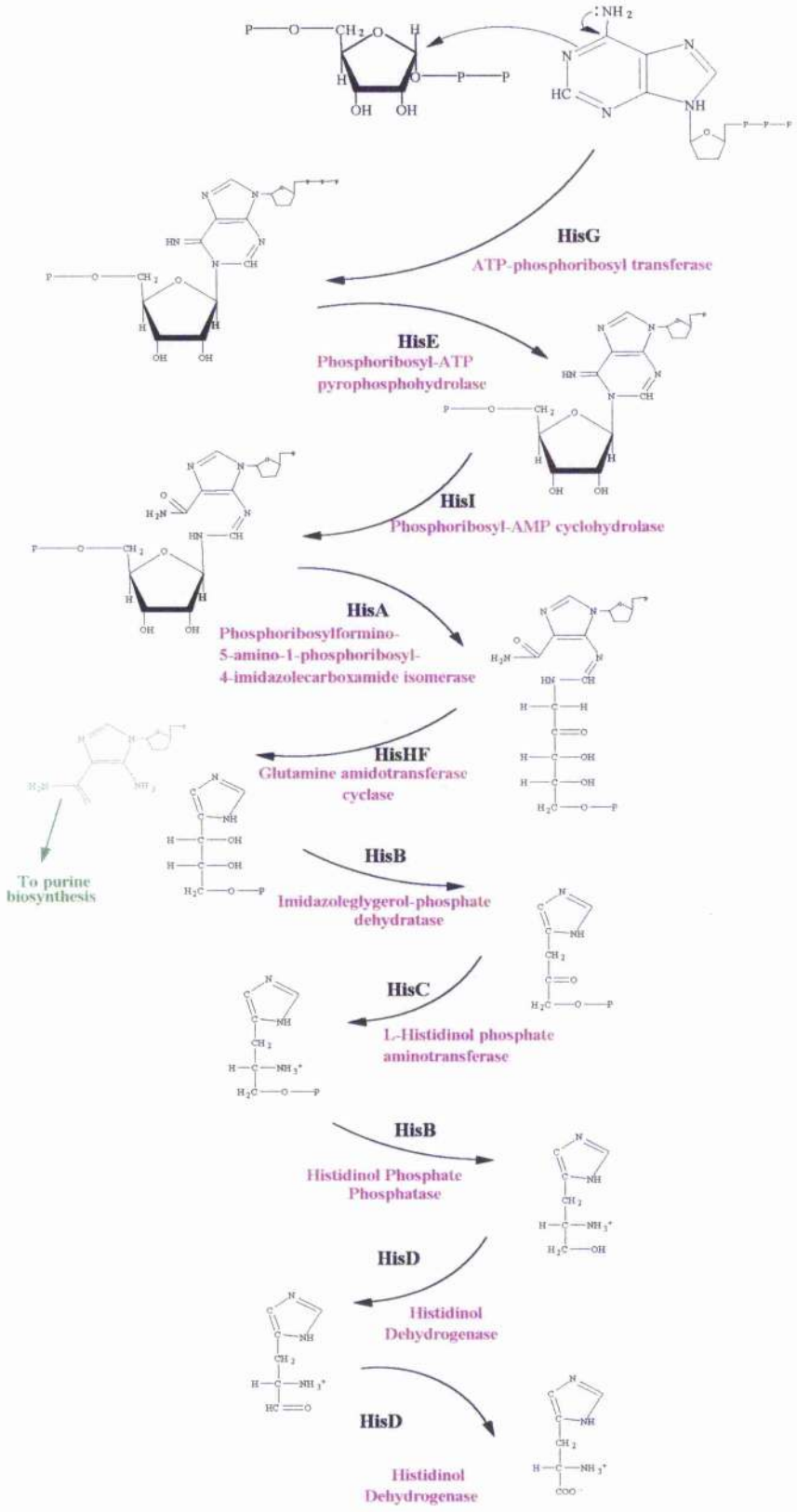


Figure 1.1 The Histidine Biosynthetic Pathway

1.2 Individual enzymes involved in Histidine Biosynthesis

1.2.1 HisG (Phosphoribosyl-ATP transferase)

The initial substrates of the histidine pathway are 5-phosphoribosyl pyrophosphate (PRPP) and ATP (Figure 1.2). HisG (N'-5'-phosphoribosyl-ATP transferase) is the first enzyme of this pathway and joins the two substrates together displacing pyrophosphate (Martin, 1963). This enzyme is feedback inhibited in bacterial systems by histidine which provides a way of regulating the pathway. Feedback inhibition of HisG provides a rapid response to fluctuations in the histidine pool while repression control of the enzyme provides long term control of histidine levels (Martin, 1963, Bell *et al.*, 1974).

Most of the information about the structure and regulation of the activity of the transferase comes from studies on the homologous enzymes from *E. coli* and *S. typhimurium*. In both organisms the purified enzyme is a hexamer composed of identical subunits of 34kDa (Voll *et al.*, 1967, Parsons and Koshland, 1974[a], Klungsoyr and Kryvi, 1971). Various aggregation states of the molecule have been observed under different conditions. The dimer is the basic oligomeric unit (Parsons and Koshland, 1974[b]) and is the most active species of the enzyme isolated from *E. coli* (Dall-Larson, 1988). Histidine and AMP shift the equilibrium of the subunits towards the hexameric state or inactive state (Dall-Larson and Klungsoyr, 1976). PR-ATP, the product of the enzyme, is known to be the most powerful ligand for shifting the protein into the hexameric form. The different aggregation states of the molecule appear to play a role in the regulation of enzyme activity. The *E. coli* enzyme for example upon binding one of the substrates PRPP, dissociates from the hexameric form to form three active dimers (Tebar *et al.*,1973, 1975). PR-ATP and other ligands

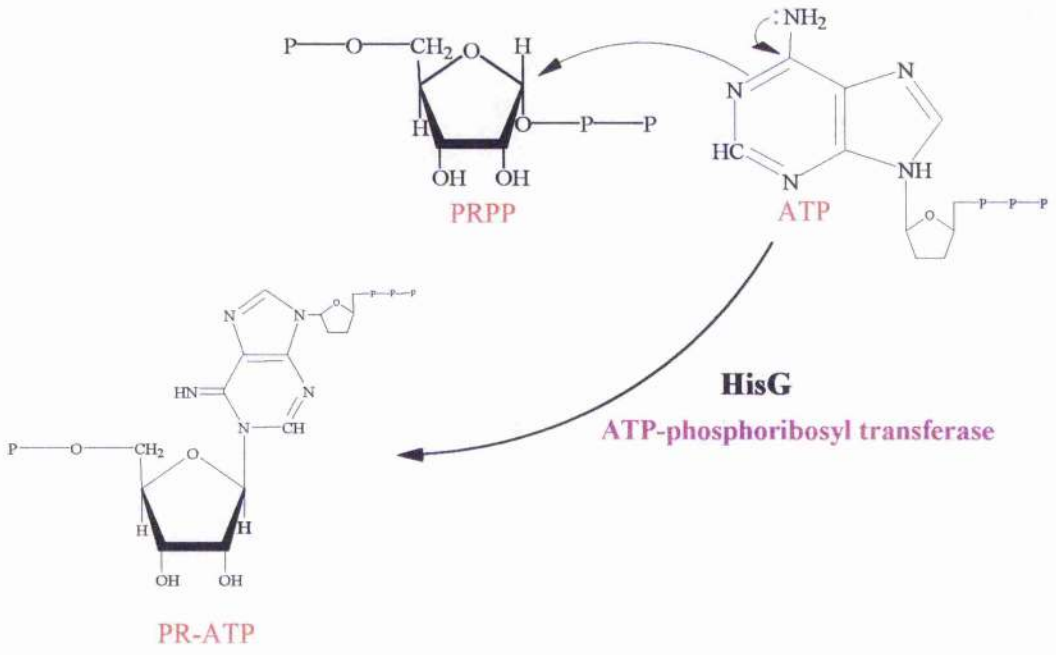


Figure 1.2 Reaction of HisG enzyme combining ATP and PRPP to form PR-ATP

that stabilise the hexameric form play a role in inhibition of enzyme activity (Dall-Larson, 1988). This appears to be connected to the mechanism of feedback inhibition observed with this enzyme. Feedback control of the HisG enzyme was first shown in 1961 by Ames *et al* who found that the *E. coli* enzyme was only sensitive to histidine in the presence of the reaction product PR-ATP.

Evidence also suggests that the HisG enzyme is involved in the repression of the histidine operon and may also be involved in regulation of the levels of Histidyl tRNA synthetase. Histidyl tRNA synthetase is thought to play a key role in the regulation of the operon (Ames *et al.*, 1983). tRNA synthetases are thought to be involved in the regulation of the biosynthetic operons for tryptophan and branch chain amino acids (Landick *et al.*, 1996, Umbarger, 1987). Recent evidence suggests that an aminoacyl tRNA synthetase like molecule has a direct involvement with and forms an

essential subunit of phosphoribosyl-ATP transferase (Sissler *et al.*, 1999). These associated proteins have been called HisZ and are based on the catalytic core of the class II histidyl tRNA synthetase but they lack aminoacylation activity. They appear to play a key role in regulation of histidine biosynthesis (Sissler *et al.*, 1999).

1.2.2 HisIE

HisIE (Phosphoribosyl-ATP pyrophosphohydrolase, phosphoribosyl-AMP cyclohydrolase) is the second enzyme of the pathway. It carries out the next two steps. Firstly it hydrolyses the two phosphates linked to the ATP part of the molecule and then the second stage of the reaction is to open the purine ring (Smith and Ames, 1965) (Figure 1.3). It was originally thought that the HisI and HisE activities were encoded by two separate genes but it is now known that they are bifunctional. There are however a few exceptions. In the archaebacterium *Methanococcus vannielii* (Beckler and Reeve, 1986, Bult *et al.*, 1996) and the gram negative organism *Azospirillium brasilense* (Fani *et al.*, 1993) the two enzyme activities appear to be on two different proteins which are encoded by two different genes. Also in *Rhodobacter sphaeroides* it appears that the *hisI* gene is monofunctional and is not linked to the *hisE* gene in any way (Oriol *et al.*, 1996). In the bacterial genes it appears that the amino terminal domain encoding the HisI activity is responsible for the third step of the pathway (cyclohydrolase activity) and that the carboxyl terminal domain encodes the HisE activity and is responsible for the second step of the pathway (the pyrophosphate hydrolase activity) (Donahue *et al.*, 1982). In some organisms such as the fungal species *S. cerevisiae* (Donahue *et al.*, 1982) and *Pichia pastoris* (Crane and Gould, 1994) the enzyme is multifunctional encoding three enzyme activities, HisI, HisE and HisD (Donahue *et al.*, 1982). This multifunctional activity is also observed

in the *his3* gene of *Neurospora crassa* (Legerton and Yanofsky, 1985). In simple terms it appears as though in the fungal species the enzyme is multifunctional encoding three enzyme activities, in the archaebacterium the activities are present individually and in bacteria and plants it appears that the enzyme activities are present as a bifunctional unit (Fani *et al.*, 1995).

The *hisIE* gene from *Arabidopsis* has been cloned independently of this project by a group in Japan led by Ko Fujimori and they published details of some cloning and genetic analysis of this gene in September 1998. They identified the *hisIE* gene as a single copy in the *Arabidopsis* genome and the product was identified as a bifunctional enzyme (Fujimori and Ohta, 1998[a]). The *hisIE* gene is composed of five exons and four introns. The splice sites between the intron exon boundaries follow the normal GU-AG rules observed in most eukaryotes including higher plants. Exon 1 encodes the transit peptide, exons 2 and 3 encode the HisI domain and exons 4 and 5 encode the HisE domain (Fujimori and Ohta, 1998[a]).

There has been little done in the way of characterisation of the *Arabidopsis* enzyme activity and no details of kinetic parameters have been published so far.

The monofunctional HisI enzyme from *M. vannielli* has been overexpressed and characterised (D'Ordine *et al.*, 1999). The enzyme is a dimer and has associated with it one equivalent of zinc per subunit that can only be removed by extensive dialysis with the zinc chelating agent 1,10-phenanthroline. The Zn^{2+} is essential for enzyme activity. Both Mg^{2+} and Zn^{2+} are required by PR-AMP cyclohydrolase for the hydrolysis of PR-AMP (D'Ordine *et al.*, 1999).

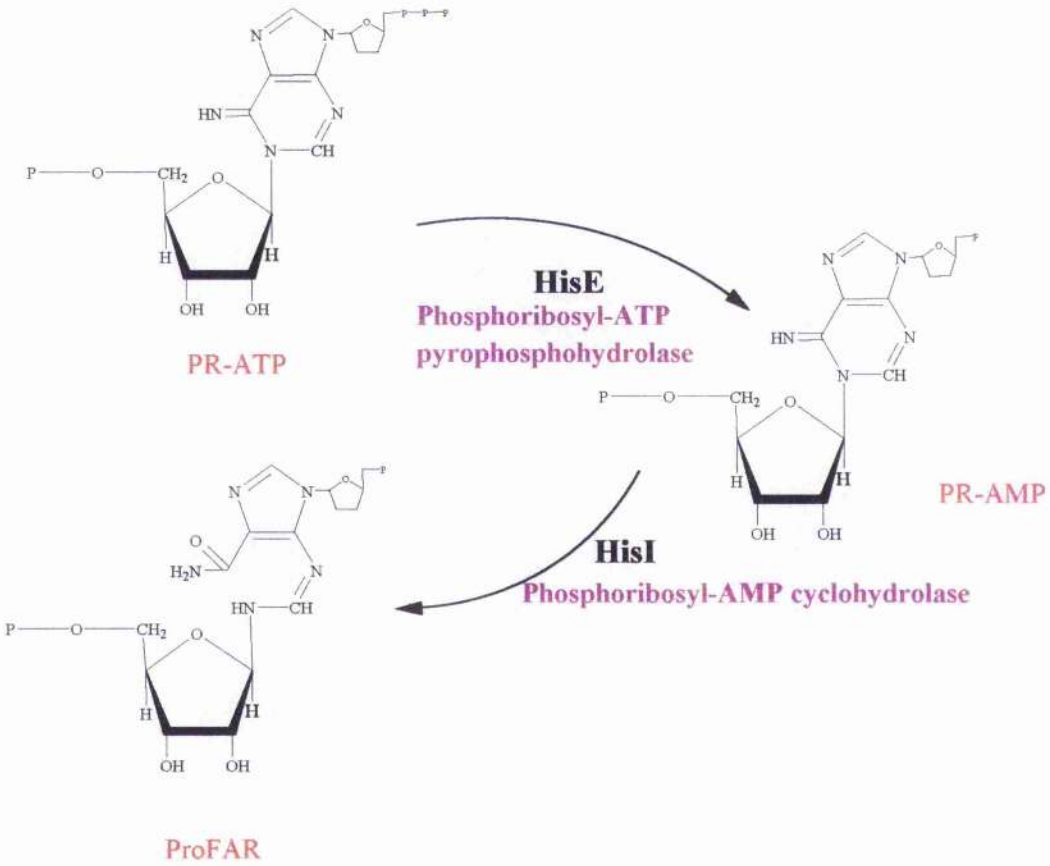


Figure 1.3 Reaction of the HisIE enzyme

1.2.3 HisA

HisA (N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase) is the third enzyme of the pathway and catalyses the opening of the PRPP ring (Smith and Ames, 1964) (Figure 1.4). The reaction catalyses the conversion of N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (BBMII or Pro-FAR) to N'-[(5'phosphoribulosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBMIII or PRFAR) (Margolis and Goldberger, 1966, 1967). This is an internal redox reaction known as an Amadori rearrangement and involves the isomerisation of the aminoaldose 5'Pro-

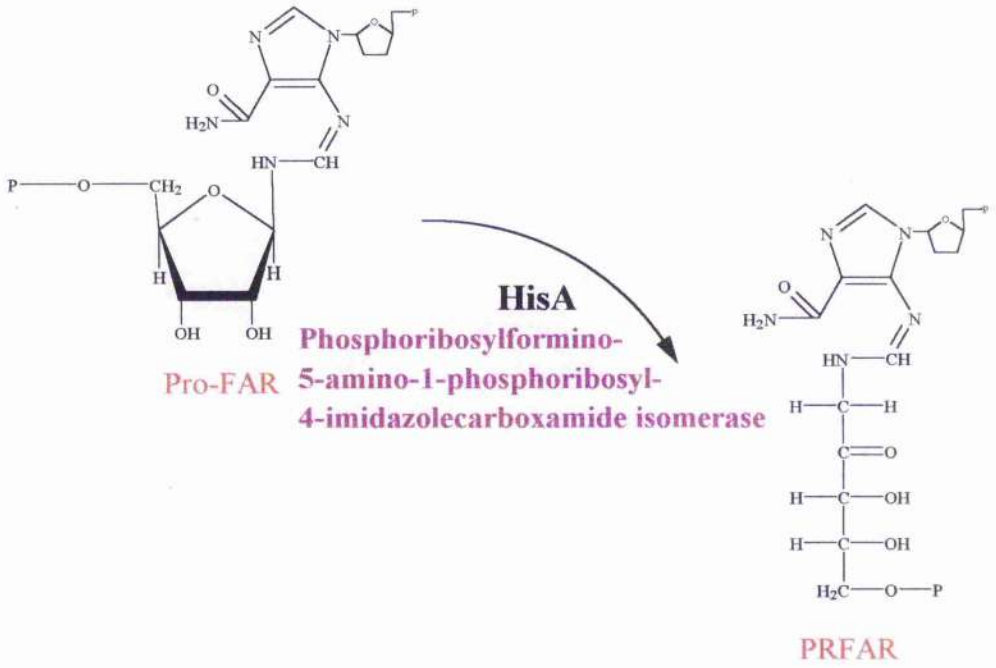


Figure 1.4 Reaction of the HisA Enzyme

FAR to the amino ketose 5' PRFAR (Smith and Ames, 1964). Little research has been done on this enzyme although recently the Japanese group led by Fujimori has cloned the cDNA encoding HisA from *Arabidopsis*, by means of complementation of an *E. coli* auxotrophic mutant. The isolated cDNA encodes a polypeptide containing 304 amino acids and has a calculated molecular weight of 33,363 Da (Fujimori *et al.*, 1998). Sequence homology with the His6 protein of yeast revealed the presence of an N terminal extension of 40 amino acids which appears to be the targeting sequence for the chloroplast. The primary structure of the mature protein was 50% identical to the *Schizosaccharomyces pombe* protein and 42% identical to the *S. cerevisiae* protein. They also identified it as a single copy gene in the *Arabidopsis* genome (Fujimori *et al.*, 1998).

A number of *hisA* homologues from other organisms have been isolated including those from *Lactococcus lactis* (Delorme *et al.*, 1992) and *Azospirillium brasilense* (Fani *et al.*, 1993), by complementation of the *E. coli hisA* mutant. HisA protein from the thermostable organism *Thermatoga maritima* has recently been purified and crystallised. Four different crystal forms have been isolated and will be used for further structure determination (Thoma *et al.*, 1999).

1.2.4 HisHF

The HisH (Glutamine amidotransferase) and HisF (cyclase) enzymes catalyse the fifth and sixth steps of the pathway converting glutamine and PRFAR to AICAR, IGP (imidazole glycerol phosphate) and glutamate (Smith and Ames, 1964) (Figure 1.5). The two enzyme activities work together as a holoenzyme complex to carry out these two reactions. The HisF protein has an ammonia dependent activity and is responsible for the conversion of PRFAR to AICAR and IGP while the HisH protein has no detectable activity (Klem and Davisson, 1993).They form a stable one to one complex that constitutes the IGP synthase enzyme (Klem and Davisson, 1993). The individual proteins are not capable of functioning individually to convert nucleotide substrate to free metabolic intermediates. The bacterial systems appear to contain separate genes for the two enzyme activities whereas in the plant system a bifunctional cDNA containing both enzyme activities has been cloned from *Arabidopsis* (Fujimori and Ohta, 1998[b]). The N and C terminal domains of the bifunctional protein show homology to the glutamine amidotransferase and cyclase of micro-organisms, respectively. This is similar to the *S. cerevisiae* enzyme where one gene encodes both activities (Kuenzler *et al.*, 1993).

The HisF domain from *T. maritima* has been purified and crystallised and is known to diffract to 1A° resolution. Precise details of the crystal structure are still to be published (Thoma *et al.*, 1999).

The two main products of the HisHF reaction are AICAR and IGP. The glutamine which is reduced to glutamate during the reaction is the third product. The AICAR formed in the reaction branches off into the pathway concerned with purine biosynthesis. It was suspected for a long time that there was an interdependence of the histidine and purine biosynthetic pathways in microbial systems. The identification of AICAR as a precursor of purine biosynthesis and a by-product of histidine biosynthesis confirmed the theory that the two pathways were closely connected (Shedlovsky and Magasanik, 1962). Therefore interrupting these early stages of the pathway would not only affect essential amino acid biosynthesis but also synthesis of purines.

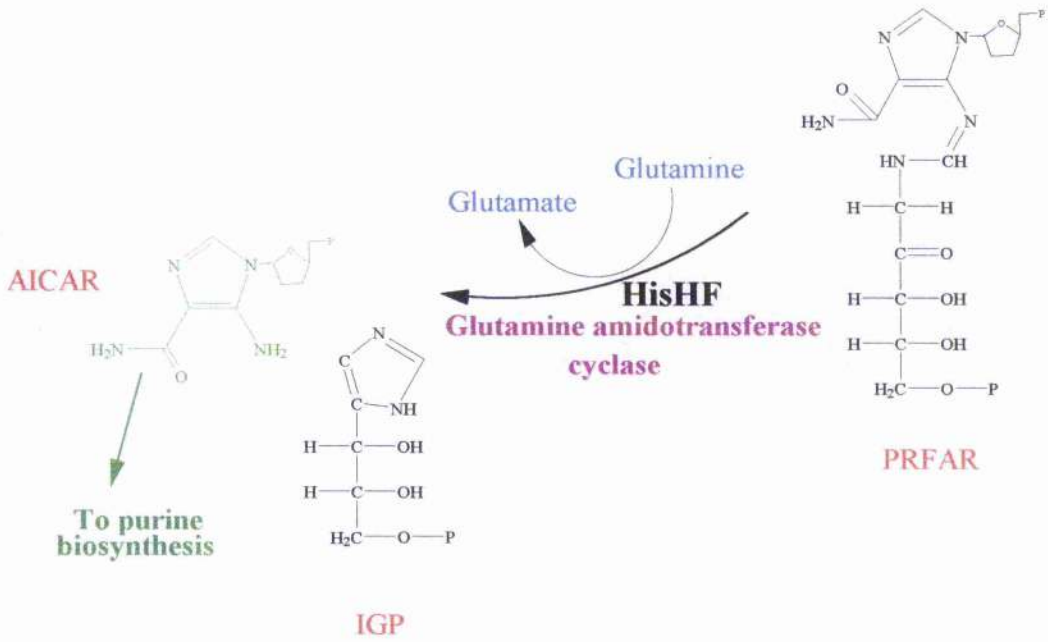


Figure 1.5 Reaction of the HisHF enzyme

1.2.5 HisB

The *hisB* gene encodes the bifunctional enzyme IGPD and histidinol-phosphate-phosphatase which catalyses the sixth and eighth steps of the pathway (Figures 1.6 and 1.8). In certain species such as the bacteria *E. coli* and *S. typhimurium* the enzymes are bifunctional and are encoded by a single gene. In plants and certain other bacterial species the two enzyme activities appear to be monofunctional and are encoded by different genes, while in *S. cerevisiae* the activities are encoded by two separate genes. The IGPD activity isolated from *Arabidopsis* and wheat is monofunctional (Mano *et al.*, 1993, Tada *et al.*, 1994).

IGPD at least in bacterial systems occurs as the C terminal domain of the *hisB* gene product. IGPD (EC 4.2.1.19) is the sixth enzyme of the pathway and catalyses the dehydration of IGP to IAP (imidazole-acetol-phosphate) (Ames and Mitchell, 1955). (Figure 1.6). The mechanism by which the dehydration occurs is unclear. In most cases dehydration would occur as a β-elimination where the hydrogen to be removed is relatively acidic due to the presence of an adjacent carbonyl or imine functional group. In this case the hydrogen which is removed is relatively non acidic and therefore the mechanism may not be a straight forward β-elimination. It has proved useful as a novel herbicide target in recent years (Mori *et al.*, 1995). Triazole phosphonates are effective inhibitors of IGPD and act as herbicides *in vivo* with a similar potency to Glyphosate™ (Hawkes *et al.*, 1993).

Digests and southern blot analysis appears to indicate the possibility of two IGPD genes in *Arabidopsis* (Tada *et al.*, 1994). These studies have also revealed potential targeting sequences. It is presumed that the first 30-70 amino acids of the plant protein sequences are the chloroplast targeting sequences. This further supports the idea that histidine biosynthesis takes place in the chloroplast within plant tissues.

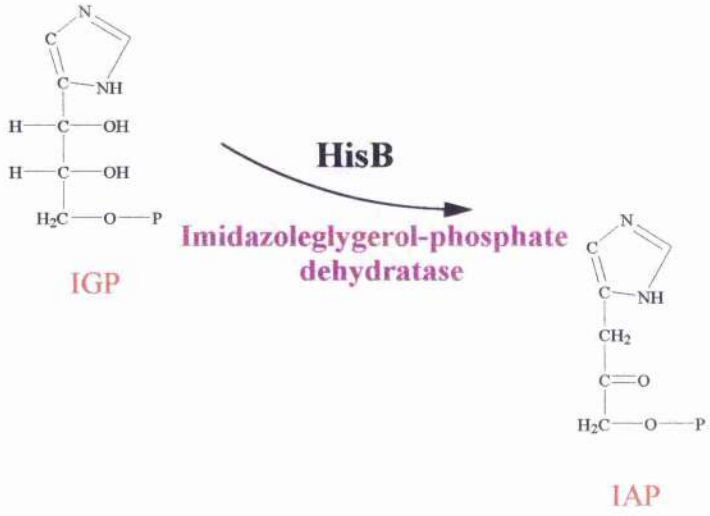


Figure 1.6 First Reaction of the HisB enzyme

1.2.6 HisC

The seventh step of the pathway is carried out by the product of the *hisC* gene IAP aminotransferase or histidinol phosphate amino transferase. This reaction involves the reversible transamination of IAP and a nitrogen atom from glutamate leading to the production of α-ketoglutarate and L-histidinol-phosphate (Figure 1.7). This enzyme is pyridoxyl-phosphate dependent (Brenner and Ames, 1971).

The gene has been recently cloned from *Nicotiana tabacum* by complementation of an *E. coli* auxotroph (El Malki *et al.*, 1998). In yeast this enzyme is under the process of general amino acid control (Hinnesbusch, 1988). General amino acid control is a mechanism which allows the expression of a complex set of genes or a group of pathways to be controlled via a small subset of proteins. The *N. tabacum* enzyme appears to possess a chloroplast targeting sequence (El Malki *et al.*, 1998). Frequent database searching has failed to yield a corresponding gene for histidine phosphate amino transferase in *Arabidopsis.*

In *S. typhimurium* histidinol phosphate amino transferase has been extensively characterised (Martin, 1970, Albritton and Levin, 1970). The enzyme is a homodimer with a native molecular weight of 59kDa. Both *E. coli* and *N. tabacum* have molecular weights of around 40kDa and comparison of amino acid sequences show convincing evidence for homology despite a low degree of amino acid identity (El Malki *et al.*, 1998).

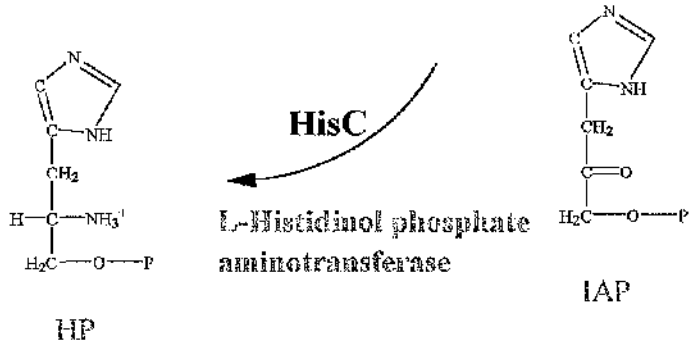


Figure 1.7 Reaction of the HisC enzyme

1.2.7 HisB enzyme (Histidinol phosphate phosphatase)

The histidinol phosphate is converted to L-histidinol by the phosphatase activity of the N terminal region of the bacterial *HisB* gene product (Figure 1.8). In *E. coli and S. typhimurium* the histidinol phosphate phosphatase activity is associated with the N terminus of the HisB bifunctional enzyme (Carlomagno *et al.*, 1988, Chiarotti *et al.*, 1986). The histidinol phosphate phosphatase activity from *S. cerevisiae* has been characterised at the genetic and biochemical levels (Millay and Houston, 1973) but the sequence has no similarity to the *E. coli* enzyme (Malone *et al.*, 1994).

Recently the gene *ytvP* encoding a histidinol phosphate phosphatase activity has been identified in *Bacillus subtilis* (Le Coq *et al.*, 1999). This completes the pathway in *B. subtilis.*

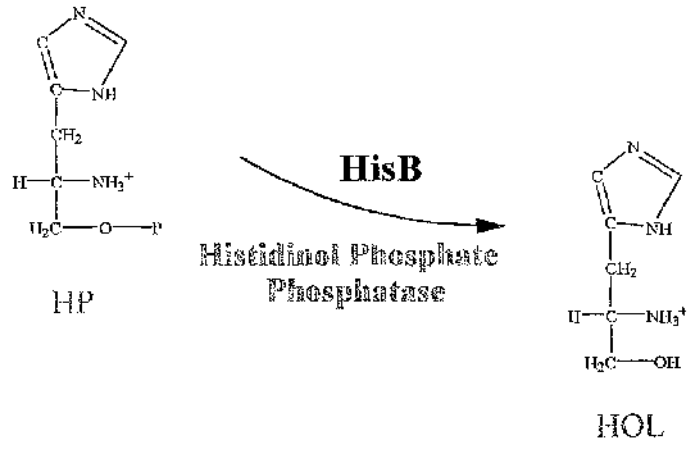


Figure 1.8 Second reaction of the HisB gene product

1.2.8 HisD enzyme (Histidinol dehydrogenase)

Histidinol dehydrogenase catalyses the oxidation of L-Histidinol to the amino acid L-Histidine (Adams, 1954). The reaction proceeds via the unstable amino aldehyde L-Histidinal which is not found as a free intermediate (Adams, 1954, Görisch and Hölke 1985). (Figure 1.9). The bifunctional enzyme Histidinol dehydrogenase is a NAD linked dehydrogenase (Bürger and Görisch, 1981, Kirschner and Bisswanger, 1976) which in the case of the *S. typhimurium* enzyme is a homodimeric zinc metalloenzyme. In fungal species such as *S. cerevisiae* and *N. crassa* the HisD enzyme activity forms part of a multifunctional enzyme product containing the HisIE activity which is encoded by the *His4* gene (Fani *et al.*, 1994). A cDNA encoding histidinol dehydrogenase has been isolated from cabbage (*Brassica oleracea*) and the protein has been overexpressed and partially characterised (Nagai *et al.*, 1992[b]). The histidinol

dehydrogenase activity was detected in chloroplasts isolated from the leaves. This gives further support to the idea that the histidine biosynthetic pathway occurs in the chloroplasts (Nagai *et al.*, 1992 [b]).

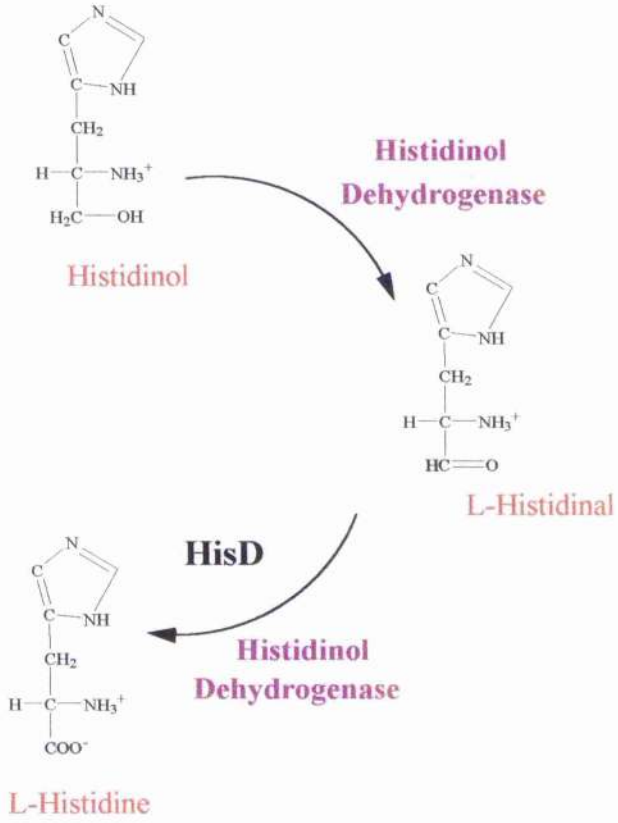


Figure 1.9 Reaction of the HisD enzyme

1.3 The aims of the project

The aim of this project was to clone the genes or cDNAs for the first three enzymes of histidine biosynthesis from *Arabidopsis* and subsequently overexpress and purify at least one of these enzymes. It transpired that the most successful results were obtained with the clones of the second enzyme HisIE. Therefore this enzyme was overexpressed and purified. Many of the intermediates and substrates required by the enzymes of this pathway are not commercially available and so in order to characterise

the HisIE enzyme and determine the kinetic parameters it was necessary to produce the substrate for the enzyme. This became a key part of the project and was achieved by biotransformation from PRPP and ATP using a purified extract of HisG from *E. coli*. Once there was a suitable supply of substrate the final third of the project was to determine the kinetic parameters of the enzyme and to use chemical modification with substrate protection and site directed mutagenesis to identify essential residues in the active site.

Chapter 2

Materials and methods

2.1 Materials and Methods

2.1.1 Chemicals

Chemicals and biochemicals were generally Analar grade or else the highest grade available. Unless otherwise stated all chemicals were purchased from the Sigma Aldrich Company Ltd, England or Fisher Scientific, UK. Bacto-tryptone, bacto-yeast extract and bacto-agar were obtained from Difco Laboratories, Detroit, USA Agarose and low melting point agarose were obtained from GIBCO BRL Life Technologies, Scotland. Complete™ EDTA free Protease inhibitor cocktail tablets were obtained from Boehringer Mannheim, UK.

2.1.2 Proteins and Enzymes

Restriction enzymes and their buffers were obtained either from New England Biolabs. Incorporated, England; Promega Corporation, England or Boehringer Mannheim, UK.

Vent DNA polymerase was obtained from New England Biolabs. Bacteriophage T4 DNA ligase and Calf Intestinal Alkaline Phosphatase (CIAP), were obtained from Promega Corporation.

Phosphoribosyl ATP Transferase was purified to homogeneity by Mr J. Greene from an overexpressing strain of *E. coli* according to the method of Elwell and Coggins (Unpublished results, see Appendix VII).

2.1.3 Oligonucleotides

Oligonucleotides for use both in the polymerase chain reaction and for DNA sequencing were purchased from Cruachem Ltd, Glasgow or from Genosys Biotechnologies Inc., England.

2.1.4 Chromatography media

DEAE Sephacel, Phenyl Sepharose, Q-Sepharose, Sephacel S-200 and Sephadex G15 were obtained from Pharmacia Biotech, England. Reactive Red 120 agarose resin was purchased from Sigma Aldrich Company Ltd, England.

2.1.5 Pre-packed media

Pre-packed Mono Q, Resource Q, and Superdex Peptide columns, were purchased from Pharmacia Biotech. These columns were used on a Pharmacia FPLC system or a Shimadzu LC-10 HPLC system. Partisil ODS3 columns were purchased from Phenomenex, England and a Supelcosil LC-18-T column was purchased from Supelco Inc. These columns were used on the Shimadzu HPLC system.

2.2 Media and supplements

2.2.1 Media for bacterial growth

E. coli strains BL21(DE3)pLysS and DH5α were routinely cultured in Luria-Bertani medium (LB). Solid media was obtained by adding 1.5% bacto agar (Difco, USA) to the liquid LB (Table 2.1).

2.2.2 Minimal media

The histidine mutants were grown in a minimal media prepared as in Goldschmidt *et al.*, (1970). The following was added to an autoclaved solution of 900ml of distilled water and 1% agar: 100ml 10X salt solution (Table 2.2), 2ml of a 125mg/ml thiamine stock and 4ml of a 50% glucose solution. Additional requirements were met with the following final concentrations of supplements: histidine, 100μg/ml;

arginine, 20μg/ml; leucine, 20μg/ml; methionine, 20μg/ml. Antibiotics were added as required at the concentrations listed below.

2.2.3 Antibiotics

Antibiotics were used at the following final concentrations in all experiments

Ampicillin 50-100μg/ml

Tetracycline 12.5μg/ml

Chloramphenicol 25μg/ml

a) Ampicillin

A 25mg/ml stock solution of ampicillin dissolved in distilled water was sterilised by filtration through a 0.22μm filter. 1ml aliquots were stored at -20°C. The growth medium was allowed to cool to 50-55°C before the addition of ampicillin to a final concentration of 50-100μg/ml.

b) Tetracycline

A stock solution of 12.5mg/ml dissolved in 50% ethanol was stored at -20°C in the dark. Filter sterilisation was not required. Medium was allowed to cool to 50-55°C before the addition of antibiotic at 12.5μg/ml final concentration.

c) Chloramphenicol

A stock solution of 25mg/ml was dissolved in 100% ethanol and stored at -20°C. Filter sterilisation was not required. Media was allowed to cool to 50-55°C before the addition of antibiotic at a final concentration of 25μg/ml.

All media and plates containing antibiotics were stored at 4°C and used within 4 weeks.

Media	Composition per litre	Comments
LB (Luria-Bertani broth)	10g bacto-tryptone 5g yeast extract 10g NaCl	Sterilised by autoclaving at 15psi for 30 minutes
LB agar	LB plus 15g bacto agar	Sterilised by autoclaving at 15psi for 30 minutes
Minimal media	100ml 10X salt solution (See Table 2.2) 2ml thiamine (125mg/ml) 4ml 50% glucose solution 15g bacto agar	Additional requirements histidine 100μg/ml arginine 20μg/ml leucine 20μg/ml methionine 20μg/ml.

Table 2.1 Growth media

Table 2.2 Buffer solutions

Buffer	Composition	Comments
TFB1	100mM RbCl 50mM $MnCl_2$ 30mM KCl 10mM $CaCl_2$ 15%(v/v) glycerol	Components are added to water in sequence and then made to pH 5.8 with dilute acetic acid and filter sterilised
TFB2	10mM MOPS 10mM RbCl 75mM $CaCl_2$ 15% (v/v) glycerol	Components added in sequence and adjusted to pH 7.0 using HCl and filter sterilised
10X Salt solution	70g K_2HPO_4 30g KH_2PO_4 5g Na citrate.$2H_2O$ 1g $MgSO_4.7H_2O$ 10g $(NH_4)_2SO_4$	Components are dissolved in water brought to 1 litre and autoclaved
TE buffer	10mM Tris HCl pH 7.2 1mM EDTA	
10X TBE Buffer	900mM Tris base 900mM boric acid 25mM EDTA pH 8.0	Used at 1X concentration in most cases
Agarose gel loading buffer	500µl glycerol 50µl bromophenol blue 100µl β-mercaptoethanol 250µl 20% SDS	Total volume 1ml
SDS-PAGE sample buffer	60mM Tris-HCl pH6.8 2% (w/v) SDS 20% (v/v) glycerol 0.1%(v/v) β-mercaptoethanol 0.0025% (w/v) bromophenol blue	

2.2.4 Isopropyl-β-D-thiogalactopyranoside (IPTG)

A 1M stock solution was made by dissolving 2g of IPTG in a final volume of 10ml. The solution was filter sterilised through a 0.22μm filter and stored as 1ml aliquots at -20°C.

2.3 General methods

2.3.1 General laboratory methods

General molecular biological techniques were carried out as described in Sambrook *et al.,* (1989). General laboratory methods for handling proteins and enzymes were as described in (Deutscher, 1990).

2.3.2 French pressure cell

Cells were broken by two passages through an automatic French pressure cell at 950psi. The pressure cell was pre-cooled on ice before use.

2.3.3 pH measurement

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

2.3.4 Conductivity measurements

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

2.3.5 Protein estimation

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

2.3.6 Lyophilization

Substrate solutions were either collected into polypropylene tubes (small samples <1ml) or into 25-100ml glass round bottomed flasks (larger samples) and the contents shell frozen in a dry ice/ethanol mixture before lyophilization on an FTS Systems (Stone Ridge, New York, USA) Flexi-Dry freeze dryer.

2.3.7 Spectrophotometric measurement of nucleic acids

Nucleic acid concentrations were determined spectrophotometrically at 260nm (Sambrook *et al.*, 1989) in a Hewlett Packard 8453 UV-Visible Spectrophotometer using Quartz cuvettes. In a 1cm path length cuvette an absorbance of 1.0 corresponds to 50μg/ml for double stranded DNA and 20μg/ml for single stranded oligonucleotides.

2.4 Bacterial strains and cloning vectors

2.4.1 Storage of Bacterial Strains

Bacterial strains were stored as glycerol stocks. These were made by the addition of 80% (v/v) glycerol/water to growing bacterial cultures to a final glycerol concentration of 50% (v/v). Duplicate collections of stock cultures were stored at-20°C and -80°C. Some cultures were temporarily stored by streaking onto agar plates and maintaining at 4°C for up to 4 weeks.

2.4.2 Bacterial strains

The bacterial strains and plasmids used in this project are listed in Tables 2.3 and 2.4.

2.4.3 *Escherichia coli* Histidine auxotrophic mutants

E. coli JC411 [*leuB6, fhuA2, lacY1, supE44, gal-6,* λ⁻, *hisG1, rfbD1*?, *galP63*?. *argG6, rpsL104, malT1* (λ^R), *xyl-7, mtl-2, metB1*], UTH903 [*ara-14, galK2,* λ⁻, *hisI903, rpsL145, malT1*(λ^R), *xyl-5, mtl-1*] and HfrG6 [*hisA323*, λ⁻] were provided by Barbara Bachmann of the Coli Genetics Stock Centre, Dept. of Biology, Yale University, New Haven, CT, USA. These strains were confirmed to be auxotrophic for histidine as they only grew on minimal media which was supplemented with histidine.

2.4.4 Storage of mutant strains

Working stocks of the mutant strains were stored at -20°C in LB in 50% (v/v) glycerol (see 2.4.1). Working cultures were obtained by streaking from the stock cultures onto LB-agar. Single colonies of the mutant strains were stabbed into 3ml LB-agar in bijouets. These "stabs" were stored in the dark at room temperature. Mutants were streaked onto minimal media plates plus and minus histidine before use in the complementation experiments to ensure that they were still auxotrophic for histidine.

Bacterial strain	**Genotype**	**Reference**
Escherichia coli DH5α	F⁻ ϕ80d*lacZ*ΔM15 *recA1 endA1 gyrA96 thi-1 hsdR17*($r_k^- m_k^+$) *supE44 relA1 deoR* Δ(*lacZYA-argF*)U169	Hanahan, 1985
Escherichia coli BL21(DE3)pLysS	F- *ompT hsdS*$_B$ r_B^-,m_B^- *dcm gal* λDE3 pLysS Cm^r	Studier and Moffatt, 1986
Escherichia coli Y1090	Δ(*lac*U169) *proA*$^+$ Δ(*lon*) *ara*D139 *strA supF* [*trp*C22::Tn10(tet^r)] (pMC9) *hsdR* ($r_K^- m_K^+$)	Huynh *et al.*, 1985
Escherichia coli JC411 (*HisG* mutant)	*leuB6, fhuA2, lacY1, supE44, gal-6,* λ⁻ *, hisG1, rfbD1?, galP63?. ArgG6, rpsL104, malT1* (λ^R), *xyl-7, mtl-2, metB1*	Low; 1968 Clark *et al.*, 1969
Escherichia coli UTH903 (*HisI* mutant)	*ara-14, galK2,* λ⁻, *hisI903, rpsL145, malT1*(λ^R), *xyl-5, mtl-1*	Goldschmidt *et al.*, 1970
Escherichia coli HfrG6 (*HisA* mutant)	*hisA323,* λ⁻	Matney *et al.*, 1964

Table 2.3 Bacterial Strains used for general cloning purposes.

Plasmid	Antibiotic resistance	Reference
pTB 361	tetracycline resistance	T7 expression plasmid Brockbank and Barth, 1993
pGEM® -5Zf(+/-)	ampicillin resistance	T7 cloning vector Promega Corporation
pUC19	ampicillin resistance	General cloning vector Yanish and Peron, 1982
pLysS	chloramphenicol resistance	T7 lysozyme plasmid Moffatt and Studier, 1987
pSACIE25	tetracycline resistance	HisIE expression plasmid containing truncated HisIE construct (this study)
pSACIE13	tetracycline resistance	Full length HisIE construct (this study)
pSACIEH146A	tetracycline resistance	Histidine 146 to alanine mutant of truncated HisIE construct
pSACIEH245A	tetracycline resistance	Histidine 245 to alanine mutant of truncated HisIE construct

Table 2.4 Plasmids used for Molecular cloning

2.4.5 cDNA libraries

The λYES cDNA library was prepared as in Elledge *et al.* (1991). mRNA was prepared from stem and leaf tissue of *Arabidopsis thaliana* plants. The cloning vector rpSE937 (GenBank accession number U02436) is the plasmid released from λYES-P, and is 7798bp in length. The sequence begins with the restriction sites EcoR1, Xho1, EcoR1 adjoining each other. The cDNAs were inserted into the Xho1 site and could be excised by digestion with EcoR1. This library was a kind gift from Dr Nigel Urwin, Division of Biochemistry and Molecular Biology, University of Glasgow.

Two λPRL libraries, one from wheat and the other from *Arabidopsis*, where obtained from Dr Danielle Werck, Strasbourg. These libraries are generated in a λZIPLOX expression vector which allows the cloning and screening of cDNAs by conventional nucleic acid screening techniques (D'Alessio *et al.*, 1992), but also permits the excision of the cDNA as a self replicating plasmid for easier manipulation (Lin *et al.*, 1992).

2.4.6 Growth of Plasmid containing *E. coli*

Growth of plasmid-containing *E. coli* cells was achieved on LB or minimal medium containing the appropriate antibiotic. For pUC19 and λYES ampicillin was added to the medium at a final concentration of 50μg/ml. Isopropyl β-D-thiogalactoside (IPTG) was added to the medium at a final concentration of 120μg/ml to induce expression of the cDNA in the λYES vector. For growth of pTB361 and derivatives such as the overexpression construct pSACIE25 tetracycline was added to the media.

2.4.7 Growth of cells and Isolation of plasmids

This was carried out using standard techniques of inoculating a 100ml overnight culture of LB (containing appropriate antibiotics), with a single colony of the correct strain. These were then grown overnight in a shaking incubator at 37°C. Plasmid isolation was done using a Qiagen Midi plasmid preparation kit (Qiagen Ltd. See Appendix V) as described in the manufacturers instructions, and the resulting plasmid DNA resuspended in 150µl of sterile water.

Plasmid minipreps were performed either by Wizard Mini preps from a 3ml overnight culture, (See Appendix II) or by lysis of cells grown in a 1ml overnight culture of LB (containing appropriate antibiotics), by a lysozyme/boiling method (Holmes and Quigley, 1981).

2.4.8 Production of competent cells

A single colony of an *E. coli* strain was grown overnight at 37°C in 10ml of LB. 2ml of this overnight culture was then used to inoculate 100ml LB in a 250ml flask. The cells were grown at 37°C on a microbiological shaker at 200rpm until they reached mid log phase (A_{550}=0.35≤0.5). The cells were chilled on ice and kept at a maximum of 4°C throughout the rest of the procedure. The culture was transferred to chilled centrifuge pots and centrifuged at 2500rpm in a Beckman Model J-6B for 5mins to pellet the cells. The pellet was resuspended in 10.5ml of ice cold TFB1 (Table 2.2). The cells were then incubated on ice for 90 minutes and re-pelleted. The pellet was resuspended in 2.8ml of TFB2 (Table 2.2). Cells were aliquoted into 100µl or 200µl aliquots and were snap frozen in a dry-ice/ ethanol bath and stored at -80°C.

Competent cells can be stored at -80°C for an indefinite period without showing any significant reduction in transformation efficiency. The transformation efficiency of the competent cells was measured by transforming 200µl aliquots of cells with 1µg pUC19 and plating onto medium containing the appropriate antibiotics and other additives.

2.4.9 Transformation protocol

An aliquot of frozen cells was allowed to thaw on ice for 30 minutes prior to use. DNA was added to the cells and following gentle mixing the cells were incubated on ice for 30minutes. The cells were heat shocked at 37°C for 60 seconds before being returned to ice for a further 30 minutes. 1ml of LB was added to the cells and they were then incubated at 37°C for 40-60 minutes to aid recovery. The cells were pelleted for 30 seconds at high speed and resuspended in 100µl of either LB or minimal medium and then plated out onto selective media.

2.4.10 cDNA Library Screen

A 2ml aliquot of competent mutant strain cells was thawed and incubated with 10µg cDNA library on ice for 20mins. The cells were given a heat shock of 42°C for 90secs and then returned to ice for 2 mins; LB was added to a final volume of 10ml and the cells shaken at 150rpm and 37°C for 1hr before being pelleted for 10 mins at 2500rpm in a Beckman Model J-6B centrifuge. The pellet was resuspended in 1ml minimal media. Aliquots of 50µl cells were spread on plates of selective media and grown in a 37°C incubator overnight.

Competent cells were plated on selective media with and without histidine supplementation in order to show that the cells had no capacity for antibiotic resistance before being transformed with the plasmid. Also, transformed cells were grown on selective media with histidine supplementation to allow all the cells carrying antibiotic resistance to grow. This was also the method used to measure the efficiency of transformation. Plating known dilutions of transformed cells on his^+ selective media showed how many transformants were derived from adding 1µg of plasmid DNA to 200µl of cells.

2.5 Manipulations of DNA

2.5.1 Plasmid extraction

Bacterial colonies growing on his^+ selective media were grown in 5ml LB containing ampicillin and plasmid isolated by Wizard Mini Preps or in 100ml LB (ampicillin) for Qiagen™ midi plasmid extraction. (See Appendix V).

2.5.2 Gel electrophoresis

Gel electrophoresis of DNA samples was performed using a 1% agarose gel in 1X TBE (Table 2.2) containing 0.5µg/ml ethidium bromide. Samples were loaded in a 4:1 volume ratio with agarose gel sample loading buffer (Table 2.2). A standard 1kb ladder (Gibco-BRL) was used on all gels as a marker to identify the size of the separated bands. The gels were run at 100V and room temperature until the bromophenol blue neared the end of the gels; DNA bands were located by viewing under UV light. Photographs were taken where appropriate.

2.5.3 Restriction digests

Digests contained 3-5μg of DNA with 1-2 units of each enzyme along with the appropriate reaction buffer. Digests were carried out at 37°C for one hour. In the case of the NdeI, BamHI double digests, since NdeI is particularly unstable, this enzyme was added 30 mins prior to the addition of BamHI; the total time for the digest was therefore one hour 30 mins.

The cDNA insert in λYES vector is flanked by EcoRI sites. Plasmid DNA was digested with EcoRI in Buffer H (Promega) at 37°C for 2hrs.

2.5.4 Gel purification of DNA fragments

The restricted plasmid or DNA fragment was separated on a low melting point agarose gel and the band excised. An equal volume of TE buffer (Table 2.2) was added along with an equal volume of TE saturated phenol and incubated at 65°C for 5 minutes. This was vortexed and spun at high speed for 5 mins. The aqueous layer was removed quickly and transferred to a fresh tube. Further extractions of the aqueous layer with phenol, TE saturated phenol/chloroform, and TE saturated chloroform were performed. The DNA was precipitated by adding 1 volume of isopropanol and 0.1 volumes of 1M NaCl and then vortexing briefly before centrifuging at full speed in a microfuge for 15 minutes. The pellet was air dried and then resuspended in a small volume of sterile water.

Alternatively the gel fragment was purified using the Wizard PCR Cleanup kit. (See Appendix IV).

2.5.5 Use of Alkaline Phosphatase to dephosphorylate DNA

Alkaline phosphatase (Promega) was used to dephosphorylate restricted plasmid DNA. 1µl calf intestinal alkaline phosphatase (CIAP) (10U/µl) and 10µl of CIAP buffer (Promega) were added to the plasmid digest and the volume made up to 100µl with sterile water. The reaction was incubated for 40 minutes at 37°C to allow dephosphorylation of the restricted plasmid.

2.5.6 DNA purification

DNA digests of both plasmid and insert DNA were purified using a Wizard™ DNA Clean-Up System (Promega), in order to remove traces of restriction enzymes and other contaminating enzymes such as alkaline phosphatase (See Appendix III). The DNA was eluted with 50µl of sterile water.

2.5.7 DNA ligation

Prior to ligation, concentrations of DNA were measured either spectrophotometrically or following gel electrophoresis by comparison with a known amount of molecular weight marker. Several ratios of vector to insert DNA [1:3, 3:1 (w/w)] were used in separate ligations. Approximately 400ng of DNA was used per ligation reaction. DNA ligation reactions were carried out using 1µl T4 DNA ligase and 1µl of the appropriate buffer along with plasmid and/or insert DNA. All ligations were carried out for 16 hours at 4°C in a final volume of 10µl.

2.6 Cloning by PCR

2.6.1 Primers for PCR

PCR primers were designed to incorporate two restriction sites that would facilitate easier cloning of the insert DNA into the new vector system; the included sites were for the enzymes NdeI and BamHI. The NdeI site is extremely useful as it incorporates the ATG start codon and ensures that the DNA is inserted in the correct reading frame (Table 2.5).

2.6.2 PCR reactions

PCR reactions were carried out using Vent DNA polymerase along with appropriate buffers and salts. 1-2 units of enzyme were added per reaction along with 10µl of 10X Vent reaction buffer, 10µl of 10X dNTP mix (10mM stock), 100pmol of each primer, 100ng of DNA and also $MgSO_4$ in varying concentrations from 1-5mM (final). 100µl reactions were set up in 0.2ml PCR tubes.

The following temperature profile was used for 30 cycles in a Biorad Gene Cycler:

3 min @ 94° C (initial denaturation)

1 min @ 94° C

1 min @ 55° C

1 min 30 @ 72° C

A final elongation step of 5 minutes at 72°C was performed at the end of the amplification reaction to ensure that all amplified material was full length.

Oligonucleotide	Sequence	Comments
Nde1	5'GTT TCT CGT AGC GTT CAT ATG GCG GTA TCG	Used to insert NdeI site to permit cloning of the full length HisIE cDNA into pTB361 expression vector; also used for cloning into the Intein™ System
Nde2	5' GTA TTC GCG TGC AAT GAT CAT ATG AAC AAT G	Removes the targeting sequence before the residues NIAL in the HisIE protein sequence. Also used for cloning into the Intein™ vector system
BamIE	5' CAA TGA ACA ACA GGA TCC GTT CCG AG	Used to generate the 3' restriction site for cloning
CtermForIE	5' CAA TTA AAC AAT CAT ATG GCT TCA GG	For cloning of *hisE* C-terminal domain into the Intein™ cloning system
NtermRevIE	5' GCT TGT TTC CTG AGA ATT CAT CAT TG	For cloning of *hisI* N-terminal domain into the Intein™ cloning system
CtermRevIE	5' CGA AGA AAA GAA TTC CTT TGT ACG G	For cloning of the *hisE* C-terminal domain of *hisIE* into the Intein™ cloning system (See Appendix VI)

Table 2.5 Primers used for PCR based molecular cloning

Oligonucleotide	Sequence	Comments
2869	5' TTG TGT GGA ATT GTG AGC GG	Primer reading from the *lac* promoter region of the λYES vector
2873	5' TCT ATA CTT TAA CGT CAA GGA G	Primer reading from the *Gal* promoter region of the λYES vector towards the insert
Reverse1	5' TGA CAG GTA GGG TCC ATC AG	To sequence back through the *hisIE* sequence
Forward1	5' CAA GGC TTT GTT AAT AGG GAG	To sequence further into the *hisIE* clone
Forward 2	5' AGA ATC AAT CAT TTC CAA GCG G	For further sequencing into the *hisIE* clone
T7	5' TAA TAC GAC TCA CTA TAG GG	Sequencing of constructs containing a T7 promoter
M13	5' GTA AAA CGA CGG CCA GTG	

Table 2.6 Oligonucleotides used for DNA sequencing

2.6.3 DNA purification methods

DNA from PCR reactions was purified and separated from contaminating products such as primer-dimers and amplification primers using a Wizard™ PCR Preps DNA purification system. (See Appendix IV). The purified DNA was eluted with 50µl of sterile water for use in further cloning procedures.

2.6.4 DNA Sequencing

Double-stranded plasmid DNA was sequenced using the Perkin Elmer ABI Prism™ Dye Terminator Cycle Sequencing Core Kit with Amplitaq®DNA Polymerase, FS. The samples were loaded and run by the DNA sequencing service run by Dr Veer Math in the Molecular Biology Support Unit, IBLS, University of Glasgow. Sequencing of overexpression and mutant constructs was carried out by Genome Express, Grenoble, France. Plasmid DNA for sequencing was prepared using a Qiagen DNA isolation kit. (See Appendix V).

The initial cycling reaction was performed on a DNA Thermal Cycler Model 480 with 0.5µg plasmid DNA and 3.2pmoles primer in a total reaction volume of 20µl. The reaction involved 25 cycles of 96°C for 30secs, 50°C for 15secs and 60°C for 4mins with a rapid thermal ramp at all steps.

The results were analysed on ABI's Sequence Navigator computer program and GCG with Wisconsin Package 8.1-Unix.

2.7 Overexpression analysis

2.7.1 Overexpression studies

A 10ml overnight culture inoculated in LB was grown overnight at 30°C. A 100ml culture was inoculated with 5ml of this overnight culture. This was then grown to an A_{600} of 0.6 before inducing expression with the addition of IPTG to a final concentration of 0.4mM. Samples were removed at various time points and the A_{600} measured and plotted. 1ml samples were pelleted and the cells resuspended in 10µl of sample buffer for each 0.1A unit. These were then subjected to SDS PAGE analysis to look for a band of overexpressed protein.

2.7.2 SDS- polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel electrophoresis was performed by the method of Laemmli (1970), with a 5% stacking gel and a 15% running gel. The ratio of acrylamide to bisacrylamide in all PAGE experiments was 30: 0.8 and polymerisation was induced by the addition of 0.03% (v/v) TEMED and 0.05% (w/v) ammonium persulphate. Samples were denatured by boiling for 5 minutes after dilution in SDS-PAGE sample buffer (Table 2.2). After electrophoresis the gels were stained for protein by the Coomassie method.

2.7.3 Staining for protein

Gels were stained for protein with Coomassie blue for 40 minutes at 40°C. The Coomassie reagent consisted of 0.1% (w/v) Coomassie brilliant blue G250 in 50% (v/v) methanol, 10% glacial acetic acid; destaining was carried out in 10%

Buffer	**Constituents**
Buffer A	50mMTris-HCl pH 7.5 0.4mM DTT 1 protease inhibitor tablet per litre of buffer
Resuspension Buffer (for breaking cells)	Buffer A plus 1 protease inhibitor tablet per 50ml

Table 2.7 Buffers for use in the purification of the HisIE protein

methanol (v/v), 10% glacial acetic acid at 40°C until the background was fully destained and the bands visible.

2.8 Purification of HisIE protein

2.8.1 Growth of cells for protein purification

The *E. coli* overexpression construct BL21(DE3)pLysS-pSACIE25 was inoculated from a plate into LB containing tetracycline and chloramphenicol and grown overnight at 30°C. The following morning nine flasks each containing 500ml of LB plus antibiotics were inoculated with 40ml of this overnight culture and growth continued at 30°C to an A_{600} of 0.5-0.6. The cells were then induced by adding IPTG to a final concentration of 0.4mM and grown for a further 4-5 hours before harvesting in a MSE 2L centrifuge. The cell pellet was stored at -20°C until required.

2.8.2 Protein purification

Cells of *E. coli* strain BL21(DE3)pLysS-pSACIE25 (13.5g wet weight) were thawed slowly on ice and resuspended in 20ml Resuspension buffer. The cell paste was subjected to two passes through the French Press at 1000psi and then diluted to 50ml with Resuspension buffer. 0.5mg of DNAseI was added and the extract stirred at 4°C for 30 minutes before centrifuging at 18,000rpm for 1 hour at 4°C.

A 150ml DEAE Sephacel column (4.5 x 10cm) was pre-equilibrated with at least 5 column volumes of Buffer A prior to loading the protein. The column was loaded at 25ml per hour and then washed for 4 hours before eluting with a 700ml linear gradient of 0-500mM KCl in Buffer A; 10ml fractions were collected.

A 70ml Phenyl Sepharose column (2.7 x 12cm) was equilibrated with Buffer A containing 1M ammonium sulphate. The pooled HisIE activity from the DEAE Sephacel column was made to 1M in ammonium sulphate by the slow addition of finely powdered solid and the resulting solution stirred gently for an hour at 4°C and then loaded onto the pre-equilibrated Phenyl Sepharose column at 50ml per hour. The column was washed for at least 4 hours before eluting with a 500ml linear gradient of 1M to 0M ammonium sulphate in Buffer A; 10ml fractions were collected.

The pooled HisIE activity from the Phenyl Sepharose column was dialysed overnight against three changes of Buffer A containing 100mM KCl. It was then concentrated to approximately 2ml in an ultrafiltration cell using a 10K filter. Concentrated protein was applied to a S-200 column (2 x 150cm) that had been pre-equilibrated with buffer A containing 500mM KCl. The column was run at 10ml per hour and 5ml fractions were collected.

The HisIE activity from the S-200 column was pooled and dialysed for two days into Buffer A containing 50%(v/v) glycerol and finally stored at -20°C.

2.9 Generation of enzyme substrate

2.9.1 Generation of Phosphoribosyl ATP

Phosphoribosyl ATP (PR-ATP) was generated enzymatically using a purified HisG extract from *E. coli*. The starting materials for the biotransformation were the two substrates for the HisG enzyme PRPP and ATP. 244mg PRPP (0.5mmoles; Sigma 80% pure, formula weight 390.1) and 276mg of ATP (Sigma, disodium salt formula weight 551.1) were dissolved in 280ml of water and 1M Tris-HCl pH8.5 and 1M $MgCl_2$ added to give a final concentration of 100mM Tris-HCl, 10mM $MgCl_2$ in 330ml. The reaction was initiated by the addition of 100μg of inorganic

pyrophosphatase (Sigma) and 5 units of the HisG enzyme. The pH was adjusted to 8.5 with 0.1M NaOH and the reaction was stirred at 37°C for one hour. The progress of the reaction was monitored in a spectrophotometer at 290nm. The biotransformation mixture was adjusted to pH 6.5 at the end of the reaction with 1M HCl and the solution was then frozen and stored at -20°C.

2.9.2 Chromatography of PR-ATP.

A Q-sepharose column (4.5 x 12cm) was equilibrated with 5 column volumes of Buffer 1 (50mM Triethylamine bicarbonate pH 7.4). The frozen biotransformation reaction was thawed on ice and diluted to the same conductivity as the equilibrating buffer by the addition of approximately 20 volumes of Buffer 1. This was then loaded onto the Q-sepharose column and eluted with a 1 litre linear gradient of 0.05-1M Triethylamine bicarbonate pH 7.4; 10ml fractions were collected. The fractions which had absorbance at 290nm were pooled and lyophilized overnight.

2.9.3 NMR spectroscopy

Samples of the PR-ATP at different stages of the purification were analysed by NMR spectroscopy. This work was carried out by Dr David Rycroft, Department of Chemistry, Glasgow University. Lyophilized material from the chromatography steps was resuspended in D_2O and re-lyophilized in order to minimise the water peak present in the proton NMR spectrum. Approximately 2-5mg of material was used for the NMR analysis. The samples were run on a Bruker A300 NMR spectrometer.

2.10 Enzyme assay

2.10.1 *E. coli* Phosphoribosyl ATP transferase (HisG)

The enzyme activity was determined by monitoring, in a 1cm pathlength cuvette, the increase in absorbance at 290nm at 37°C. (PR-ATP ε= 3600 M^{-1} cm^{-1}) (Smith and Ames, 1964). The assay mixture contained 150mM KCl, 100mM Tris-HCl pH 8.5, 10mM $MgCl_2$, 5mM ATP and 0.5mM PRPP. The enzyme assay was carried out in a final volume of 1ml.

2.10.2 HisIE enzyme activity

The activity of the *Arabidopsis thaliana* HisIE enzyme was monitored by measuring the increase in absorbance at 290nm and 25°C in a 1ml cuvette. This corresponds to the formation of Pro-FAR (BBMII) ε= 8000 M^{-1} cm^{-1} (Martin *et al.*, 1971). The assay mixture contained 50mM Tris-HCl pH 8.5, 10mM $MgCl_2$ and 50-100μM PR-ATP (final concentrations).

2.11 Chemical modification experiments

2.11.1 Inactivation of HisIE enzyme with diethylpyrocarbonate (DEPC).

Inactivation was carried out by incubating the HisIE enzyme with DEPC in 10mM potassium phosphate buffer pH 7.2 at 25°C. A 50mM solution of DEPC was freshly prepared before each experiment by diluting the stock reagent with ice cold absolute ethanol. The inactivation reaction was terminated by the addition of 0.5M imidazole hydrochloride pH 7 to a final concentration of 0.5mM.

2.11.2 Inactivation with trinitrobenzenesulfonic acid (TNBS).

Inactivation with TNBS was carried out in the dark at 25°C. The enzyme was incubated in the presence of 10mM potassium phosphate buffer pH 7.2. TNBS was added to a final concentration of 0.1-1mM. The reaction was terminated by the addition of 0.5M lysine to a final concentration of 1mM.

Appendix I Names and Addresses of Suppliers

Bochringer Mannheim/ Roche Diagnostics, Lewes, East Sussex, UK.

Cruachem Ltd, West of Scotland Science Park, Glasgow, Scotland.

Difco Laboratories, Detroit, Michigan, USA.

E. coli Genetic Stock Centre, Dept. of Biology, Yale University, New Haven, CT, USA

Fisher Scientific, Loughborough, Leicestershire, UK.

Genome Express, Zone Astec, Cedex, F37000, Grenoble, France.

Genosys Biotechnologies Inc., Cambridge Science Park, Cambridge, England

Gibco BRL Life Technologies, Inchinnan Business Park, Paisley, Scotland.

Hewlett Packard Ltd, Cheadle Heath, Stockport, Cheshire, UK.

New England Biolabs Inc., Hitchin, Hertfordshire, UK.

Pharmacia Biotech, St Albans, Hertfordshire, England UK.

Phenomenex, Macclesfield, Cheshire, UK.

Promega Corporation, Southampton, UK.

Qiagen Ltd, Crawley, West Sussex, UK.

Sigma Aldrich Company Ltd, Poole, Dorset, UK.

Shimadzu Corporation, Japan.

Supelco Inc, Supelco UK (Sigma Aldrich Company Ltd), Poole, Dorset, UK.

Dr Danielle Werck, IBMP-CNBS, UPR 406, 28 Rue Goethe, F-67000, Strasbourg, France.

Appendix II Wizard DNA Mini Preps Kit (Promega)

This kit allows the rapid isolation of plasmid DNA by an alkaline lysis method. The DNA is bound to a silica based resin in the presence of high salt and following washing to remove contaminants can be eluted in low salt buffer. There is no need for organic solvent extractions and the DNA is of suitable quality for use in DNA sequencing and other molecular biological applications.

Step 1 Pellet 1-5ml of an overnight culture
Step 2 Resuspend cells in resuspension buffer
Step 3 Add cell lysis solution and mix until a clear lysate is formed
Step 4 Neutralise by adding potassium acetate
Step 5 Spin to pellet insoluble material for 5 minutes at full speed in a microfuge
Step 6 Add 1ml of the Wizard DNA Purification Resin® to the syringe barrel/column
Step 7 Load supernatant onto Wizard mini column
Step 8 Draw the material into the column, and wash with 2ml of wash buffer
Step 9 Spin the column in a microfuge to remove any residual buffer and elute DNA with 50µl of water

Cell Lysis Solution	0.2N NaOH
	1% SDS
Cell Resuspension Solution	50mM Tris-HCl pH 7.5
	10mM EDTA
	100µg/ml RNase A
Column Wash Solution	190mM Potassium acetate
	20mM Tris-HCl pH 7.5
	1mM EDTA

Solution is diluted with 95% ethanol before use. Final ethanol concentration 55%.

Neutralisation Solution	1.32M potassium acetate pH 4.8

Appendix III Wizard DNA Cleanup kit (Promega)

This kit can be used as an alternative to organic solvent extractions and ethanol precipitations for the purification of DNA from contaminating restriction enzymes or phosphatases.

The DNA is bound to a silica based resin in the presence of high salt and is eluted in low salt buffer in a similar manner to the previous protocol.

Step 1 Take restriction digest and add 1ml of Wizard DNA Cleanup Resin.

Step 2 Mix gently by inversion and load onto the syringe barrel /column

Step 3 Wash with 2ml of 80% isopropanol

Step 4 Spin column dry and then elute with 50µl of water

Appendix IV Wizard PCR Preps DNA purification system

Allows the purification of double stranded PCR amplified DNA from contaminating nucleotides, enzymes and primer dimers etc. The direct purification method can be used when there is only a single PCR product in the reaction. The presence of other amplification products requires gel purification of the band of interest.

Purification of PCR product from an agarose gel

Step 1 Separate PCR products by electrophoresis on a TAE low melting point agarose gel and excise the relevant band
Step 2 Incubate the sample at 70°C to melt the agarose
Step 3 Add 1ml of the DNA Purification Resin to the sample and mix thoroughly for 20 seconds
Step 4 Add the resin to the syringe barrel /column and draw through
Step 5 Wash with 2ml of 80% isopropanol and spin in a microfuge to dry
Step 6 Elute the DNA in 50µl of water

If purifying the DNA direct from the PCR reaction steps 1-3 are replaced with the following;

Step 1a For each PCR reaction transfer the aqueous layer to a fresh tube
Step 2a Add 100µl of Direct Purification Buffer and vortex briefly to mix
Step 3a Add 1ml of resin and vortex briefly 3 times over the period of 1 minute

Direct Purification Buffer 50mM KCl
10mM Tris-HCl (pH 8.8 at 25°C)
1.5mM $MgCl_2$
0.1% Triton® X-100

Appendix V Qiagen Midi kit

This protocol is again based on an alkaline lysis method and using a silica based column to bind the plasmid DNA and allow separation from contaminants.

Step 1 Pellet 20-100ml of an overnight culture

Step 2 Resuspend pellet in buffer P1 (Resuspension buffer)

Step 3 Add buffer P2 (Lysis buffer) and mix gently by inversion

Step 4 Add buffer P3 (Neutralisation buffer), mix gently and incubate on ice for 15 minutes

Step 5 Centrifuge at 15,000rpm for 30 minutes to clear lysate

Step 6 Equilibrate the Qiagen tip and add the cleared supernatant and allow to flow by gravity

Step 7 Wash the Qiagen tip with buffer QC

Step 8 Elute DNA with buffer QF

Step 9 Precipitate DNA with 0.7 volumes of isopropanol and centrifuge for 30 minutes at 4°C

Step 10 Wash DNA with cold 70% ethanol and air dry for 5 minutes and resuspend in water

Appendix VI The IMPACT™ T7 One Step Purification System-Intein system (New England Biolabs Inc)

Impact™ (Intein Mediated Purification with an Affinity Chitin binding Tag) is a new method for obtaining pure protein in a single chromatographic step in 24 hours. This system involves a protein splicing element called an intein which has been isolated from *Saccharomyces cerevisiae* (Perler *et al.*, 1994, Kane *et al.*, 1990)). The element is modified so that it can undergo self cleavage at the N terminus in the presence of thiols e.g. DTT (Chong et al., 1996,1997). The protein of interest is inserted as an in frame fusion with the intein which has itself been linked to a chitin binding tag. The protein is produced as a fusion protein which is linked to this tag. The cell extract is passed through a chitin column and the protein of interest should be the only one to bind. All other contaminants are washed through the column. The column is incubated overnight with DTT at 4°C which stimulates the intein mediated self cleavage to release the target protein.

Chapter 3

Cloning and overexpression of the hisIE cDNA

3.1 Introduction

The technique of cloning by complementation is a well established technique for the cloning of cDNAs from cDNA libraries providing that a suitable mutant is available. The mutant allows for the selection of positive clones that have been rescued and are able to grow on selective media. Genes that are involved in essential metabolic pathways are by far some of the easier targets to isolate. Here we can deprive the cells of the end product of the pathway, for example by removing an amino acid such as methionine from the medium and, providing the mutants have acquired the missing gene by complementation, and are able to express that gene, they should be able to produce methionine and thus survive in the presence of the minimal media. *E. coli* and *Saccharomyces cerevisiae* have proved to be useful model organisms for complementation studies (Delorme *et al*, 1992). There are now extensive collections of mutants available for both organisms. There are auxotrophic mutants available for the histidine biosynthetic genes for both *E. coli* and *Saccharomyces cerevisiae*. We obtained auxotrophic mutants for each of the genes involved in histidine biosynthesis from the Coli Genetic Stock Centre.

In order to isolate the cDNAs for the plant enzymes, we attempted to complement *E. coli* auxotrophic mutants for the individual genes with cDNA libraries generated from *Arabidopsis* and wheat. This process had been used previously to isolate and express the *Lactococcus lactis hisG* gene in *E. coli* (Delorme *et al*, 1992). Of the three libraries screened, the λYES library was the most extensively screened. The λYES library is a Yeast *E. coli* Shuttle vector (Elledge, 1991) containing *Arabidopsis* cDNA inserts. This vector system allows expression of the cDNA inserts in both species and is dependant on the orientation of the insert in relation to both

promoters. This library was provided by Dr Nigel Urwin from the Division of Biochemistry and Molecular Biology, Glasgow University. The library was constructed from stems and leaves of *Arabidopsis* plants at various stages of growth.

The other libraries used for complementation of the mutants were λPRL libraries, one from *Arabidopsis* and another from wheat. Both of these libraries were provided by Dr Danielle Werck, IBMP-CNBS, Strasbourg, France. The λPRL libraries are λZIPLOX derivatives (λ expression vectors), that allow excision of plasmid DNA containing cDNA inserts (Lin *et al.*, 1992). The *E. coli* mutants were screened for complementing cDNAs using plasmid DNA.

3.2 Library screening

3.2.1 Initial screening

Initial work to screen the λYES cDNA library was carried out by Miss Edith Gould who established initial screening conditions and isolated a number of potential clones. Initial screening was performed as described in the materials and methods section. The aim was to screen for cDNAs for the first three enzymes of the pathway. These were all performed simultaneously using the three mutants which were transformed individually and then selected for on selective minimal media plates. Multiple attempts at complementation were carried out with varying degrees of success.

A critical discovery was made when I attempted to repeat the complementation experiments with the plasmid DNA isolated by Miss Gould which may have influenced some of the results she interpreted as positive results. Following transformation she

was resuspending the *E. coli* in LB before plating onto minimal media plates. This was providing a highly nutritious background on which the mutants were able to grow. This may explain the number of positive colonies that Edith obtained especially when retransforming with isolated plasmid. From this point on all mutant *E. coli* were resuspended in 1X Salt solution and then plated onto minimal media.

Because of this observation the λYES library was rescreened using the modified protocol.

3.2.2 Initial transformation results

Each of the mutants was plated from a glycerol stock onto selective minimal media in the presence and absence of histidine to ensure that each was auxotrophic for histidine i.e. they were only able to grow in the presence of histidine. Competent cells were then prepared as described in section 2.4.8. These cells were then transformed with small aliquots of one of the plasmid cDNA libraries. All the plasmids contained a gene for antibiotic resistance (ampicillin) as a marker which was also used to aid selection.

Following transformation with the libraries the cells were plated onto minimal media in the presence and absence of histidine, and ampicillin was added to select for those cells containing the plasmid. The plates were placed at 37°C overnight but no growth was observed after 20 hours incubation. The plates were therefore returned to the incubator for a further 24 hours. After approximately 40 hours incubation at 37°C colonies were observed on a few of the plates. The colonies, that grew on minimal

Table 3.1 Clones isolated from complementation experiments

Gene	Clone name	Library	Insert size	Complementation following retransformation	Sequence checked	Comments
hisA	A1-2	λYES	1.6kb	no	no	
hisA	A1	λPRL	1kb	yes	yes	sequence identical to HisIE
hisA	A4	λPRL	1kb	yes	yes	sequence identical to HisIE
hisA	A6	λPRL	1kb	no	yes	sequence identical to HisIE
hisIE	IE1	λYES	1.1kb	yes	yes	
hisIE	IE2	λYES	1.1kb	yes	yes	
hisIE	IE15	λYES	1.1kb	yes	yes	DNA template for all other manipulations
hisIE	IE20	λPRL	1.1kb	yes	yes	
hisG	G3	λPRL	1kb	no	yes	sequence identical to HisIE
hisG	G4	λPRL	1kb	no	yes	sequence identical to HisIE
hisG	G5	λPRL	1kb	yes	yes	sequence identical to HisIE

media in the absence of histidine, were assumed to contain DNA which was able to complement the missing gene in that mutant.

3.2.3 Isolation of plasmids

Colonies were isolated from the minimal media plates and plasmid DNA was prepared. The DNA was then analysed by restriction analysis for the presence of an appropriately sized insert. The DNA was digested with the restriction enzyme EcoRI revealing insert fragments of approximately 1-1.1Kb in length (Table 3.1).

3.2.4 Retransformation

To confirm that the isolated clones were able to complement the mutants a second round of screening was carried out. This involved re-transforming the mutants with the isolated plasmid and then subsequent selection on minimal media containing ampicillin in the presence and absence of histidine. Those clones which were able to produce a significant number of colonies on the plates lacking histidine were deemed to be viable in that they were able to complement the auxotrophy.

Clones complementing the *hisIE* mutant were isolated as well as some clones for hisG and hisA. Plasmid DNA was isolated from the complemented clones and was re-transformed into the appropriate mutant. Those plasmids which appeared to be able to complement were sequenced. This revealed something strange. The inserts in the plasmid DNA isolated from the hisG and hisA clones had the same sequence as the insert obtained from the hisIE clones (Table 3.1). One initial theory to explain this was that the hisIE cDNA encoded all three enzyme activities. This theory was dismissed however once the entire hisIE clone had been sequenced, since it was homologous only to the microbial *hisIE* genes and contained no additional DNA homologous to the

microbial *hisG* and *hisA* genes. A second theory was that the HisIE protein was able to provide a way to rescue these mutants by other means. Both the *hisG* and *hisA* mutants are point mutations that have not been characterised at the DNA level. It is possible that there is some association of the enzyme activities *in vivo* and that the presence of extra HisIE protein can "repair" the deficiency for one of these activities.

In the case of *hisG* there would need to be a restoration of the ability to make PR-ATP, while in the case of *hisA* the activity required is a sugar isomerase activity. It is possible that the HisIE protein which has a binding site for PR-ATP might be able to bind PRPP and ATP at adjacent sites and catalyse their slow condensation. The reaction catalysed by the HisA protein is an Amadori rearrangement (an isomerisation); this reaction does proceed very slowly in the absence of enzyme. The presence of extra HisIE protein would cause greater than normal levels of substrate for HisA to accumulate and this might allow sufficient spontaneous reaction to occur and result in sufficient flux through the pathway to allow for some bacterial growth and hence the observation of complementation.

3.2.5 Sequencing of the hisIE clone

The clone for hisIE was fully sequenced by automated thermal cycle sequencing. The clones were sequenced using the oligonucleotides listed in Table 2.4. The DNA sequence of the forward strand of the clone is presented in Figure 3.1. A translation of this sequence revealed a protein of 281 amino acids in length. This was compared to other known sequences for *hisIE* gene products and appears to line up very well. (Figure 3.2). Many of the highly conserved residues across all species are

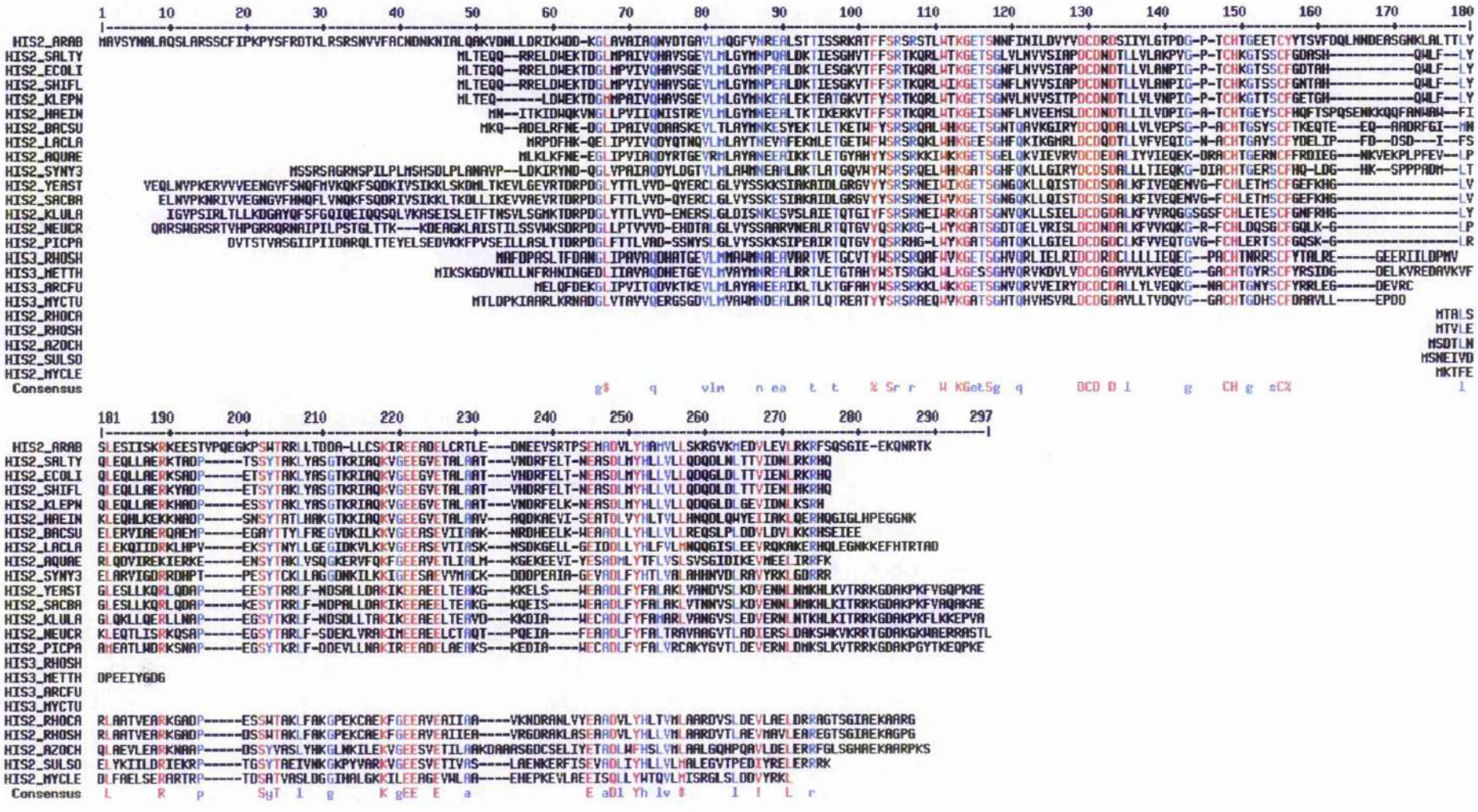
Figure 3.1 DNA sequence of *Arabidopsis thaliana* hisIE clone and the amino acid sequence predicted from the longest open reading frame

```
   1 AGCTACGTCCGGGGCCACTTTTCGTTTCTCGTAGCGTTAAAATGGCGGTATCGTACAATG
                                                 M  A  V  S  Y  N  A    7
  61 CATTAGCTCAGTCTTTAGCGAGAAGTAGCTGCTTCATCCCCAAACCTTATTCCTTTAGAG
     L  A  Q  S  L  A  R  S  S  C  F  I  P  K  P  Y  S  F  R  D   27
 121 ATACTAAGCTGAGAAGCAGATCCAATGTCGTATTCGCGTGCAATGATAATAAGAACATTG
     T  K  L  R  S  R  S  N  V  V  F  A  C  N  D  N  K  N  I  A   47
 181 CTCTTCAAGCTAAGGTAGATAACTTGTTGGACCGCATTAAATGGGATGACAAAGGATTAG
     L  Q  A  K  V  D  N  L  L  D  R  I  K  W  D  D  K  G  L  A   67
 241 CTGTGGCAATAGCACAAAACGTTGATACGGGAGCAGTATTGATGCAAGGCTTTGTTAATA
     V  A  I  A  Q  N  V  D  T  G  A  V  L  M  Q  G  F  V  N  R   87
 301 GGGAGGCCCTCTCCACAACCATCAGTTCTCGGAAAGCTACATTCTTTAGTCGATCAAGAT
     E  A  L  S  T  T  I  S  S  R  K  A  T  F  F  S  R  S  R  S  107
 361 CTACCTTATGGACTAAGGGAGAGACATCCAATAACTTCATCAATATTCTTGATGTGTATG
     T  L  W  T  K  G  E  T  S  N  N  F  I  N  I  L  D  V  Y  V  127
 421 TTGATTGTGATCGTGATTCGATTATTTACCTTGGAACACCTGATGGACCTACCTGTCACA
     D  C  D  R  D  S  I  I  Y  L  G  T  P  D  G  P  T  C  H  T  147
 481 CAGGGGAAGAGACTTGTTACTACACATCGGTTTTTGATCAATTAAACAATGATGAGGCTT
     G  E  E  T  C  Y  Y  T  S  V  F  D  Q  L  N  N  D  E  A  S  167
 541 CAGGAAACAAGCTAGCATTAACAACATTGTACTCGCTAGAATCAATCATTTCCAAGCGGA
     G  N  K  L  A  L  T  T  L  Y  S  L  E  S  I  I  S  K  R  K  187
 601 AAGAAGAATCAACAGTTCCTCAAGAAGGTAAACCATCATGGACTCGACGGTTGTTGACGG
     E  E  S  T  V  P  Q  E  G  K  P  S  W  T  R  R  L  L  T  D  207
 661 ATGACGCTCTGCTTTGCTCAAAGATCAGGGAAGAAGCTGACGAGTTATGCAGAACACTGG
     D  A  L  L  C  S  K  I  R  E  E  A  D  E  L  C  R  T  L  E  227
 721 AGGATAATGAGGAAGTTTCAAGAACACCATCAGAGATGGCTGATGTTTTATACCACGCAA
     D  N  E  E  V  S  R  T  P  S  E  M  A  D  V  L  Y  H  A  M  247
 781 TGGTGCTTCTATCTAAAAGGGGTGTGAAGATGGAAGATGTTCTTGAAGTTCTTAGGAAAC
     V  L  L  S  K  R  G  V  K  M  E  D  V  L  E  V  L  R  K  R  267
 841 GCTTCTCTCAATCTGGAATCGAGGAGAAGCAAAACCGTACAAAGTAACCATTTTCTTCGA
     F  S  Q  S  G  I  E  E  K  Q  N  R  T  K  *                  281
 901 TTGTCTGTTTAGTATTTGCGAGAATTTTTTTAGTTACTCGGAACTTATACTGTTGTTCAT
```

Figure 3.2 Alignment of known HisIE protein sequences including the predicted HisIE protein isolated from *Arabidopsis.*

Red- Highly conserved, Blue- Moderately conserved

ARAB	*Arabidopsis thaliana*
SALTY	*Salmonella typhimurium*
ECOLI	*Escherichia coli*
SHIFL	*Shigella flexnieri*
KLEPN	*Klebsiella pneumoniae*
HAEIN	*Haemophilus influenzae*
BACSU	*Bacillus subtilis*
LACLA	*Lactococcus lactis*
AQUAE	*Aquifex aeolicus*
SYNY3	*Synechocystis sp. Strain PCC6803*
YEAST	*Saccharomyces cerevisiae*
SACBA	*Saccharomyces bayanus*
KLULA	*Kluyveromyces lactis*
NEUCR	*Neurospora crassa*
PICPA	*Pichia pastoris*
RHOSH	*Rhodobacter sphaeoides*
METTH	*Methanococcus thermoautotrophicum*
ARCFU	*Archaeoglobus fulgidus*
MYCTU	*Mycobacterium tuberculosis*
RHOCA	*Rhodobacter capsulatus*
AZOCH	*Azotobacter chroococcum*
SULSO	*Sulfolobus solfataricus*
MYCLE	*Mycobacterium leprae*


1 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
HIS2_ARAB MAVSYNALAQSLARSSCFIPKPYSFRDTKLRSRSNVVFACNDNKNIALQAKVDNLLDRIKWDD-KGLAVAIAQNVDTGAVLMQGFVNREALSTTISSRKATFFSRSRSTLWTKGETSNNFINILDVYVDCDRDSIIYLGTPDG-P-TCHTGEETCYYTSVFDQLNNDEASGNKLALTTLY
HIS2_SALTY MLTEQQ--RRELDWEKTDGLMPAIVQHAVSGEVLMLGYMNPQALDKTIESGHVTFFSRTKQRLWTKGETSGLVLNVVSIAPDCDNDTLLVLAKPVG-P-TCHKGTSSCFGDASH----------QWLF--LY
HIS2_ECOLI MLTEQQ--RRELDWEKTDGLMPVIVQHAVSGEVLMLGYMNPEALDKTLESGKVTFFSRTKQRLWTKGETSGNFLNVVSIAPDCDNDTLLVLANPIG-P-TCHKGTSSCFGDTAH----------QWLF--LY
HIS2_SHIFL MLTEQQ--RRELDWEKTDGLMPVIVQHAVSGEVLMLGYMNPEALDKTIESGKVTFFSRTKQRLWIKGETSGNFLNVVSIAPDCDNDTLLVLANPIG-P-TCHKGTSSCFGNTAH----------QWLF--LY
HIS2_KLEPN MLTEQ------LDWEKTDGMMPAIVQHAVSGEVLMLGYMNKEALEKTEATGKVTFYSRTKQRLWTKGETSGNVLNVVSITPDCDNDTLLVLVNPIG-P-TCHKGTTSCFGETGH----------QWLF--LY
HIS2_HAEIN MN--ITKIDWQKVNGLLPVIIQNISTREVLMLGYMNEEALTKTIKERKVTFFSRTKQRLWTKGEISGNFLNVEEMSLDCDNDTLLILVDPIG-A-TCHTGEYSCFHQFTSPQSENKKQQFANWAW--FI
HIS2_BACSU MKQ--ADELRFNE-DGLIPAIVQDARSKEVLTLAYMNKESYEKTLETKETWFYSRSRQALWMKGETSGNTQAVKGIRYDCDQDALLVLVEPSG-P-ACHTGSYSCFTKEQTE---EQ--AADRFGI--MN
HIS2_LACLA MRPDFHK-QELIPVIVQDYQTNQVLMLAYTNEVAFEKMLETGETWFWSRSRQKLWMKGEESGHFQKIKGHRLDCDQDTLLVFVEQIG-N-ACHTGAYSCFYDELIP----FD--DSD----I--FS
HIS2_AQUAE MLKLKFNE-EGLIPVIAQDYRTGEVRMLAYANEEAIKKTLETGYAHYYSRSRKKIWKKGETSGELQKVIEVRVDCDEDALIYVIEQEK-DRACHTGERNCFFRDIEG---NKVEKPLPFEV--LP
HIS2_SYNY3 MSSRSAGRNSPILPLMSHSDLPLANAVP--LDKIRYND-QGLVPAIAQDYLDGTVLMLAHMNEAALAKTLATGQVWYWSRSRQELWHKGATSGHFQKLLGIRYDCDSDALLLTIEQKG-DIACHTGERSCFHQ-LDG----HK--SPPPADM--LT
HIS2_YEAST VEQLNVPKERVVVEENGVFSNQFMVKQKFSQDKIVSIKKLSKDMLTKEVLGEVRTDRPDGLYTTLVVD-QYERCLGLVYSSKKSIAKRIDLGRGVYYSRSRNEIWIKGETSGNGQKLLQISTDCDSDALKFIVEQENVG-FCHLETHSCFGEFKHG----------LV
HIS2_SACBA ELNVPKNRIVVEGNGVFHNQFLVNQKFSQDRIVSIKKLTKDLLIKEVVAEVRTDRPDGLFTTLVVD-QYERCLGLVYSSKESIAKRIDLGRGVYYSRSRNEIWVKGETSGNGQKLLQISTDCDSDALKFIVEQENVG-FCHLETHSCFGEFKHG----------LV
HIS2_KLULA IGVPSIRLTLLKDGAYQFSFGQIQEIQQSQLVKASEISLETFTNSVLSGMKTDRPDGLYTTLVVD-EMERSLGLDISNKESVSLAIETQTGIYFSRSRNEIWRKGATSGNVQKLLSIELDCDGDALKFVVRQGGSGSFCHLETESCFGNFRHG----------LY
HIS2_NEUCR QARSHGRSRTVHPGRRQRNARIPILPSTGLTTK---KDEAGKLAISTILSSVWKSDRPDGLLPTVVVD-EHDTALGLVYSSAARVNEALRTQTGVYQSRKRG-LWYKGATSGDTQELVRISLDCDNDALKFVVKQKG-R-FCHLDQSGCFGQLK-G----------LP
HIS2_PICPA DVTSTVASGIIPIIDARQLTTEYELSEDVKKFPVSEILLASLTTDRPDGLFTTLVAD-SSNYSLGLVYSSKKSIPEAIRTQTGVYQSRHG-LWYKGATSGATQKLLGIELDCDGDCLKFVVEQTGVG-FCHLERTSCFGQSK-G----------LR
HIS3_RHOSH MAFDPASLTFDANGLIPAVAQDHATGEVLMMAWMNAEALARTVETGCVTYWSRSRQAFWVKGETSGHVQRLIELRIDCDRDCLLLLIEQEG--PACHTNRRSCFYTALRE-----GEERIILDPMV
HIS3_METTH MIKSKGDVNILLNFRHNINGEDLIIAVYAQDHETGEVLMVAYMNREALRRTLETGTAHYWSTSRGKLWLKGESSGHVQRVKDVLVDCDGDAVVLKVEQEG--GACHTGYRSCFYRSIDG-----DELKVREDAVKVF
HIS3_ARCFU MELQFDEKGLIPVITQDVKTKEVLMLAYANEEAIKLTLKTGFAHYWSRSRKKLWKKGETSGNVQRVVEIRYDCDCDALLYLVEQKG--NACHTGNYSCFYRRLEG-----DEVRC
HIS3_MYCTU MTLDPKIAARLKRNADGLVTAVVQERGSGDVLMVAWMNDEALARTLQTREATYYSRSRAEQWVKGATSGHTQHVHSVRLDCDGDAVLLTVDQVG--GACHTGDHSCFDAAVLL-----EPDD
HIS2_RHOCA MTALS
HIS2_RHOSH MTVLE
HIS2_AZOCH MSDTLN
HIS2_SULSO MSNEIVD
HIS2_MYCLE MKTFE
Consensus g$ q vln n ea t t % Sr r W KGetSg q DCD D l g CH g sC% l
181 190 200 210 220 230 240 250 260 270 280 290 297
HIS2_ARAB SLESIISKRKEESTVPQEGKPSWTRRLLTDDA-LLCSKIREEADELCRTLE---DNEEVSRTPSEMADVLYHAMVLLSKRGVKMEDVLEVLRKRFSQSGIE-EKQNRTK
HIS2_SALTY QLEQLLAERKTADP------TSSYTAKLYASGTKRIAQKVGEEGVETALAAT---VNDRFELT-NEASDLMYHLLVLLQDQDLNLTTVIDNLRKRHQ
HIS2_ECOLI QLEQLLAERKSADP------ETSYTAKLYASGTKRIAQKVGEEGVETALAAT---VHDRFELT-NEASDLMYHLLVLLQDQGLDLTTVIENLRKRHQ
HIS2_SHIFL QLEQLLAERKYADP------ETSYTAKLYASGTKRIAQKVGEEGVETALAAT---VHDRFELT-NEASDLMYHLLVLLQDQDLDLTTVIENLHKRHQ
HIS2_KLEPN QLEQLLAERKHADP------ESSYTAKLYASGTKRIAQKVGEEGVETALAAT---VNDRFELK-NEASDLMYHLLVLLQDQGLDLGEVIDNLKSRH
HIS2_HAEIN KLEQHLKEKKNADP------SNSYTATLHAKGTKKIAQKVGEEGVETALAAV---AQDKAEVI-SEATDLVYHLTVLLHNQDLQHYEIIAKLQERHQGIGLHPEGGNK
HIS2_BACSU ELERVIAERQAEMP------EGAYTTYLFREGVDKILKKVGEEASEVIIAAK---NRDHEELK-WEADDLLYHLLVLLREQSLPLDDVLDVLKKRHSEIEE
HIS2_LACLA ELEKQIIDRKLHPV------EKSYTNYLLGEGIDKVLKKVGEEASEVTIASK---NSDKGELL-GEIDDLLYHLFVLMNQQGISLEEVRQKAKERHQLEGNKKEFHTRTAD
HIS2_AQUAE RLQDVIREKIERKE------ENSYTAKLVSQGKERVFQKFGEEAVETLIALM---KGEKEEVI-YESADMLYTFLVSLSVSGIDIKEVMEELIRRFK
HIS2_SYNY3 ELARVIGDRRDHPT------PESYTCKLLAGGDMKILKKIGEESAEVYMACK---DDDPEAIA-GEVADLFYHTLVALAHHNVDLRAVYRKLGDRRR
HIS2_YEAST GLESLLKQRLQDAP------EESYTRRLF-NDSALLDAKIKEEAEELTEAKG---KKELS----WEAADLFYFALAKLVAMDVSLKDVENNLNMKHLKVTRRKGDAKPKFVGQPKAE
HIS2_SACBA GLESLLKQRLQDAP------KESYTRRLF-NDPALLDAKIKEEAEELTEAKG---KQEIS----WEAADLFYFALAKLVTMNVSLKDVENNLNMKHLKITRRKGDAKPKFVAQAKAE
HIS2_KLULA GLQKLLQERLLNAP------EGSYTKRLF-NDSDLLTAKIKEEAEELTEAVD---KKDIA----WECADLFYFAMARLVANGVSLEDVERNLNTKHLKITRRKGDAKPKFLKKEPVA
HIS2_NEUCR KLEQTLISRKQSAP------EGSYTARLF-SDEKLVRAKIMEEAEELCTAQT---PQEIA----FEAADLFYFALTRAVAAGVTLADIERSLDAKSWKVKRRTGDAKGKWAERRRASTL
HIS2_PICPA AMEATLWDRKSNAP------EGSYTKRLF-DDEVLLNAKIREEADELAEAKS---KEDIA----WECADLFYFALVRCAKYGVTLDEVERNLDMKSLKVTRRKGDAKPGYTKEQPKE
HIS3_RHOSH
HIS3_METTH DPEEIYGDG
HIS3_ARCFU
HIS3_MYCTU
HIS2_RHOCA RLAATVEARKGADP------ESSWTAKLFAKGPEKCAEKFGEEAVEAIIAA----VKNDRANLVYEAADVLYHLTVMLAARDVSLDEVLAELDRRAGTSGIAEKAARG
HIS2_RHOSH RLAATVEARKGADP------DSSWTAKLFAKGPEKCAEKFGEEAVEAIIEA----VRGDRAKLASEAADVLYHLLVMLAARDVTLAEVMAVLEAREGTSGIAEKAGPG
HIS2_AZOCH QLAEVLEARKNAAP------DSSYVASLYHKGLNKILEKVGEESVETILAAKDAAARSGDCSELIYETADLWFHSLVMLAALGQMPQAVLDELERRFGLSGHAEKAARPKS
HIS2_SULSO ELYKIILDRIEKRP------TGSYTAEIVNKGKPYVARKVGEESVETIVAS----LAENKERFISEVADLIYHLLVLMALEGVTPEDIYRELERRRK
HIS2_MYCLE DLFAELSERARTRP------TDSATVASLDGGIHALGKKILEEAGEVWLAA----EHEPKEVLAEEISQLLYWTQVLMISRGLSLDDVYRKL
Consensus L R p SgT l g K gEE E a E aDl Yh lv $ l ! L r

also highly conserved in the *Arabidopsis* enzyme. The sequence showed 70% identity and similarity between the plant and the bacterial sequences and 50% between the plant and the fungal species. Comparison of the sequence with the sequences of the monofunctional proteins from *Rhodobacter sphaeroides* (Oriol *et al.*, 1996) and *Methanococcus jannaschii* (Bult *et al.*, 1996) suggests that the *Arabidopsis* enzyme had two functional domains one containing the HisI and the other the HisE activities. There is also an N-terminal extension not seen in the bacterial proteins suggesting the presence of a chloroplastic targeting sequence. This correlates with the theory that the enzymes of histidine biosynthesis are found in the chloroplast. The sequence represented here was subsequently confirmed by the work of Fujimori *et al.*, (1998).

3.2.6 Chloroplast targeting sequences

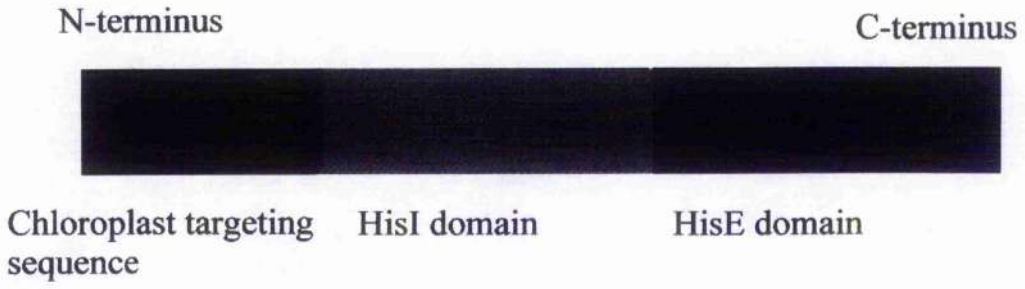
There appears to be no standard amino acid sequence for a chloroplastic targeting sequence. It is believed that the transit peptide, as it is often called, is composed of three regions. An amino terminal region that is devoid of proline, glycine or charged residues followed by a central region containing several hydroxylated and basic residues and finally a C-terminal region which is predicted as β sheet (Filho *et al.*, 1996). From the sequence a pattern similar to this can be seen within the first fifty amino acids which is believed to be the targeting sequence (Figure 3.4). A secondary structure prediction for this protein (produced by Dr A. Lapthorn using the PHD [Profile fed neural network systems from HeiDelberg] program, EMBL, Heidelberg, Germany) (Rost and Sander, 1993 [a, b]) has also revealed a probable β sheet within the C-terminal portion of the first fifty amino acids (Figure3.3). It can be observed from this secondary structure prediction that the HisI domain contains a mixture of

Figure 3.3 Secondary structure prediction of the deduced amino acid sequence of the *Arabidopsis thaliana* HisIE protein

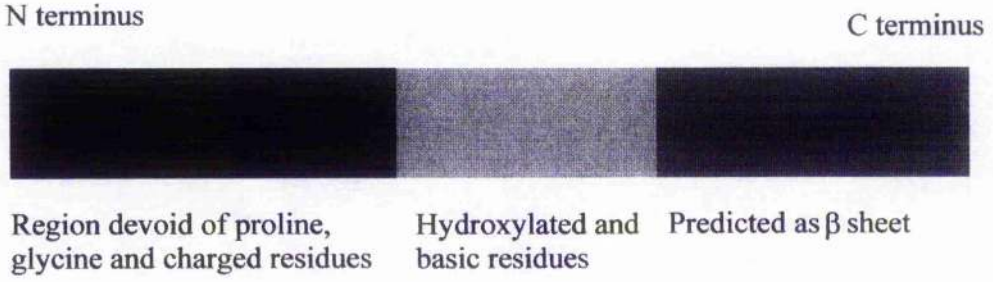
Secondary structure prediction generated using PHD program EMBL, Heidelberg, Germany.

AA- amino acid, prH- probability of helix, prE- probability of extended helix, prL- probability of loop


AA
PHD sec
Rel sec
detail:
prH sec
prE sec
prL sec
subset: SUB sec
accessibility
3st: P_3 acc
10st: PHD acc
Rel acc
subset: SUB acc



a) Simple block diagram showing the three regions that make up the HisIE protein from *Arabidopsis thaliana*



b) Expansion of the region containing the proposed chloroplast targeting sequence of the HisIE protein, and details of its component parts

Figure 3.4 Schematic structure of the *Arabidopsis thaliana* HisIE protein and an expanded view of the components of the proposed targeting sequence

α-helix and β-sheet whereas the HisE domain is predicted to be mainly α-helical (Figures 3.3 and 3.4).

A major aim of this project was to overexpress the *Arabidopsis thaliana* HisIE protein in *E. coli* so that sufficient protein would be available for kinetic characterisation and also for structural and mechanistic studies. Removal of the targeting sequence may well prove to be important for the successful overexpression of this protein in *E. coli*, since the enzymes normally found within the chloroplast to remove the targeting sequence are not present in bacteria. Failure to remove this sequence may result in either incorrect folding of the protein or instability when it is produced in *E. coli*.

3.3 Generation of overexpression constructs

3.3.1 T7 expression vectors

T7 expression systems allow the efficient overexpression of a protein in a controlled manner. The cloning vector has a promoter for a T7 RNA polymerase which is located upstream from the multiple cloning site used to clone the insert DNA. T7 RNA polymerase is able to transcribe DNA five times faster than *E. coli* RNA polymerase and as it is specific for T7 promoters it results in induction of high level expression from such promoters (Chamberlin and Ring, 1973, Golomb and Chamberlin 1974). This allows the specific expression of the gene product of interest. The bacterial strains used for expression purposes can contain a gene for T7 lysozyme which binds to and inhibits T7 RNA polymerase (Moffatt and Studier, 1987). This is used as a means of controlling basal levels of protein expression prior to induction with IPTG. This is particularly important if the protein is toxic to the bacterial strain. The gene for the T7 lysozyme is present on the pLysS plasmid (Studier and Moffatt, 1986), which

also contains a chloramphenicol resistance gene which is transformed separately into the expression host.

3.3.2 Cloning into pGEM expression vector

To facilitate expression of the *Arabidopsis* enzyme in *E.coli* the insert from the λYES clone was inserted into a simple pGEM vector. This vector is a T7 based vector and since there were suitable restriction sites this allowed the digestion of the λYES plasmid to release the insert for direct ligation into the pGEM vector. The cloning was very successful and produced a number of clones which were able to complement the mutant *E. coli*. Small scale expression experiments failed to show bands of the correct molecular weight on SDS-PAGE corresponding to the overexpressed protein. This failure to achieve overexpression in the bacterial system may be due to the presence of the targeting sequence which would be naturally removed upon entering the chloroplast. The presence of this sequence might interfere with the proper folding of the protein and make it susceptible to proteolysis. An alternative explanation may involve the ability of bacterial systems to express a eukaryotic protein. Prokaryotes including *E. coli* use a purine rich Shine Delgarno sequence to recognise the appropriate initiation site for translation on the mRNA. Eukaryotes, however use the AUG codon nearest the 5' end of the mRNA as the start sequence for translation (Stryer, 1988). It may be the case that the *E. coli* machinery required for translation is unable to recognise the appropriate start site for translation on the hisIE mRNA and as a result fails to produce protein of the appropriate size. Starting transcription from the plasmid promoter may add a Shine Delgarno sequence to the mRNA, it may however lie too far from the initiation codon of the hisIE mRNA to allow translation to start.

Clearly in order to achieve overexpression alternative constructs would have to be made.

3.3.3 pTB361 plasmid

The expression plasmid used for the purpose of these experiments is pTB361 which was obtained courtesy of Dr M Horsburgh, University of Glasgow. The T7 promoter is located immediately upstream of the multiple cloning site and the restriction site NdeI is used to clone the 5 prime end. The NdeI site incorporates the ATG start codon of the protein to ensure that the protein is inserted in the correct reading frame as well as to ensure it is in the correct position to obtain maximum overexpression.

3.3.4 Cloning into pTB361

Cloning into the expression vector pTB361 required the generation of restriction sites within the hisIE cDNA to facilitate cloning in the correct orientation and reading frame. The NdeI site was used as the 5 prime site and BamHI was used for cloning the 3 prime end. These two sites were generated by PCR mutagenesis using primers containing the appropriate mutations (Table 3.2).

3.3.5 Generation of two constructs

Since initial experiments failed to show a band of overexpressed protein attempts were made to remove the targeting sequence. Two separate constructs were made one containing the complete open reading frame for the *Arabidopsis* HisIE protein (long construct) and the other containing a shorter sequence lacking the majority of the targeting sequence, but still containing the complete sequence of the

presumed enzyme domains (short construct). The shorter construct required the generation of a new start codon in front of the N terminal sequence NIAL of the protein (Figure 3.5). Both constructs were made using a PCR-based strategy. The question as to whether or not the targeting sequence has been removed in the correct place will only be established later once the native plant enzyme has been isolated and it's N-terminal sequence determined.

The λYES vector containing the hisIE cDNA insert was used as the template for the PCR reactions. The primers used to engineer the NdeI and BamHI sites are given in Table 3.2.

The PCR reaction was carried out as described in section 2.6.2 and a clean single band was observed on a 1% agarose gel in 1X TBE. A fragment of approximately 850bp was observed for the full length "long" construct and a fragment of approximately 700bp was seen for the "short construct" (Figure 3.6).

Oligonucleotide	Sequence
Nde 1	5' GTT TCT CGT AGC GTT CAT ATG GCG GTA TCG 3'
Nde 2	5' GTA TTC GCG TGC AAT GAT CAT ATG AAC ATT G 3'
BamHI	5' CAA TGA ACA ACA GGA TCC GTT CCG AG 3'

Table 3.2 Oligonucleotides used for introduction of restriction sites for the cloning of hisIE cDNA into pTB361

The restriction sites used for cloning purposes have been underlined.

PCR primer for introducing NdeI site into the "long" construct

5′ TTTTCGTTTCTCGTAGCGTTAAAATGGCGGTATCGTACA

5′ GTTTCTCGTAGCGTTCATATGGCGGTATCG 3′
NdeI

M A V S

PCR Primer for introducing a new start codon into the "short" construct

5′ GTATTCGCGTGCAATGATAATAAGAACATTGCT

5′ GTATTCGCGTGCAATGATCATATGAACATTG 3′
NdeI

M N I A

Figure 3.5 PCR primers for the generation of the 5' end of the two HisIE expression constructs

3.3.6 Cloning of the HisIE fragments

The PCR products were purified using the Wizard PCR cleanup kit (Appendix IV) and were digested with both restriction enzymes. NdeI was added first and the digest allowed to proceed for 30 minutes before the addition of BamHI. The DNA was then digested for a further hour. The digestion of the vector pTB361 was carried out in exactly the same way. The plasmid DNA was then dephosphorylated for 40 minutes with CIAP. Both the vector and insert DNA were purified using the Wizard DNA cleanup system (Appendix III) and the resulting material was used to set up ligation reactions which were left overnight at 4°C.

The following morning the ligation reactions were transformed into *E. coli* DH5α cells and plated onto LB-Agar plates containing tetracycline. The plates were grown overnight at 37°C and the number of colonies noted.

3.3.7 Initial Cloning results

Following transformation of DH5α, a number of colonies were observed on each plate. Individual colonies were picked and plasmid DNA prepared from them. The DNA was then analysed by restriction analysis for the presence of an insert of the appropriate size. The results are summarised in Table 3.3.

The clones identified as having inserts of the appropriate size were then transformed into the *hisIE E. coli* to ensure that they could express a functional protein and thus rescue the mutant. The constructs were also sequenced to check that the DNA was in the correct reading frame and that no errors had been incorporated during the PCR process. One of the clones for the short construct minus

Figure 3.6 Agarose gel of the PCR reaction products for the HisIE expression constructs

Markers (sizes are given in bp)

Lane 1	full length hisIE 2mM Mg^{2+}	band at 850bp
Lane 2	full length hisIE 3mM Mg^{2+}	band at 850bp
Lane 3	full length hisIE 4mM Mg^{2+}	band at 850bp
Lane 4	hisIE minus targeting sequence 2mM Mg^{2+}	
Lane 5	hisIE minus targeting sequence 3mM Mg^{2+}	band at 700bp
Lane 6	hisIE minus targeting sequence 4mM Mg^{2+}	band at 700bp

1% agarose gel in TBE buffer containing 0.5μg/ml ethidium bromide.

Name	Construct type	Insert size	Complementation	Sequence check	Expression
pSACIE11	long	850bp	no	no	no
pSACIE 12	long	none	no	no	no
pSACIE 13	long	850bp	yes	yes	no?
pSACIE 14	long	850bp	yes	no	no
pSACIE 15	long	850bp	yes	yes	no?
pSACIE 21	short	none	no	no	no
pSACIE 22	short	none	no	no	no
pSACIE 23	short	700bp	yes	single mutation	yes
pSACIE 24	short	none	no	no	no
pSACIE 25	short	700bp	yes	yes	yes

Table 3.3 Generation of both constructs in pTB361

Summary of the clones obtained

the targeting sequence pSACIE23 was found to contain a mutation at position 151 (See Figure 3.1). The other short clone pSACIE25 was found to be perfect. This was the construct that was used in all further stages of work.

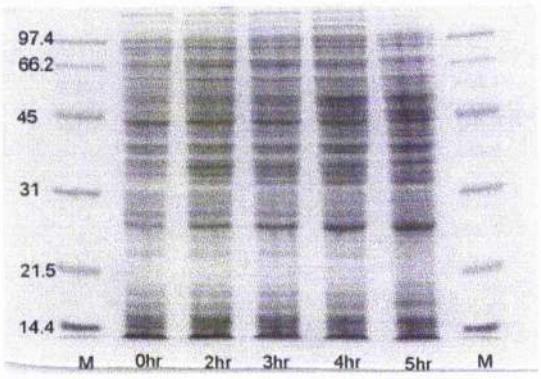
3.3.8 Overexpression analysis

The *E. coli* strain BL21(DE3)pLysS was used for the expression of the hisIE constructs. Both the long and the short constructs were subjected to overexpression analysis. Initial experiments growing strains harbouring the selected recombinant plasmids at 37°C proved to be frustrating, in that they failed to grow at this temperature. The cells did not grow at 37°C and decreasing the concentrations of the antibiotics tetracycline and chloramphenicol did not have any effect. When the temperature was dropped to 30°C it was found that the recombinant strains grew normally. This temperature dependency may be a result of using the BamHI site within the vector. This site lies next to the *cer* gene on the plasmid which is thought to have a role in plasmid stability. As a result of the insert being so close to this gene it may have an effect on the pTB361 plasmid making it less stable at higher temperatures.

3.3.9 Expression

Small scale induction time courses were monitored by removing samples on an hourly basis from a growing culture, harvesting the cells, resuspending in SDS-PAGE sample buffer (2.7.1) and storing frozen at -20°C for future analysis by SDS-PAGE. This was carried out for both the full length and the short construct. The short construct, lacking the targeting sequence, gives rise to a protein species of approximately 31kDa on SDS PAGE, which increases to approximately 30% total cell

Gel A



Gel B

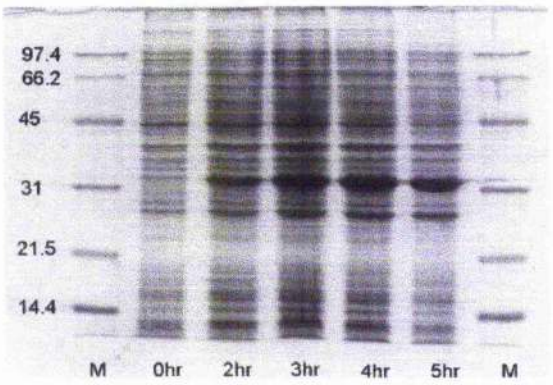


Figure 3.7 SDS PAGE gels showing the overexpression of hisIE constructs at 30°C

Protein expression was induced by the addition of 0.4mM IPTG to a growing culture and samples were removed at various time points. 1ml samples were pelleted and the cells resuspended in 10μl of sample buffer for each 0.1A unit. 5μl of this material was then analysed on a 15% acrylamide gel.

Gel A Long construct showing no evidence of overexpression

Gel B Short construct shows presence of a band of overexpressed protein at 31kDa

Overexpression studies were carried out as described in sections 2.7.1 and 3.3.9. Gels were run and stained as described in sections 2.7.2 and 2.7.3.

protein after 4 hours (Figure 3.7 B). The full length construct, containing the targeting sequence, shows no real evidence of overexpression (Figure 3.7 A). This would indicate that the targeting sequence does interfere with the overexpression of the protein.

As this targeting sequence would naturally be removed it seemed obvious to use the shorter construct which gave high levels of expression of active enzyme suitable for purification and characterisation studies. Therefore all future work relating to protein purification was performed using this short construct, pSACIE25.

3.4 Cloning into a tagged vector system

In order to simplify the purification protocol for the HisIE enzyme an attempt was made to clone the hisIE cDNA into the Impact™T7 cloning system (Appendix VI). This vector system expresses the protein as a fusion protein containing an affinity tag which aids purification. A one step affinity column and a specific cleavage of the fusion protein means that the target protein is purified to homogeneity in a one step process. The aim of this piece of work was to clone the DNA encoding the individual domains of the protein into this vector to simplify the purification of the domains. This was primarily because there may have been problems expressing and purifying the domains from the regular T7 system. It was doubtful if the domains would have folded independently to produce active protein that could be used for characterisation of the domains. The intention was to use this system to generate pure HisIE bifunctional protein for both crystallography trials and also for the generation of antibodies.

Unfortunately this cloning was unsuccessful and it was not possible to obtain clones containing inserts from these experiments.

3.5 Attempts to clone the hisG and hisA cDNA's

3.5.1 Consensus PCR approach

As there was no success in cloning the hisG and hisA cDNAs by complementation of the *E. coli* mutants a PCR based strategy was attempted. Taking known HisG and HisA protein sequences, regions of high homology were determined. To these regions of high homology degenerate oligonucleotides were designed using the codon usage table from *Arabidopsis* as a guide. The oligonucleotides were then used to PCR fragments of the cDNA from the library. This approach yielded a few possible fragments of varying lengths. When these fragments were sequenced however they showed no homology to either a HisA or HisG sequence. All libraries were investigated but with little success.

3.6 Publication of cloning of HisIE by Ko Fujimori

Approximately 18 months into my PhD after I had cloned, overexpressed and purified the HisIE enzyme a group from Japan led by Ko Fujimori published details of the cloning of a number of genes involved in histidine biosynthesis from *Arabidopsis*. This included details of the hisIE cDNA and genomic sequences. The published data confirmed the DNA and predicted protein sequences for my overexpression clones as well as the presence of a targeting sequence within the first 50 amino acids.

The published data contained details of the intron /exon structure of the gene which is composed of five exons and four introns. The first exon encodes the targeting sequence. Exons two and three then code for the HisI domain and exons four and five code for the HisE domain. From the intron /exon structure the predicted junction for

the targeting sequence is six amino acid residues in front of the cleavage site that I engineered into my constructs.

The Japanese workers expressed a number of different constructs and managed to determine which domains were involved in the individual steps of the reaction (Fujimori and Ohta, 1998[a]). Analysis of *Arabidopsis* genomic DNA by Southern blotting revealed that the *hisIE* gene is present as a single copy gene within the *Arabidopsis* genome (Fujimori and Ohta, 1998[a]).

Chapter 4

Purification of the HisIE enzyme

4.1 Purification of the HisIE enzyme

4.1.1 Development of the purification protocol

The purification of this enzyme proved to be difficult and a number of different methods of purification were attempted before the final purification protocol was established. Even though the starting material was a crude extract containing substantial amounts of over-expressed protein several chromatographic steps were required and for some time it proved difficult to remove all traces of other proteins until a final gel filtration step was introduced.

4.2 Large scale growth

Following the success of the overexpression experiments a large scale growth of the *E. coli* expression strain BL21(DE3)pLysS pSACIE25 was carried out to provide material for protein purification. The cells were grown as described in section 2.8.1. From a 4.5 litre growth the average yield of cells was around 15 grams (wet weight). These cells were stored at -20°C until needed.

4.3 Problems with purification

The lack of available substrate for the HisIE enzyme complicated the development of a purification strategy for the enzyme. Initially the only way of monitoring the enzyme activity during the purification was to perform a coupled assay with the first enzyme of the pathway HisG. Both substrates for HisG are commercially available. The problem with the coupled assay was that although it provided an indication of the presence of HisIE activity it was not suitable for determining the number of units and specific activity at each stage. A particular problem of the

coupled assay was that the products of both the HisG and HisIE enzymes absorb at 290nm.

The success of the assay was also highly dependent on the concentrations of salts which made it difficult to use on fractions from the ion exchange columns. Often the simplest way to monitor the protein purification was to run SDS PAGE gels at each stage of the purification to identify the fractions containing the 30kDa HisIE protein band. The lack of a simple assay also meant that it was difficult to determine the stability of the enzyme during the purification procedure and also during storage.

One method, which was developed at a later stage in this work, for providing substrate for the HisIE enzyme was to use a partially purified biotransformation product as the substrate. This was produced by using what was effectively a large scale enzyme assay for HisG. This was incubated until the A_{290} had reached a steady level and then the reaction mix was passed through a 10K filter to remove the HisG enzyme. This allowed the direct monitoring of the activity of HisIE since a constant amount of substrate could be added to each assay.

4.4 Protein Purification

4.4.1 Ion exchange chromatography

The initial step to purify the HisIE enzyme was to use ion exchange chromatography. The resin chosen for chromatography was DEAE-Sephacel. This is a positively charged ion exchange resin. The *E. coli* BL21(DE3)pLysS pSACIE25 cells were slowly thawed on ice before resuspending in Resuspension buffer (Buffer A plus one protease inhibitor tablet per 50ml of buffer). Buffer A (50mMTris-HCl pH 7.5, 0.4mM DTT, and one protease inhibitor tablet per litre of buffer).

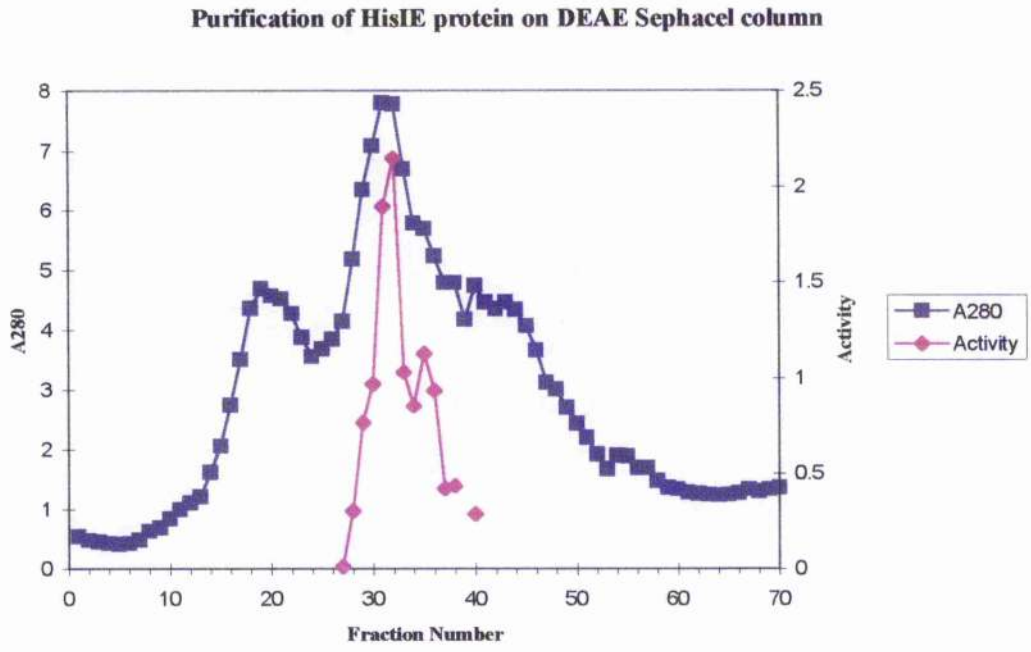


Figure 4.1 Elution profile from the DEAE Sephacel column.

The HisIE protein is eluted approximately half way through the gradient at around 250mM KCl. A 700ml linear gradient of 0-500mM KCl was applied to the DEAE Sephacel column (4.5 x 10cm), at a flow rate of 40ml per hour; 10ml fractions were collected. The A_{280} and the enzyme activity are plotted and appear to correlate well. Although the HisIE enzyme is always eluted at the same position in the gradient, minor discrepancies in the enzyme activity profile are sometimes present which are not reproducible. The HisIE enzyme is assayed as described in 2.10.2.

20ml of Resuspension buffer was added to about 7-10g of cells and the cells were mixed until a smooth paste was formed. This paste was then French pressed twice at 950psi. The extract was diluted to 50ml with the addition of Resuspension buffer. 0.5mg of DNaseI was added and the suspension incubated with stirring at 4°C for 30 minutes before centrifuging at 18,000rpm for 1 hour at 4°C. The supernatant was decanted off and loaded onto a pre-equilibrated DEAE column. The column had been pre-equilibrated with 8 column volumes of Buffer A. The cell extract was loaded onto the column at 25ml per hour. Following loading the column was washed with 3-4 column volumes of Buffer A to remove unbound material etc. A gradient of 0-500mM KCl was applied overnight and 10ml fractions were collected. The gradient consisted of 350ml of Buffer A1 (50mM Tris-HCl pH 7.5, 0.4mM DTT and one protease inhibitor tablet) and 350ml of Buffer A2 (Buffer A1 with 500mM KCl).

The following morning the absorbance at 280nm for each fraction was measured and plotted. This revealed a number of protein peaks but the main enzyme activity was concentrated in the large peak that came off approximately half way through the gradient at around 250mM KCl (Figure 4.1). The fractions were assayed for the presence of enzyme activity. The enzyme activity was plotted on the same graph and correlated well with the main protein peak.

The fractions containing enzyme activity were then pooled and the volume, protein concentration and enzyme activity were measured. The extract was then made to 1M Ammonium sulphate by the slow addition of ground solid. This was added slowing with stirring at 4°C.

4.4.2 Hydrophobic interaction chromatography

The next chromatographic method employed was hydrophobic interaction chromatography on Phenyl Sepharose which separates proteins by virtue of their differing hydrophobicities. A 70ml Phenyl Sepharose column was pre-equilibrated with 5-10 column volumes of Buffer B(50mM Tris-HCl pH 7.5,1M $(NH_4)_2SO_4$, 0.4mM DTT and one protease inhibitor tablet per litre). The pooled fractions from the DEAE column that had been made to 1M $(NH_4)_2SO_4$ were loaded onto the column at 40ml per hour. The column was washed with several column volumes or for at least four hours before applying a gradient of 1M-0M $(NH_4)_2SO_4$ overnight. The gradient consisted of 250ml Buffer B1 (50mM Tris-HCl pH 7.5, 1M Ammonium sulphate, 0.4mM DTT and one protease inhibitor tablet per litre) and 250ml of Buffer B2 (50mM Tris-HCl pH 7.5, 0.4mM DTT and one protease inhibitor tablet per litre). 10ml fractions were collected.

The A_{280} and the activity of the fractions were measured the following morning and the results plotted on a graph. The peak containing the HisIE enzyme activity was found at the beginning of the gradient. It appeared to elute from the column at around 780mM $(NH_4)_2SO_4$.

After two chromatography steps the HisIE protein was approximately 75-80% pure. The remaining purification to get homogeneous protein proved to be difficult and a number of different chromatographic methods were attempted.

4.5 Gel filtration chromatography

The third chromatographic method used to separate the HisIE protein from the remaining contaminants was gel filtration chromatography which separates proteins

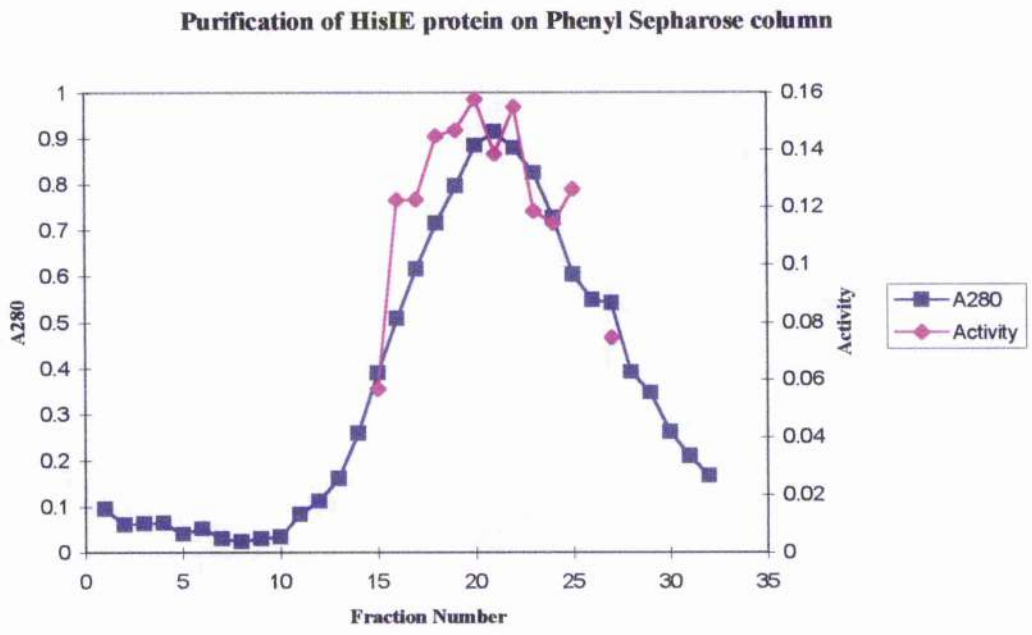


Figure 4.2 Elution profile from the Phenyl Sepharose column

Only the first part of the gradient is represented in this graph. The HisIE protein starts to elute at around 780mM $(NH_4)_2SO_4$ and is off the column by the 500mM point in the gradient. A linear gradient of 1M-0M $(NH_4)_2SO_4$ was applied to the Phenyl Sepharose column (2.7 x 12cm) at a flow rate of 25ml per hour; 10ml fractions were collected. The A_{280} and the enzyme activity are plotted and appear to correlate well. Although the HisIE enzyme is always eluted at the same position in the gradient, minor discrepancies in the enzyme activity profile are sometimes present which are not reproducible. The HisIE enzyme is assayed as described in 2.10.2.

on the basis of size. The resin used for this work was Sepharose S-200 which separates proteins in the molecular weight range 5,000- 250,000. The fractions containing HisIE activity from the phenyl sepharose column were pooled and dialysed overnight against three changes of Buffer C (50mM Tris-HCl pH7.5, 0.4mM DTT and 100mM KCl) in order to remove the $(NH_4)_2SO_4$. The dialysed protein was then concentrated in an Amicon concentrator through a 10K cut off membrane. This was to reduce the volume of the sample to around 2ml for loading onto the S-200 column. The protein was retained by the membrane while the buffer and other solutes were forced through by the pressure of the nitrogen gas. The column was pre-equilibrated for 2 days with Buffer D (50mM Tris-HCl pH 7.5, 0.4mM DTT, and 500mM KCl). The concentrated protein was applied to the column which was run at a speed of 10ml per hour. 5ml fractions were collected during the 48 hours required for chromatography.

The A_{280} of the fractions was monitored and this was plotted on a graph (Figure 4.3). The peaks were then assayed for HisIE activity. The fractions were also analysed by SDS-PAGE in order to determine if the protein was homogeneous.

The fractions containing HisIE activity from the S-200 column were pooled and dialysed for two days into storage buffer (50% Glycerol, 50mM Tris-HCl pH 7.5, 0.4mM DTT) and finally stored at -20°C.

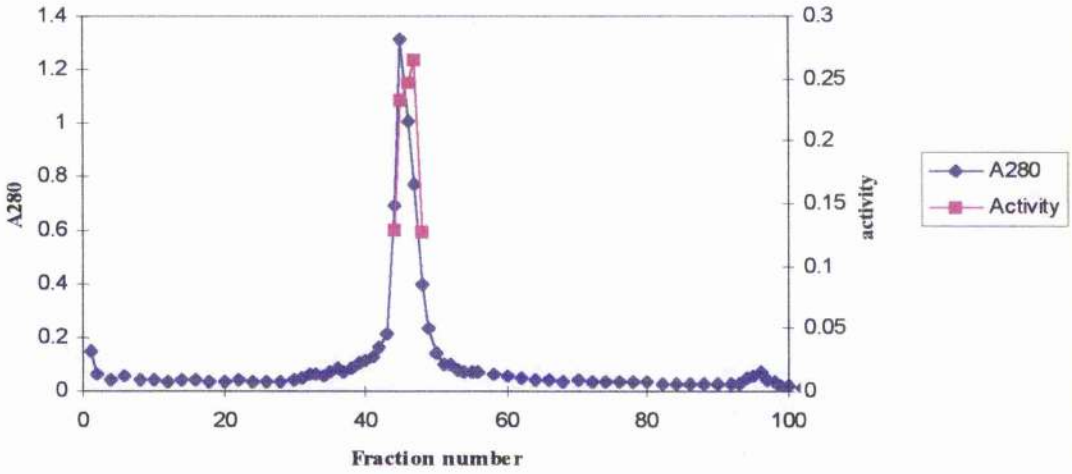
The SDS-PAGE indicated that the protein still contained a number of faint contaminants which were significantly different in subunit molecular weight and should have been separated on the S-200 column from the 31kDa HisIE protein (Figure 4.4). There was in particular a band in line with the 96kDa marker which is likely to be a trimer of the HisIE protein which has somehow become crosslinked

Figure 4.3a and 4.3b Show the elution profiles of the HisIE protein from the S-200 gel filtration column under differing conditions.

Figure 4.3a) shows the elution profile from the column when the protein has been concentrated in the absence of salt. Figure 4.3b) shows the elution profile when the protein had been concentrated in the presence of 100mM KCl. The Sephacel S-200 column (2 x 150cm) was run in Buffer D (50mM Tris-HCl pH 7.5, 0.4mM DTT and 500mM KCl) at a flow rate of 10ml per hour. 5ml fractions were collected.

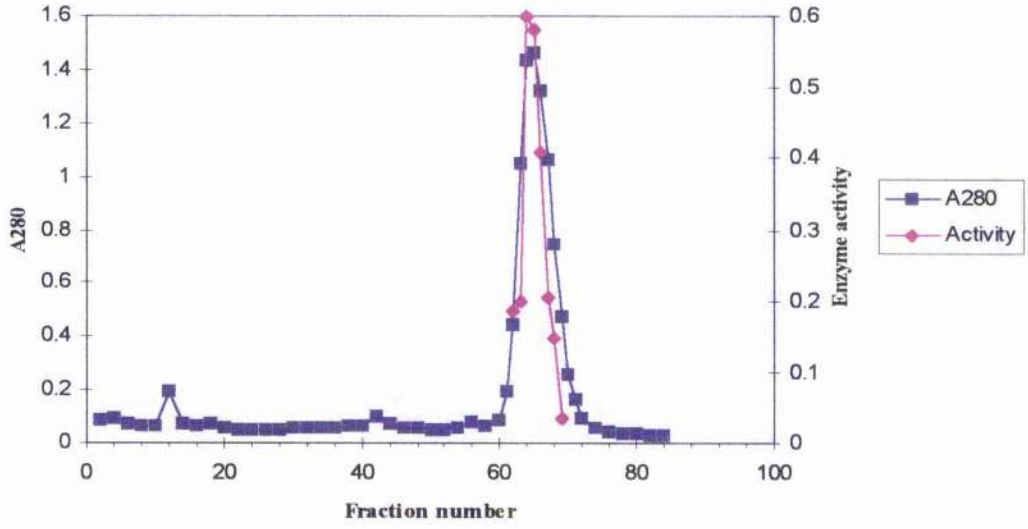
a)

Purification of HisIE using gel filtration following concentration of the protein in the absence of salt



b)

Purification of HisIE using gel filtration following concentration of protein in the presence of 100mM KCl



such that the presence of SDS and mercaptoethanol are unable to separate the subunits. Increasing the concentration of both SDS and mercaptoethanol in the gels and in the sample buffer did show a reduction in the amount of protein present in the band at 96kDa, but it did not eliminate it. The presence of the other proteins in the sample was a little disturbing as by this time and stage of purification the protein should have been homogeneous. Best estimates are that the protein is 95% pure at this stage.

4.6 The key discovery

A simple procedure was the key to unlocking the secret of obtaining pure protein. It became evident that concentrating the sample before applying it to the S-200 column was causing the protein to form an aggregate of some sort. This aggregate was then stable despite the presence of 500mM KCl in the S-200 column. This aggregated protein ran on the column as a higher molecular weight entity. It was clear from the elution position that the protein was running at a molecular weight considerably greater than the 31kDa predicted molecular weight of the monomeric form of the protein. To avoid this aggregation 100mM salt was added to the protein before concentrating it. This addition of salt prevented the formation of the aggregates, so much so that the elution position of the HisIE protein was noticeably shifted. The protein now emerged from the column later indicating that it was running as a lower molecular weight species, predicted to be the dimer. When analysed by SDS-PAGE there was now one major band at 31kDa with two faint bands present below that (Figure 4.5).These two lower molecular weight bands are believed to be degradation products of the full length protein.

4.7 Ammonium sulphate precipitation

The first two columns appeared to be the most successful in removing the bulk of the contaminating proteins from the HisIE. The hydrophobic interaction step which was used as the second column appeared to be a vital step in the purification process. I decided to try use the phenyl sepharose column first to investigate if this gave a more efficient purification step. This step was combined with another means of purification, salt fractionation using ammonium sulphate which is a classical method of protein purification that allows the fractionation of proteins depending on their solubility.

After the cells had been fractionated using the French pressure cell and the extract clarified by centrifugation, varying amounts of solid $(NH_4)_2SO_4$ were added to bring the concentration of the solution to a known percentage of $(NH_4)_2SO_4$. The first fraction was a 0-30% cut. This precipitated much of the DNA and a few other proteins. Cuts of 30-40%, 40-50%, 50-60% and 60-70% were carried out. The various pellets and supernatants were then examined for the presence of HisIE activity. Due to the difficulties of carrying out the coupled assay in such high salt concentrations, the fractions were analysed by SDS-PAGE. The HisIE protein was not confined to one fraction but appeared to be split between the 40-50 and the 50-60% fractions.

These fractions were then diluted and applied to the phenyl sepharose column which had been previously equilibrated with buffer B. A gradient of 1-0M $(NH_4)_2SO_4$ was applied overnight. The pooled fractions were dialysed overnight against buffer A to remove the $(NH_4)_2SO_4$. The following day the protein was applied to the DEAE Sephacel ion exchange column and chromatographed as described previously. When the fractions containing HisIE protein were analysed by SDS-PAGE the protein appeared to be of similar purity to the fractions obtained after chromatography on a DEAE column followed by a phenyl sepharose column (Figure 4.4). Following gel

filtration on the S-200 column the protein did not appear to be more pure than on previous occasions. (The gel filtration had been carried without the 100mM KCl present during concentrating.) All future preparations of enzyme were carried out using the DEAE Sephacel column followed by the Phenyl Sepharose column as this was a shorter protocol.

4.8 Dye binding columns

Many dyes are known to mimic certain molecules found in nature such as certain nucleotides. The HisIE protein binds a nucleotide, PR-ATP, which it then hydrolyses to produce PR-AMP and finally ProFAR. The idea that perhaps one of the dye resins may mimic the substrate of HisIE and therefore would allow an affinity type interaction with the protein was investigated. A test kit purchased from Sigma, which contained a number of different dye resins, was used to identify a suitable dye resin for the purification. In order to test the resins a small amount of resin was removed into an eppendorf tube. The resin was washed several times with 50mM Tris-HCl pH 7.5 buffer. This was performed by shaking the resin with the buffer and then pelleting the resin, removing the supernatant and repeating. A solution of partially purified protein was then applied to the pelleted resin and agitated gently for 20 minutes. The resin was re-pelleted and the supernatant was removed and stored for assaying for HisIE activity. 500mM KCl was added and again the resin was agitated for 20 minutes. The resin was again pelleted and the supernatant was stored for assays.

Several resins appeared to bind the HisIE enzyme with slightly different efficiencies. The percentage of activity that bound to the resin was calculated and the

total recovery was calculated. The Reactive Red 120 resin was chosen as it bound the protein very well and we also had a stock of this material which could be used to prepare a preparative column. When this was tried as a large scale enzyme preparation the protein did bind to the column and was eluted with a 0-500mM KCl gradient. When the fractions were analysed on SDS-PAGE the HisIE protein was not homogeneous (Figure 4.4). The Reactive Red 120 column offered no significant advantage over the anion exchange and hydrophobic columns.

4.9 Determination of the purity of the HisIE enzyme

Samples of the protein at various stages of the purification were subjected to SDS-PAGE analysis to illustrate the progress of the purification and to establish the purity of the final material. Samples from the crude extract, DEAE column, Phenyl Sepharose column and finally the S-200 column were run on a 15% Acrylamide running gel with a 5% stacker (Figure 4.5). The protein was judged to be pure by SDS-PAGE analysis.

4.10 Calibration of the S-200 column in high salt conditions and determination of the subunit molecular weight of the HisIE protein

A number of proteins of known molecular weight were used as molecular weight markers to calibrate the S-200 column. This was to allow the determination of the molecular weight of the HisIE protein as it was eluted from the column. The following molecular weight markers were used; Blue Dextran 2000kDa, Alcohol dehydrogenase 150kDa, Albumin 66kDa, Carbonic anhydrase, 29kDa and Cytochrome C 12.5kDa. Measurement of the A_{280} absorbance of the samples revealed the elution profile from the column. All measurements relating to the elution volume

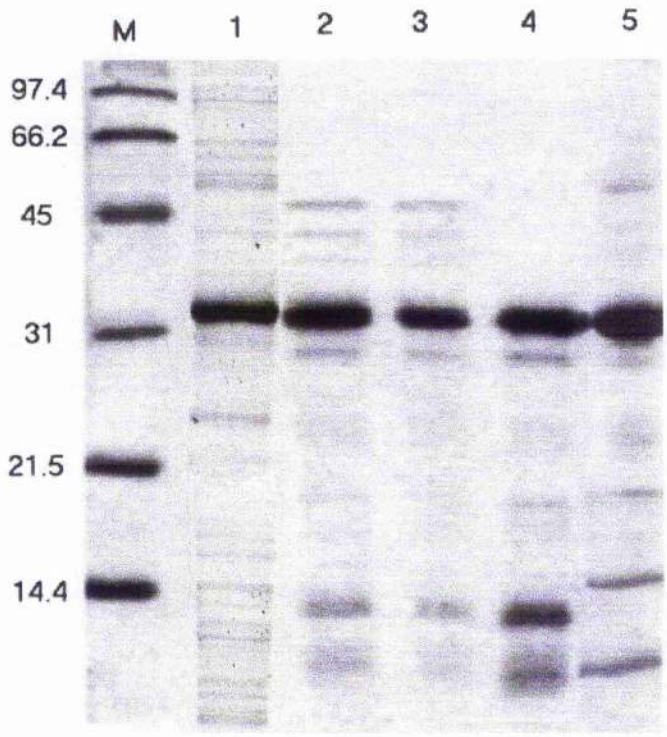


Figure 4.4 Composite SDS PAGE gel showing samples of HisIE protein after various failed attempts to obtain pure protein.

Markers

Lane 1 Following DEAE column

Lane 2 DEAE column followed by Phenyl Sepharose column

Lane 3 Ammonium Sulphate precipitation and Phenyl Sepharose column followed by DEAE column

Lane 4 Following chromatography on dye binding column

Lane 5 Protein sample from S-200 column following concentration in the absence of salt.

A 15% Acrylamide running gel was used in all cases along with a 5% stacking gel. Following electrophoresis at room temperature the gel was stained with Coomassie Blue (section 2.7.3). Lanes 2 and 3 show that there is little difference between the protein following either DEAE/ Phenyl Sepharose or Ammonium sulphate/ Phenyl Sepharose/ DEAE treatment. Lanes 4 and 5 show that the HisIE protein is not homogeneous following either the dye binding column or gel filtration.

Sample	Volume (ml)	Protein (mg/ml)	Protein (mg)	Activity (units/ml)	Total Units	% recovery	Specific activity	Purification
Crude extract	90	23	2070	35	3150	100	1.52	1
DEAE pool	76	8.5	646	24	1820	57.9	2.82	1.86
Phenyl sepharose	80	2.5	200	7.4	591	18.6	2.96	1.94
S-200	14	8	112	26.4	370	11.7	3.30	2.17

Table 4.1 Purification Table for purification of the HisIE enzyme

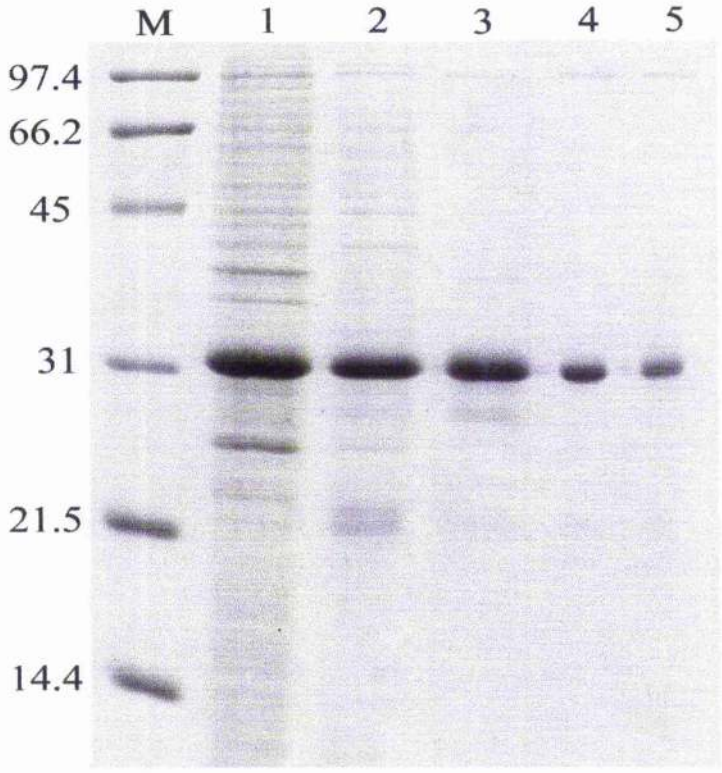


Figure 4.5 Analysis of protein purification by SDS PAGE

Markers	(molecular weights are given alongside the gel photograph in kDa)
Lane 1	Crude extract (15μg total protein)
Lane 2	DEAE pool (10μg total protein)
Lane 3	Phenyl Sepharose pool (10μg total protein)
Lane 4	S-200 pool (5μg total protein)
Lane 5	S-200 pool (5μg total protein)

A 15% Acrylamide running gel was used along with a 5% stacking gel. Following electrophoresis at room temperature the gel was stained with Coomassie Blue (section 2.7.3). A faint band running with the 97kDa marker is present in all samples and is believed to be a trimer of the HisIE protein.

were taken from the summit of the peaks in the elution profile. The results were then plotted as elution volume versus log molecular weight. A standard curve was drawn and used to determine the molecular weight of the HisIE protein (Figure 4.6). The HisIE protein which had been concentrated in the absence of salt, forming a large aggregate, has a high molecular weight which is excluded or almost excluded from the column i.e. greater than 250kDa. In comparison the protein concentrated in the presence of 100mM KCl has a calculated molecular weight of 63kDa which would indicate that the HisIE protein is running as a dimer, despite the presence of 500mM KCl in the gel filtration column.

4.11 Determination of the molecular weight of the HisIE protein under native gel filtration conditions

The same S-200 gel filtration column was then purged with low salt buffer containing 50mM Tris-HCl pH 7.5 and 0.4mM DTT. The column was calibrated using the following molecular weight markers; Blue Dextran 2000kDa, β Amylase 200kDa , Albumin 66kDa, Carbonic anhydrase, 29kDa and Cytochrome C 12.5kDa. A standard curve was drawn and used to determine the native molecular weight of the protein (Figure 4.7). 5mg of purified protein at a concentration of 10mg/ml was applied to the column. Under low salt conditions the elution position of the HisIE protein indicates the molecular weight to be around 61kDa. This suggests that the protein is a dimer under these conditions and at this protein concentration. It is likely that under physiological conditions within the cell the HisIE protein would be present as a dimer.

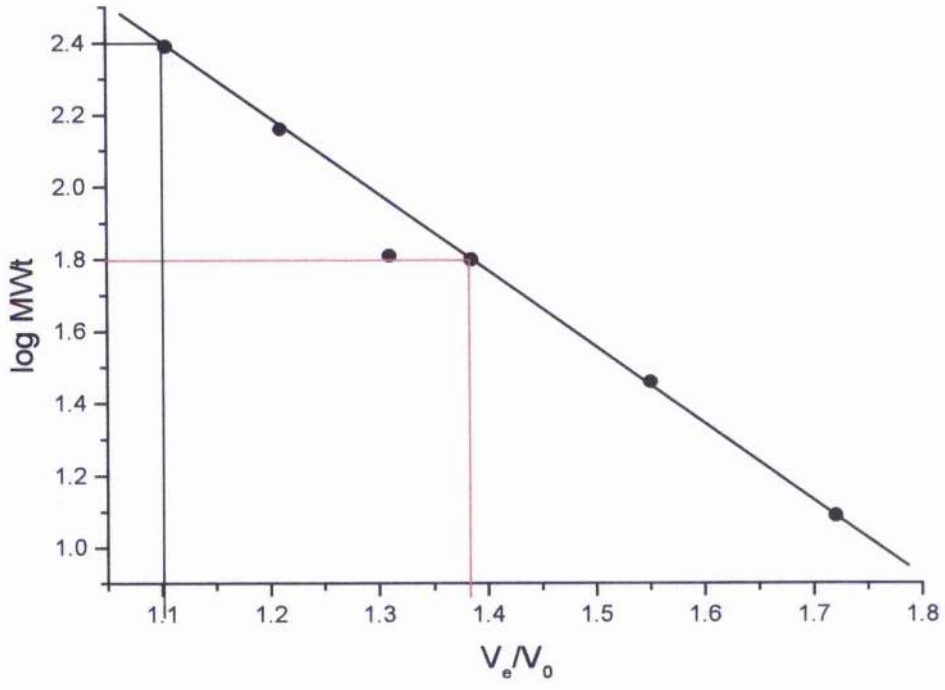


Figure 4.6 Calibration curve for S-200 column in the presence of 500mM KCl

Standard curve plotting V_e/V_0 against log molecular weight. The following molecular weight markers were used; Blue Dextran 2000kDa, Alcohol dehydrogenase 150kDa, Albumin 66kDa, Carbonic anhydrase, 29kDa and Cytochrome C 12.5kDa. The S-200 column (2 x 150cm) was run in high salt buffer (50mM Tris-HCl pH 7.5, 0.4mM DTT and 500mM KCl) at a constant flow rate of 10ml per hour. HisIE protein concentrated in the absence of salt has a calculated molecular weight greater than 250kDa (Black line). HisIE protein concentrated in the presence of 100mM KCl has a calculated molecular weight of 63kDa which would indicate that the HisIE protein is running as a dimer (Red line).

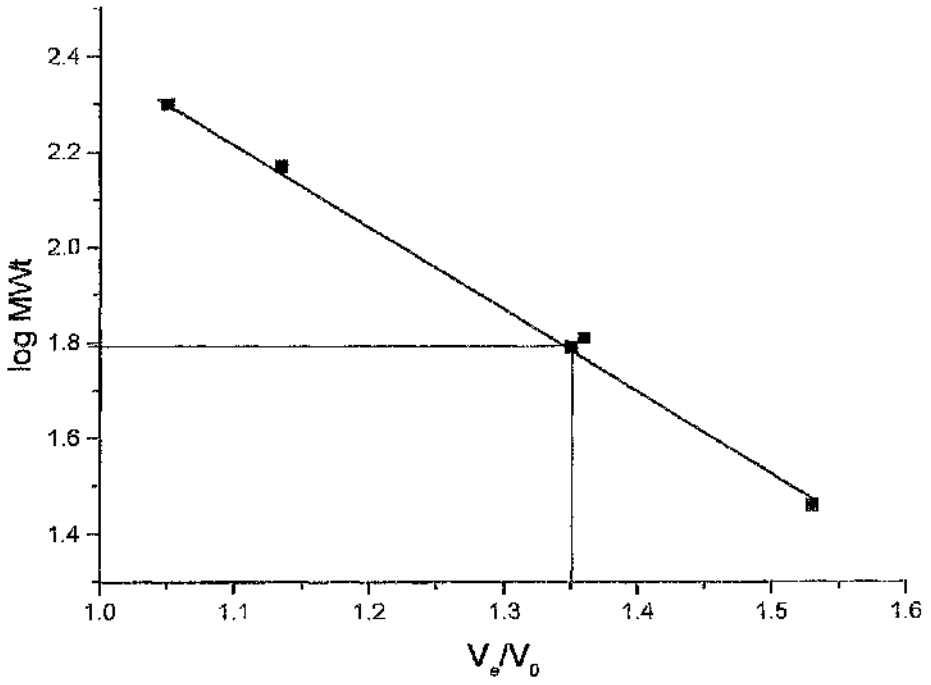


Figure 4.7 Calibration curve for S-200 column under low salt conditions

Standard curve plotting V_e/V_0 against log molecular weight. The following molecular weight markers were used; Blue Dextran 2000kDa, β Amylase 200kDa , Albumin 66kDa, Carbonic anhydrase, 29kDa and Cytochrome C 12.5kDa. The S-200 column (2 x 150cm) was run in low salt buffer (50mM Tris-HCl pH 7.5 and 0.4mM DTT) at a constant flow rate of 10ml per hour. Under low salt conditions the Cytochrome C protein bound to the resin and only the presence of high salt allowed it to move through the column. Purified HisIE protein has a calculated molecular weight of 63kDa under native conditions which would indicate that the HisIE protein is a dimer.

4.12 Mass spectrometry analysis of the HisIE protein

Electrospray mass spectrometry of the purified HisIE protein was carried out by Dr K. Lilley, University of Leicester. Preliminary results indicate that the protein has a mass in the correct range for the predicted molecular weight of 26,859Da. The exact value obtained was 26,874; this discrepancy is almost certainly due to a failure to calibrate the instrument properly. A further mass analysis will be carried out.

Chapter 5

Purification of Phosphoribosyl ATP

5.1 General introduction

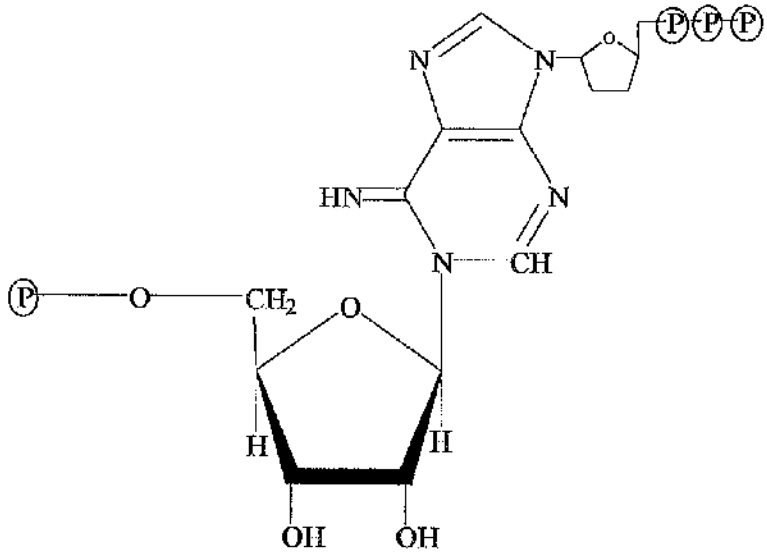
Phosphoribosyl ATP (PR-ATP, Figure 5.1) is the product of the first enzyme of the histidine biosynthetic pathway and is formed by the condensation of PRPP and ATP by the formation of an N-glycosyl linkage between the C-1 of the phosphoribosyl group and the N-1 of the purine of ATP. This reaction is catalysed by the enzyme ATP-phosphoribosyl transferase (E.C. 2.4.2.17) the product of the *hisG* gene. The reaction produces inorganic pyrophosphate from the PRPP moiety and PR-ATP. The formation of the PR-ATP can be monitored by measuring the increase in absorbance at 290nm (Martin, 1963).

ATP-phosphoribosyl transferase has been well studied from both bacterial and fungal species (Martin *et al.*, 1971, Winkler, 1988) and a strain which overproduces the *E. coli* enzyme has been constructed by Mr A. Elwell (unpublished results). Much less is known about the plant enzyme. In this work we have used the purified *E. coli* ATP-phosphoribosyl transferase to generate PR-ATP.

To characterise the bifunctional HisIE enzyme a supply of the substrate PR-ATP was essential. The most convenient means of generating PR-ATP is by biotransformation from PRPP and ATP (which are both commercially available) using ATP-phosphoribosyl transferase.

5.2 Enzymatic synthesis of PR-ATP

Since PRPP is relatively expensive and unstable (it must be stored at -80°C) the possibility of generating it enzymatically was also considered. However, because of the added complexity of having a second biotransformation step, the direct route was used to generate PR-ATP.



PR-ATP

Figure 5.1 Structure of PR-ATP

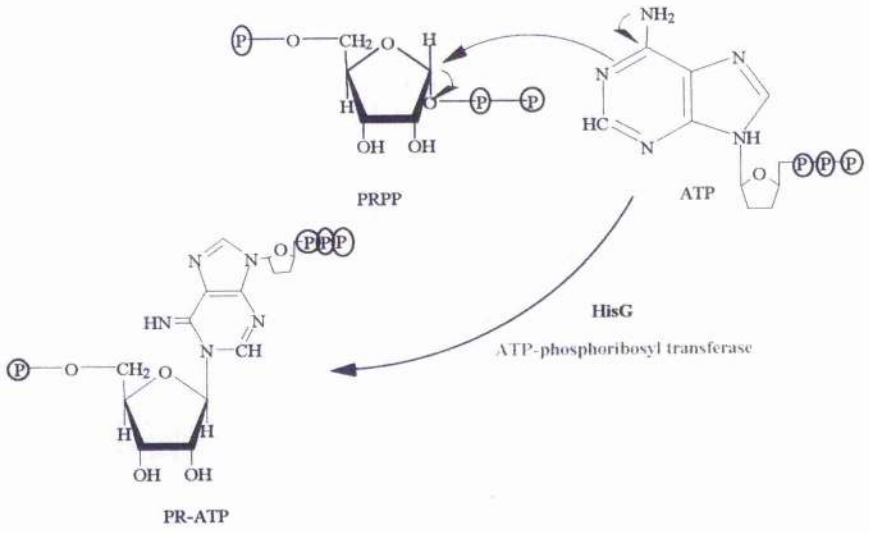


Figure 5.2 a) Generation of PR-ATP direct from PRPP and ATP

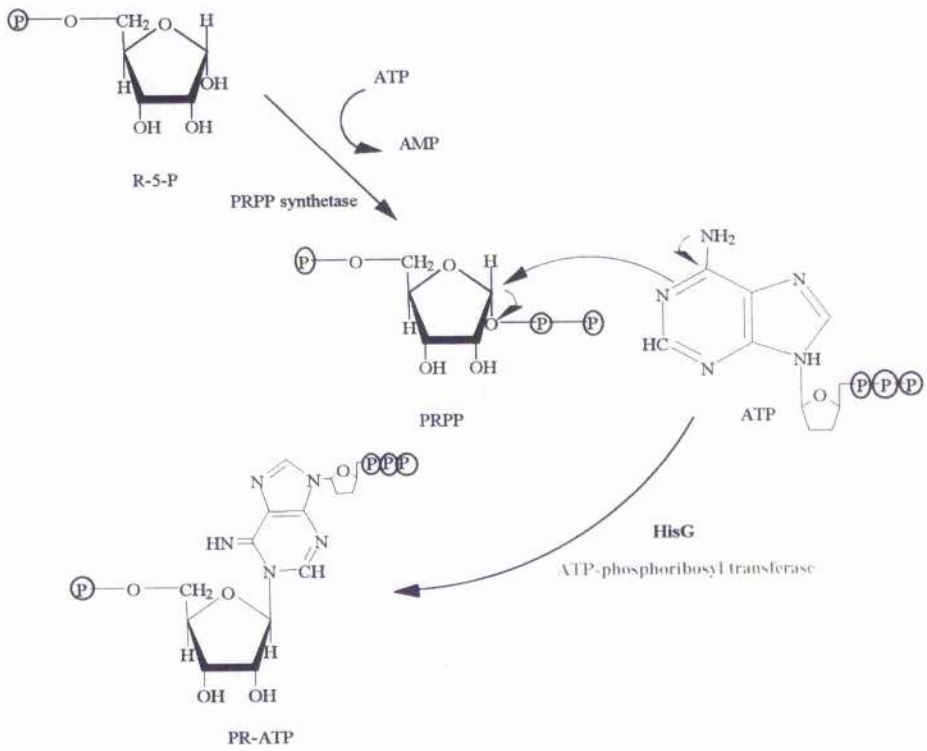


Figure 5.2 b) Generation of PR-ATP from Ribose-5- phosphate

5.3 Initial attempts at generating PR-ATP

The original biotransformation as described in Martin *et al.*, (1971) consisted of 1.5mM PRPP, 5mM ATP, 10mM $MgCl_2$, 100mM Tris pH 8.5, inorganic pyrophosphatase and a crude extract from *Salmonella typhimurium* containing ATP-phosphoribosyl transferase. Our initial experiments to produce the substrate used this method to generate PR-ATP except that purified *E. coli* enzyme was used instead of the bacterial extract.

The 330ml reaction mixture was incubated at 37°C and the progress of the reaction monitored by removing a 1ml aliquot and observing the UV spectrum continuously for 4 hours (Figure 5.3). The maximum A_{290} was observed after 40 minutes and in subsequent preparations the pH of the bulk reaction mixture was adjusted at this stage to 6 with 1M HCl and the resulting solution frozen and stored at -20°C.

Continuous monitoring of the 1ml aliquot of the reaction mixture for 4 hours showed that the maximum A_{290} absorbance occurred after 40 minutes and began to decline after 60 minutes, gradually decreasing until after about 120 minutes the A_{290} had fallen almost to zero (Figure 5.3). This suggests that the reaction product PR-ATP was unstable under these conditions, perhaps due to the high pH (8.5) of the buffer conditions or to the temperature (37°C).

5.3.1 Chromatography of the PR-ATP

The crude PR-ATP solution was split into two equal batches and each was chromatographed separately. A Q-Sepharose column was pre-equilibrated with 10mM

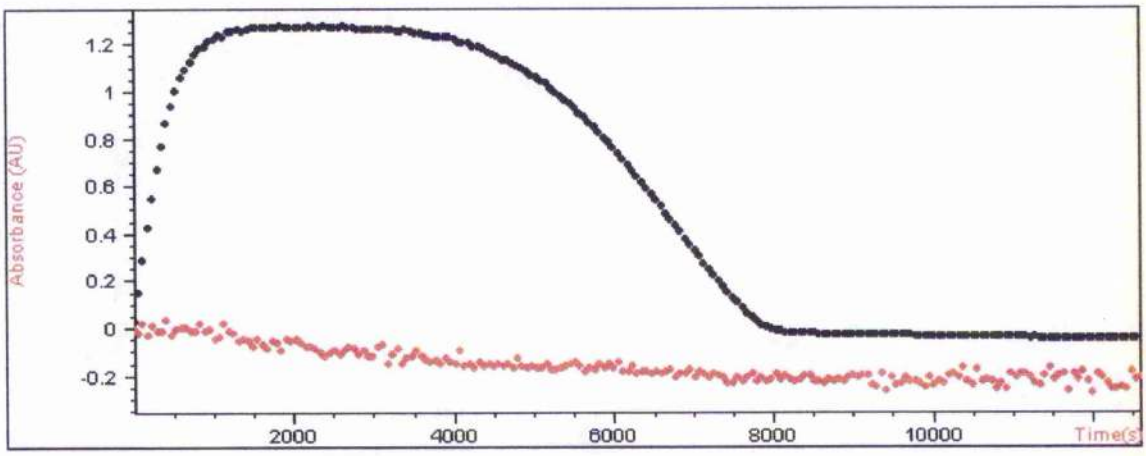


Figure 5.3 Biotransformation reaction showing the profile observed during Biotransformation to produce PR-ATP.

Spectrophotometer traces showing the changes in absorbance with time during the enzymatic synthesis of PR-ATP. The absorbance changes at both 290nm (black line) and 260nm (red line) were monitored for a 1ml aliquot of the 330ml biotransformation mixture used to generate PR-ATP. Reaction was produced as described in section 5.3 using the method of Martin *et al.*, (1971). The pH of the bulk reaction was adjusted to 6 after 40 minutes, although the reaction was monitored for several more hours using the 1ml aliquot.

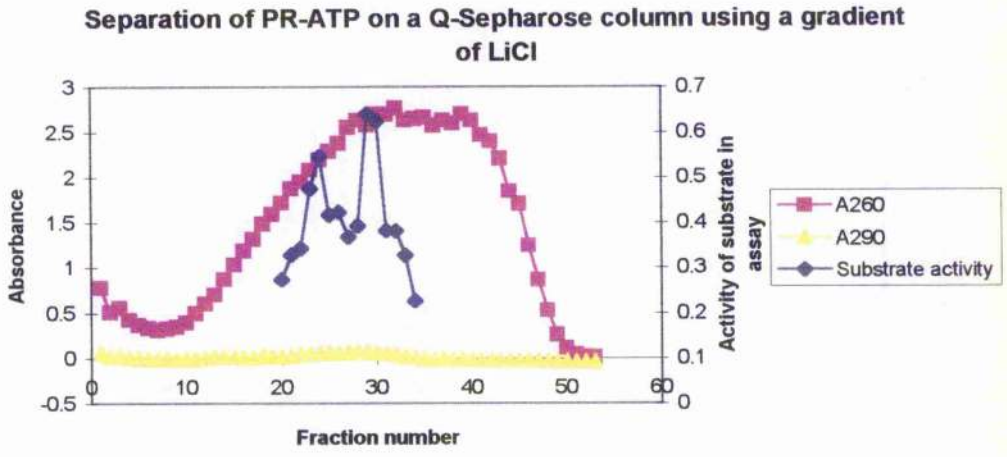


Figure 5.4 Separation of PR-ATP by Q-Sepharose ion exchange chromatography with LiCl as eluent.

Ion exchange chromatography of PR-ATP on a Q-Sepharose (Pharmacia Biotech) column (4.5 x 12cm) at 4°C. A 1 litre linear gradient of 50-400mM LiCl / 10mM Imidazole pH 6.5 was run at a constant flow rate of 50ml per hour. 10ml fractions were collected. The A_{290} and A_{260} absorbances of the fractions were measured, and the presence of PR-ATP was established by assaying fractions in the presence of HisIE enzyme. The split peak of substrate activity is not clearly understood but the presence of large quantities of ATP in the second peak is believed to contribute to the difference in activity. See discussion in section 5.3.2.

imidazole buffer pH 6.5 containing 50mM LiCl. The PR-ATP solution was diluted with this buffer until the conductivity of the solution was the same as the eluate from the column. This meant an approximately 15-fold dilution of the original sample to a total volume of almost 2 litres. The original Martin *et al.*,(1971) method used a DEAE cellulose resin for the chromatography and a gradient of 100-140mM LiCl to elute the material from the column. The Q-Sepharose resin is a stronger anion exchange resin and thus a gradient of 50 to 400mM LiCl was used to ensure that PR-ATP was eluted from the column. The absorbance of the fractions was measured at both 260 and 290nm. (Figure 5.4).

5.3.2 Preliminary characterisation of PR-ATP as a substrate for the HisIE enzyme

Assaying the fractions eluted from the Q-Sepharose with the HisIE enzyme, identified those fractions containing PR-ATP. The absorbance profile from the column shows a broad peak of material absorbing in the UV at both 260 and 290nm. The fractions were assayed in the presence of identical amounts of HisIE enzyme and the initial rate of A_{290} increase, observed for each fraction was plotted as a means of determining the presence and relative quantities of active substrate. The fractions in the first part of the peak show the normal profile observed during the HisIE enzyme assay (Figure 6.1), with both an increase in the absorbance at 290nm as well as a decrease in the absorbance at 260nm. The second part of the peak was also assayed and found to contain PR-ATP. There was however a key difference in the assay profile. The previously observed steady decrease in A_{260} did not occur; instead the A_{260} readings fluctuated. The increase in A_{290} was still observed. This different behaviour in the assays indicates that the later fractions differ from those at the beginning. It was

suspected that these later fractions were contaminated with residual ATP that had not been successfully resolved on the column. In order to test this theory the first fractions containing PR-ATP were spiked with 100μM ATP and re-assayed. The A_{260} decrease that had been previously observed was lost and replaced by a random fluctuation in the A_{260} as seen in the assay of the later fractions. This was consistent with the proposal that the later fractions were contaminated with ATP which was later confirmed by HPLC analysis of the freeze dried fractions. The fractions containing PR-ATP were divided into two pools. The first fractions, showing both absorbance changes, were designated as pool A and the later fractions, showing only the A_{290} change, as pool B.

The lyophilized pools of PR-ATP were washed three times with cold absolute ethanol (Analar Grade) to remove the LiCl and imidazole since both lithium chloride and imidazole are soluble in ethanol whereas the lithium salts of the histidine pathway intermediates are much less soluble in ethanol (Martin *et al.*, 1971). 20ml of ethanol was used for each wash and the solution mixed thoroughly by vortexing. The solution was then centrifuged at 17,000rpm to pellet the insoluble PR-ATP. After three washes, the resulting precipitates were dried exhaustively on the freeze drier to remove all traces of solvent ethanol. The freeze dried samples were then analysed by HPLC and NMR.

5.3.3 HPLC analysis

5.3.3.1 Supelcosil LC-18-S

Supelcosil LC-18-S (4.6mmIDX150mm) is a modified C18 reverse phase column from Supelco which has a special surface treatment along with an octadecylsilane bonded phase, and has been developed specifically for the efficient separation of nucleotides. This column was run with a gradient of Buffer A (0.1M-

Figure 5.5 HPLC analysis of PR-ATP pools showing the different amounts of ATP contamination

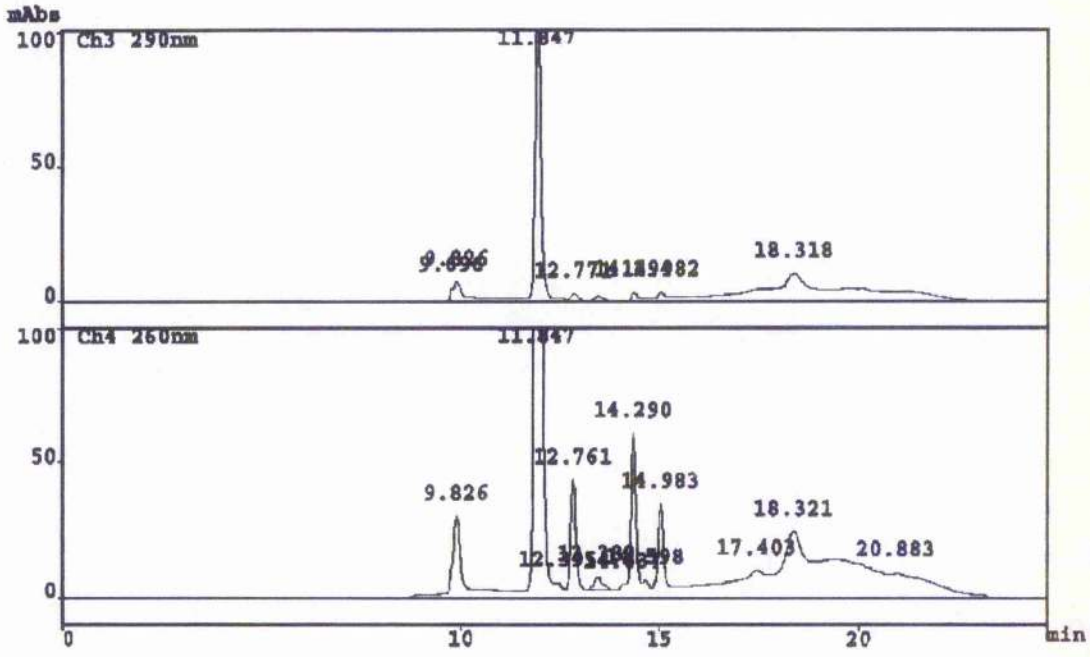
Profile observed during reverse phase chromatography of PR-ATP samples from pools A and B using a Supelcosil LC-18-S column. PR-ATP has a retention time of 11.8 minutes and ATP has a retention time of 14.2 minutes. The ATP peak at 14.2 minutes is larger in pool B confirming that this sample contained increased amounts of ATP.

The gradient was run as follows,

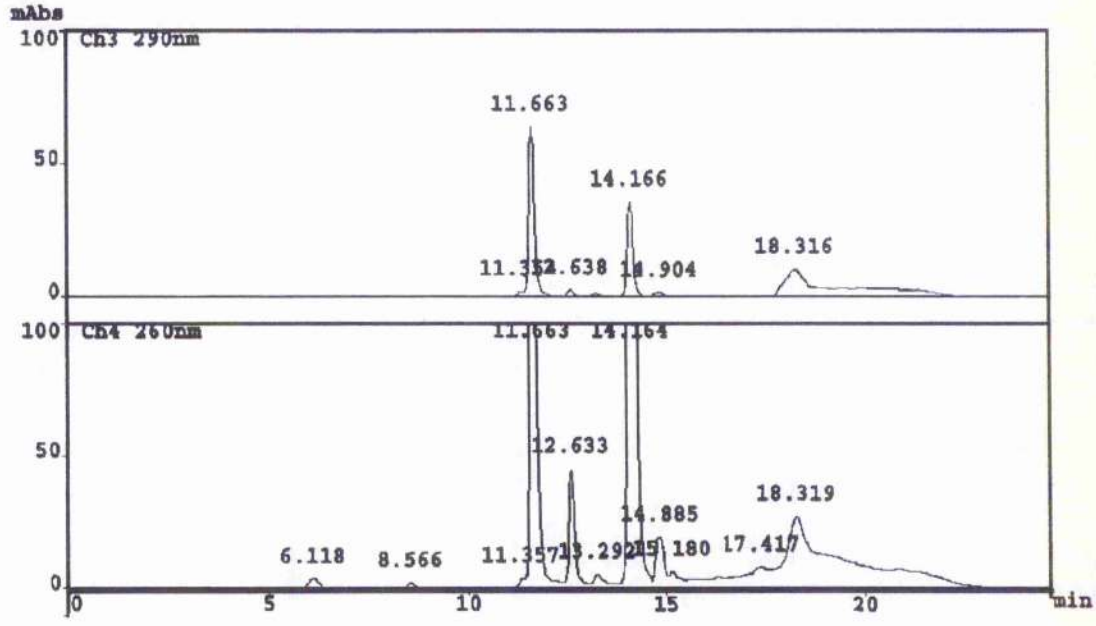
Time (min)	%B
0	0
2.5	0
5	30
10	60
13	100
18	100
20	0
25	0

The flow rate was 1.5ml per minute and the temperature was 25°C.

Pool A


mAbs
100
Ch3 290nm
11.847
50
0
18.318
100
Ch4 260nm
11.847
14.290
12.761
14.983
9.826
18.321
17.403
20.883
50
0
0
10
15
20
min

Pool B


mAbs
100
Ch3 290nm
11.663
50
14.166
18.316
0
100
Ch4 260nm
11.663
14.164
12.633
50
18.319
14.885
6.118
8.566
17.417
0
0
5
10
15
20
min

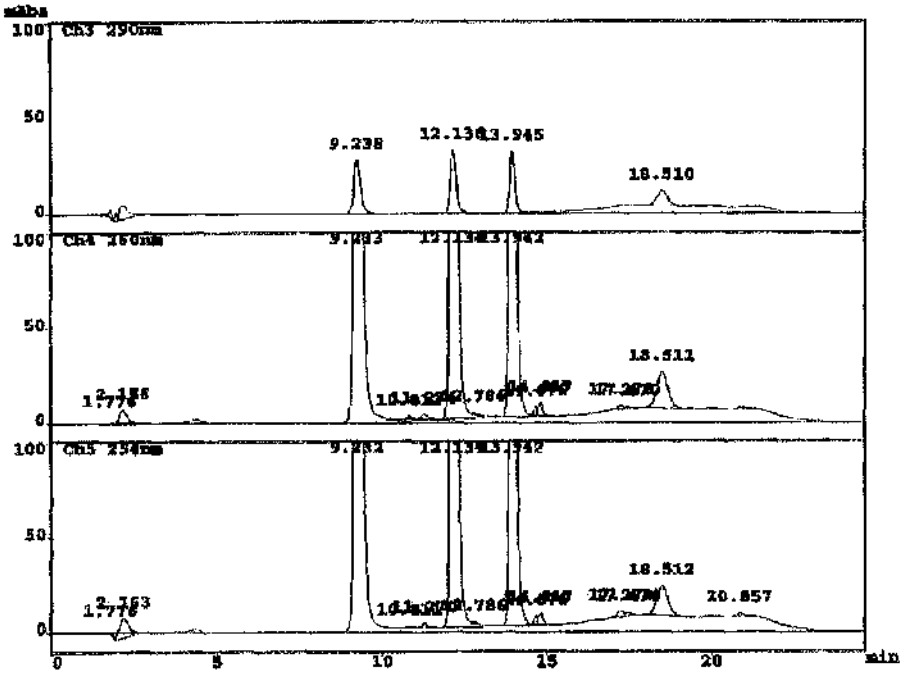


Figure 5.6 HPLC analysis of ATP, ADP and AMP on the Supelcosil LC-18-S column.

AMP, 9.2 minutes, ADP, 12.2 minutes and ATP, 13.9 minutes

The gradient was run as follows,

Time (min)	%B
0	0
2.5	0
5	30
10	60
13	100
18	100
20	0
25	0

The flow rate was 1.5ml per minute and the temperature maintained at 25°C. (Buffer A 0.1M potassium phosphate buffer pH 6.0, 4mM-tetrabutylammonium hydrogen sulphate) and Buffer B (A: methanol (70:30) pH 7.2).

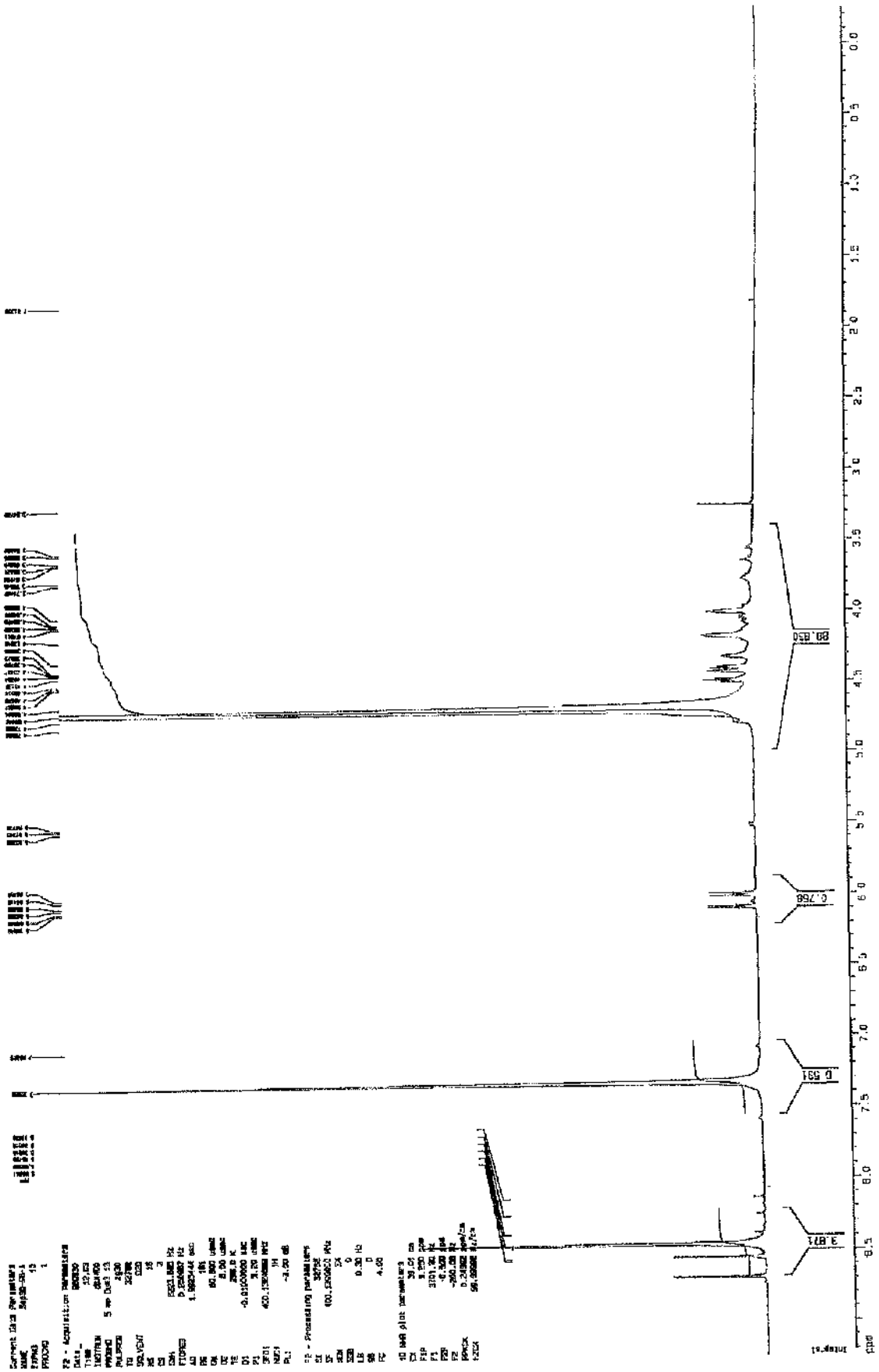
potassium phosphate buffer pH 6.0, 4mM-tetrabutylammonium hydrogen sulphate) and Buffer B (A: methanol (70:30) pH 7.2).

5.3.3.2 Results of the chromatography

The samples for analysis were made up at a concentration of 5mg/ml in distilled water. Standard solutions of ATP, ADP, AMP were made up at a concentration of 10mM to calibrate the column and to aid identification of the peaks. Standard solutions were run both individually and as mixtures to determine the retention times. The column resolved all three adenine nucleotides and they could be seen as distinct peaks that eluted from the column in the order AMP, ADP, ATP with retention times of 9.2 minutes, 12.2 minutes and 13.9 minutes respectively (Figure 5.6). The samples of PR-ATP from the Q-Sepharose column showed the presence of significant quantities of ATP as well as ADP. The peak corresponding to the PR-ATP came off at 11.8 minutes, ahead of both ADP and ATP. The LC-18-S column thus proved to be very useful as an analytical column since it allowed the identification of several of the expected contaminants in the partially purified PR-ATP pools. Analysis of pools A and B confirmed that pool B was particularly highly contaminated with ATP (Figure 5.5). As ATP appears to interfere with the assay of the HisIE enzyme it is essential to remove this contaminant from the PR-ATP. This column however is only useful as an analytical column since it requires the presence of an ion pairing agent tetrabutylammonium hydrogen sulphate. A preparative chromatographic procedure using volatile buffers had now to be developed to produce samples of pure PR-ATP.

Figure 5.7 400MHz Proton NMR spectrum of partially purified PR-ATP sample still containing imidazole

Imidazole peaks are present at 7.3 and 8.5 ppm.


88.850
6.766
6.591
3.811
ppm

5.3.4 NMR Analysis

The 400mHz NMR spectrum of the sample isolated from pool A is shown in Figure 5.7. The spectrum revealed the presence of large quantities of imidazole that had not been removed by the ethanol washes. The imidazole appeared as two large peaks in the region 7.3 and 8.5 ppm on the NMR trace. While the spectrum contained peaks consistent with the presence of PR-ATP, detailed assignment was not possible at this stage because of the high levels imidazole and the presence of some contaminating ATP.

To establish that the crude PR-ATP sample was stable for 24 hours at room temperature, approximately 11-14°C, the NMR was repeated after 24 hours. The two spectra were identical indicating that there is no significant degradation occurring over this time scale.

5.3.5 Development of a simpler chromatographic method for the analysis of PR-ATP using a Partisil ODS3 column (Phenomenex)

This is a C18 reverse phase column which has a 10.5% carbon load, is end capped and has a particle size of 10µm. A number of different solvents and gradients were investigated. Initial attempts at separation used a potassium phosphate buffer containing 4mM tetrabutylammonium hydrogen sulphate to aid the separation. This resulted in a clean and rapid separation of the PR-ATP from the contaminants such as ATP (within 5 minutes) but still required an ion pairing agent and was therefore not promising as a potential preparative method.

A simple isocratic separation was achieved on this column using 25mM potassium phosphate pH 6. The PR-ATP was resolved from the residual ATP within a

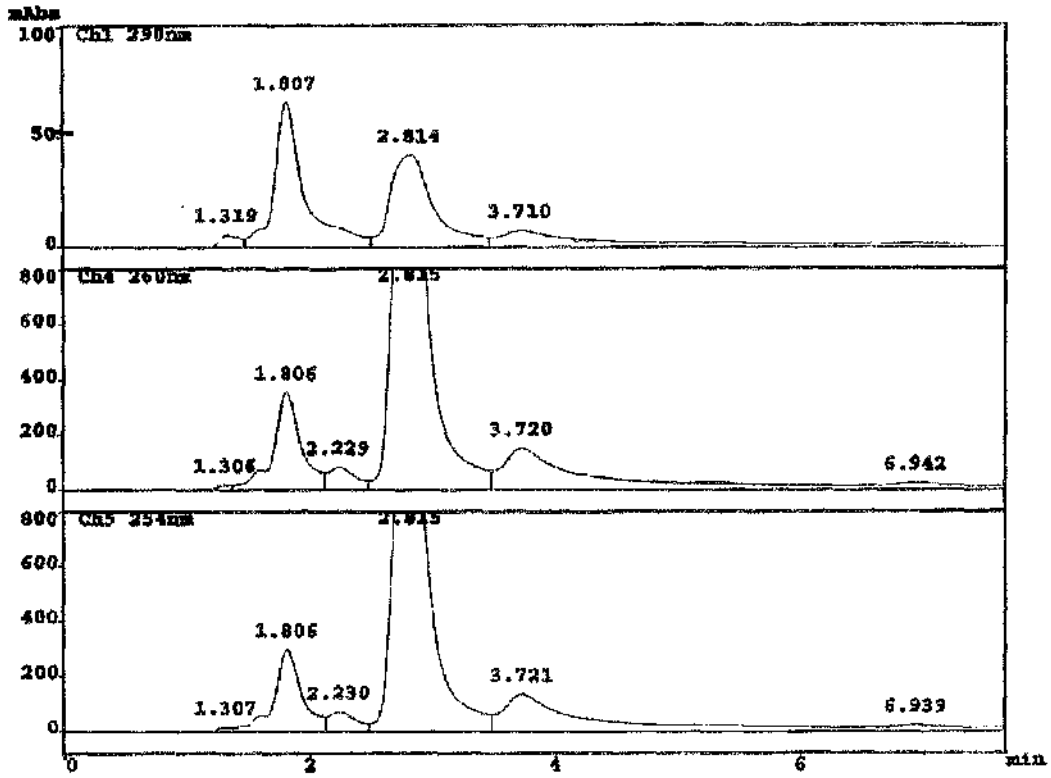


Figure 5.8 HPLC analysis of PR-ATP pool B using the Partisil ODS3 column and 25mM potassium phosphate buffer

PR-ATP has a retention time of 1.8 minutes and ATP has a retention time of 2.8 minutes.

Time (min)	%B
0	0
8	0
9	100
14	100
15	0
20	0
20	stop

Buffer A is 25mM Potassium phosphate pH 6 and the flow rate was 2ml per minute. Buffer B is 100% methanol and is used to remove residual material bound to the column

few minutes (Figure 5.8). Using the analytical column (4.6mm x 250mm) the PR-ATP eluted at 1.8 minutes and ATP eluted at 2.8 minutes. The column does retain some UV absorbing material if used for a significant time without a washing step with 100% methanol. A washing step was therefore included whenever this method was used (Figure 5.8). This method was used for all subsequent analysis of PR-ATP samples to determine if they were pure following initial chromatography and to identify the extent of any contamination with ATP. Due to the presence of 25mM potassium phosphate in the buffer this method is not satisfactory for preparative use.

5.4 Simplifying the biotransformation to reduce ATP contamination in the purified product

The major problem with the production of pure PR-ATP by the method developed here from the Martin *et al.* procedure is the significant quantities of contaminating ATP that have to be removed. Reducing the amount of excess ATP in the reaction would simplify the purification. Therefore the effect of using different amounts of ATP in the biotransformation was investigated. It was found that a 1:1 ratio of PRPP: ATP was sufficient to allow the reaction to proceed to a satisfactory extent. The presence of excess ATP does not make a great deal of difference to the amount of PR-ATP formed in the biotransformation reaction. The reaction with a 2:1 ratio of ATP:PRPP does progress more rapidly but levels off at the same final absorbance value as the 1:1 reaction implying that the amount of product formed is similar in both cases.

5.5 A simplified procedure for generating Phosphoribosyl ATP

The biotransformation was carried out as described in section 2.9.1. and the reaction mixture contained equal molar amounts of the two starting materials at a final concentration of 0.5mM. The course of the reaction was monitored at 290nm and was observed to be complete within 1 hour reaching a final A_{290} of 1.2 (Figure 5.9). Using the known extinction coefficient of PR-ATP (3600; Ames *et al.*, 1961) the reaction should have gone to a final A_{290} of 1.8; the observed A_{290} of 1.2 represents a conversion of 66% of the material. The molecular weight of PR-ATP is 719 and on a scale of 0.5mM this represents a yield of 216mg of PR-ATP.

5.5.1 Preparative chromatography of PRATP.

A Q-sepharose column (4.5 x 12cm) was equilibrated with 5 column volumes of Buffer A (50mM Triethylamine formate pH 6). This buffer was chosen as it was a volatile buffer which was able to buffer in the pH range required for the stability and purification of PR-ATP (around 4-7) (Perrin and Dempsey, 1974). The frozen biotransformation reaction was thawed on ice and was diluted to the same conductivity as the equilibrating buffer. The pH was measured to ensure that the pH of the sample was the same as that of the buffer coming off the column. It was found that no adjustment was necessary. The biotransformation reaction was diluted almost 20 fold with Buffer A and then loaded onto the Q-sepharose column. A 1 litre linear gradient of 50mM to 1M Triethylamine formate was applied and 10ml fractions collected.

Samples were analysed spectrophotometrically for the presence of material absorbing at 290nm (Figure 5.10). Only one peak has significant 290 absorbance and is eluted from the column between 750-800mM triethylamine. The fractions

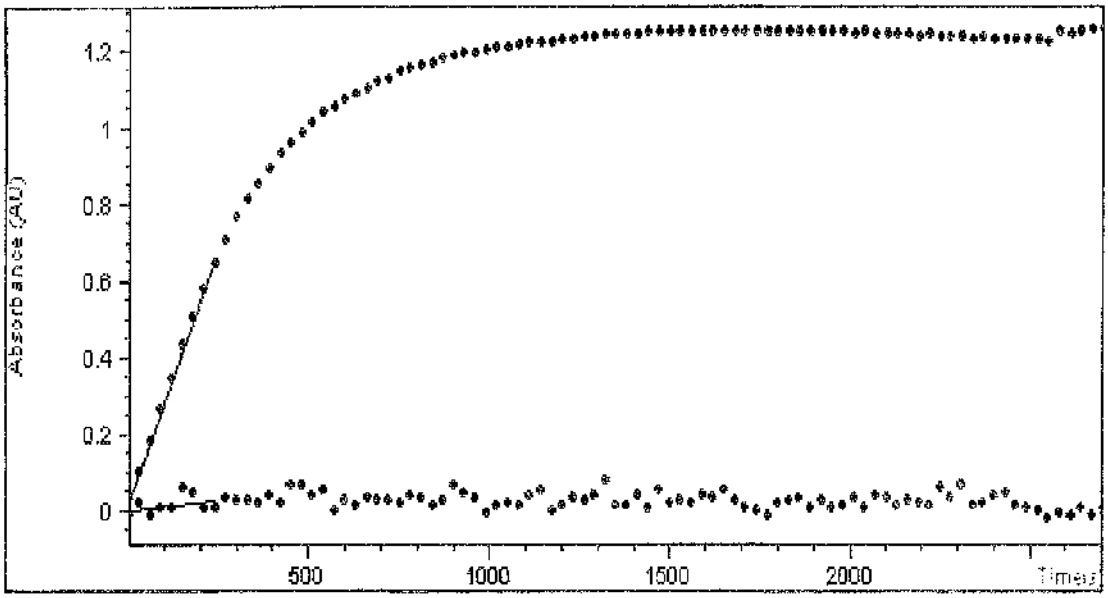


Figure 5.9 Biotransformation reaction

This figure shows the progress of the biotransformation reaction used to produce PR-ATP. The reaction is observed to have gone to completion within 50 minutes as no further increase in 290 absorbance is measured. The final A_{290} is 1.25.

associated with this peak were demonstrated to contain PR-ATP by assaying for HisIE activity and were then pooled and lyophilized overnight. Two peaks containing material which absorbed at 260nm were observed (Figure 5.10) one before and one following the PR-ATP peak. One is believed to be the unreacted ATP.

5.5.2 Lyophilization of PR-ATP samples in Triethylamine formate.

Although Triethylamine formate buffer is a volatile buffer system it was very difficult to remove during freeze drying. A brown, syrupy residue remained in the flask despite repeated freeze drying. Triethylammonium formate is a light brown syrup material which is very soluble in water (Merck Index, 9th Edition). The material in the flask was assumed to be the Triethylamine salt of PR-ATP. The pH of the material after redissolving in water was 3.8; this was adjusted to 6.5 by the addition of a small amount of 5M NaOH. An ammoniacal odour was observed from the flask indicating the displacement of the triethylamine. This material was then freeze dried overnight. A dry, white, fluffy solid was observed in the flask. Unfortunately much of this material was sodium formate. NMR analysis identified the relative amounts of both PR-ATP and sodium formate. This salty material was not ideal for kinetic analysis and therefore methods to remove the salt from the sample were investigated.

The problem observed during repeated freeze drying attempts is that the substrate breaks down to produce phosphoribose and ATP. This breakdown can be monitored by comparing samples at different stages of the drying process. The more the freeze drying process is repeated the more extensive the breakdown appears to be.

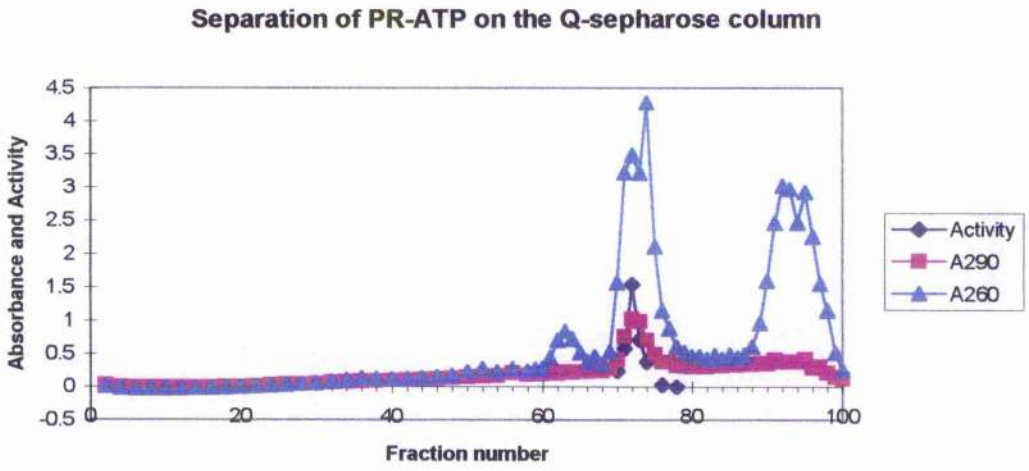


Figure 5.10 Elution profile of PR-ATP from Q-Sepharose column using the Triethylamine formate buffer pH 6

Separation on the Q-sepharose column requires a gradient to 1M Triethylamine formate. The peak containing the PR-ATP is the only peak on the elution profile containing significant A_{290} absorbance. There are two other peaks both absorbing in the A_{260} which are presumed to be unreacted ATP and a breakdown product of PR-ATP. There is also the possibility that the small peak before the PR-ATP could be ADP or AMP (these may be breakdown products of ATP)

5.5.3 Gel filtration of substrate

The lyophilized material containing sodium formate was resuspended in a small volume of distilled water. This was loaded onto a Sephadex G15 column (2 x 80cm) which had been previously equilibrated with distilled water. The column was attached to a Shimadzu LC10 HPLC system and run at a flow rate of 2ml per minute. The material coming off the column was monitored for A_{290} absorbance, collected and lyophilised overnight. This desalting step was successful in removing the sodium formate (and also LiCl, imidazole) and NMR analysis showed the material was pure PR-ATP.

5.5.4 NMR analysis of gel filtered PR-ATP

NMR analysis of the purified PR-ATP samples was carried out by Dr D. Rycroft, Dept. of Chemistry, Glasgow University. Samples of purified PR-ATP were spiked with a known quantity of sodium formate. The formate gave a signal in the region 8.4ppm (Figure 5.11). Integration of this peak and comparison with the signals for PR-ATP allowed the calculation of the amount of PR-ATP present in the sample. This sample was shown to contain only PR-ATP with no other salts present. The important signals for the PR-ATP molecule are as follows;

6.0ppm

6.1ppm

8.6ppm

8.7ppm

These four key signals appear to be characteristic for PR-ATP and allow its identification by NMR analysis. The other peaks present in the NMR trace are

associated with the ribose sugar moieties in both the ATP part of the molecule as well as the phosphoribose part. This region of the spectrum is highly complex and difficult to interpret.

5.5.5 Analysis by mass spectrometry

A mass spectrum of the purified PR-ATP was obtained using a Shimadzu Class 8000 LC-Mass spectrometer. This machine was used in negative ion mode. A molecular ion of 718, corresponding exactly to the mass of PR-ATP^{-1}, was observed along with molecular ions for the mono, di and tri sodium forms of the substrate. (Figure 5.12). The purified PR-ATP had therefore been purified as the sodium salt. The masses of a number of standards including ADP and ATP were determined. A molecular ion a 506 was observed in small amounts in the PR-ATP sample and corresponds to the molecular ion for ATP. This is believed to have arisen from the fragmentation of the PR-ATP molecule in the chamber of the mass spectrometer.

The sample was analysed by direct injection into the mass spectrometer and also by injection through a Partisil column. One clean peak was observed by this combined LC-MS technique.

5.6 Search for a more volatile buffer system

In order to simplify the chromatography attempts were made to find a more volatile buffer. There are few volatile buffers which have the capacity to buffer within the pH boundaries of these experiments. Although it was slightly outwith the desired pH range for the PR-ATP molecule (pH 4 to 7) Triethylamine bicarbonate was chosen. This buffer can buffer in the pH range 7 to 10 (Perrin and Dempsey, 1974) but when adjusting the pH of a 1M solution with solid carbon dioxide it is only

Figure 5.11 400mHz Proton NMR spectrum of purified PR-ATP

The key signals for the PR-ATP molecular are as follows;

6.0ppm

6.1ppm

8.6ppm

8.7ppm

The complex region of the spectrum around 4.5ppm is due to the ribose sugar moiety and is very difficult to interpret.

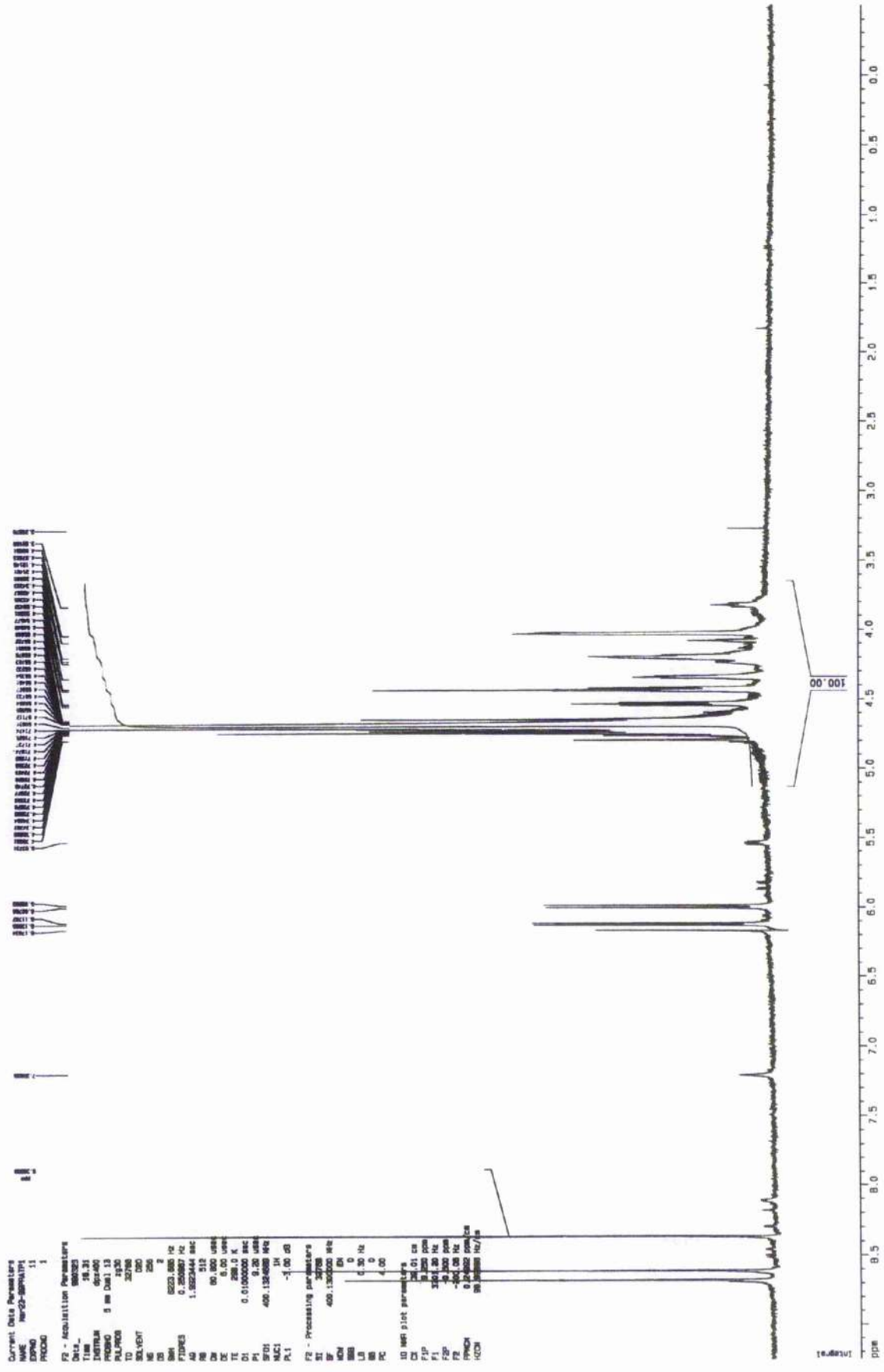

100.00
ppm

Figure 5.12 Enlargement of the 400mHz Proton NMR spectrum of purified PR-ATP showing the key peaks

The important signals for the PR-ATP molecule are as follows;

6.0ppm

6.1ppm

8.6ppm

8.7ppm

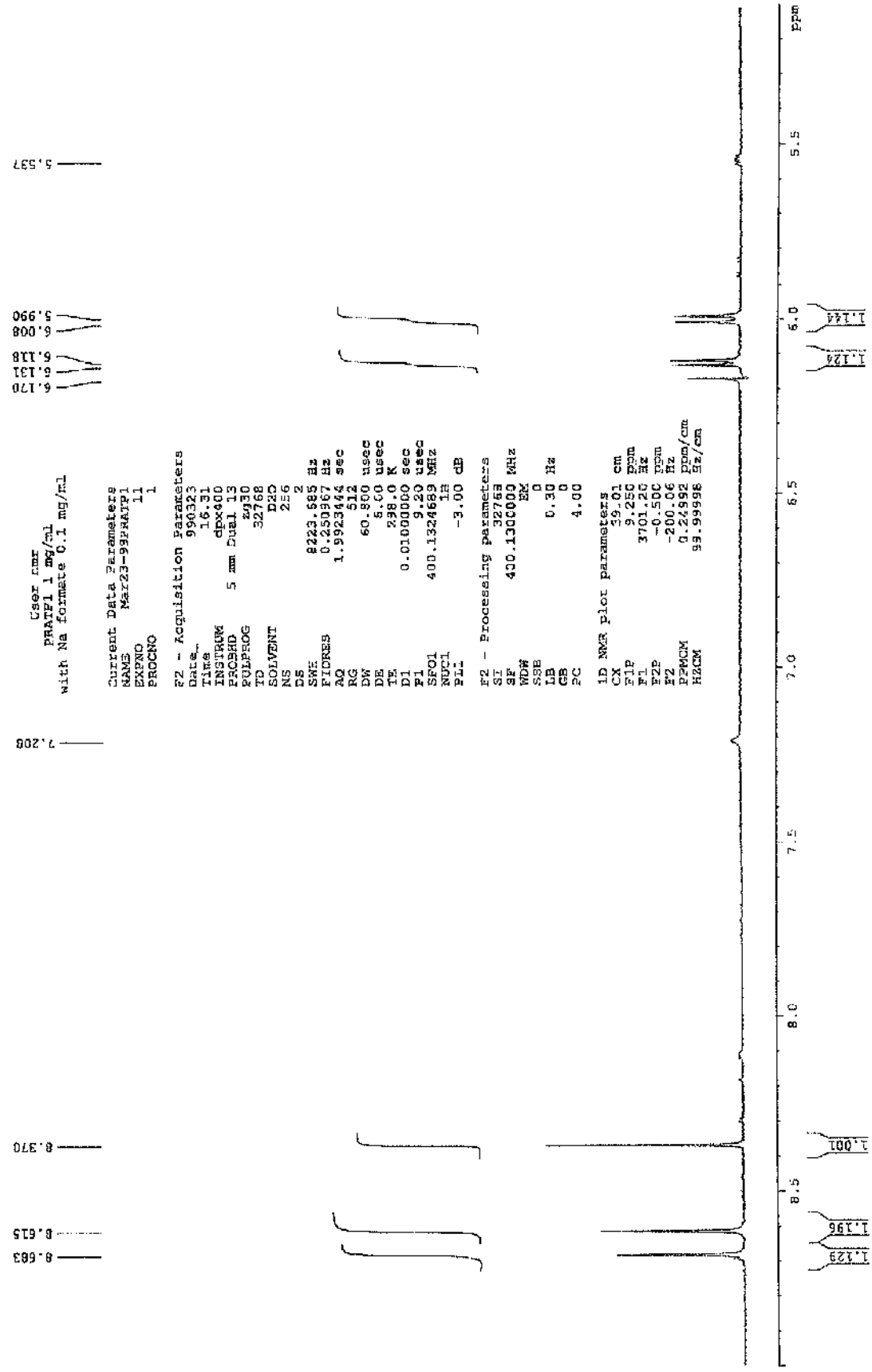
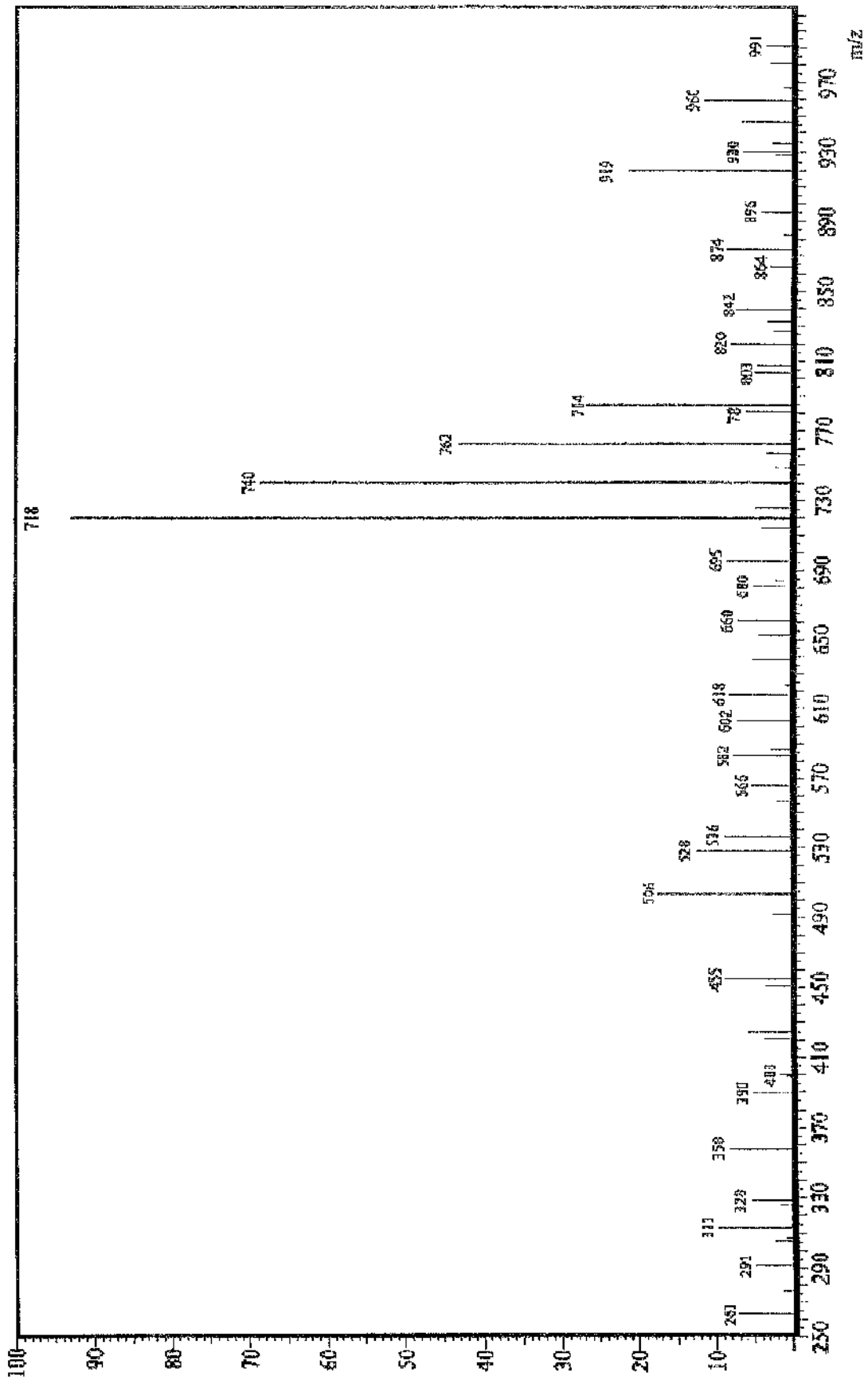

User nmr
PRATP1 1 mg/ml
with Na formate 0.1 mg/ml
Current Data Parameters
NAME Mar23-99PRATP1
EXPNO 11
PROCNO 1
F2 - Acquisition Parameters
Date_ 990323
Time 16.31
INSTRUM dpx400
PROBHD 5 mm Dual 13
PULPROG zg30
TD 32768
SOLVENT D2O
NS 256
DS 2
SWH 8223.685 Hz
FIDRES 0.250967 Hz
AQ 1.9923444 sec
RG 512
DW 60.800 usec
DE 6.00 usec
TE 298.0 K
D1 0.01000000 sec
P1 9.20 usec
SFO1 400.1324689 MHz
NUC1 1H
PL1 -3.00 dB
F2 - Processing parameters
SI 32768
SF 400.1300000 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 4.00
1D NMR plot parameters
CX 19.01 cm
F1P 9.250 ppm
F1 3701.20 Hz
F2P -0.500 ppm
F2 -200.06 Hz
PPMCM 0.24992 ppm/cm
HZCM 99.99998 Hz/cm
8.370
8.615
8.683
7.208
6.170
6.131
6.118
6.008
5.990
5.537
1.001
1.196
1.129
1.144
1.129
8.5
8.0
7.5
7.0
6.5
6.0
5.5
ppm

Figure 5.13 Mass spectrum of PR-ATP

Mass spectrum obtained from a purified extract of PR-ATP. The molecular ion at 718 corresponds to the PR-ATP molecule. The mono-sodium form can be seen at 740, the di-sodium at 762 and the tri-sodium form at 784.


100
90
80
70
60
50
40
30
20
10
250
290
330
370
410
450
490
530
570
610
650
690
730
770
810
850
890
930
970
m/z
718
740
762
784
919
960
991

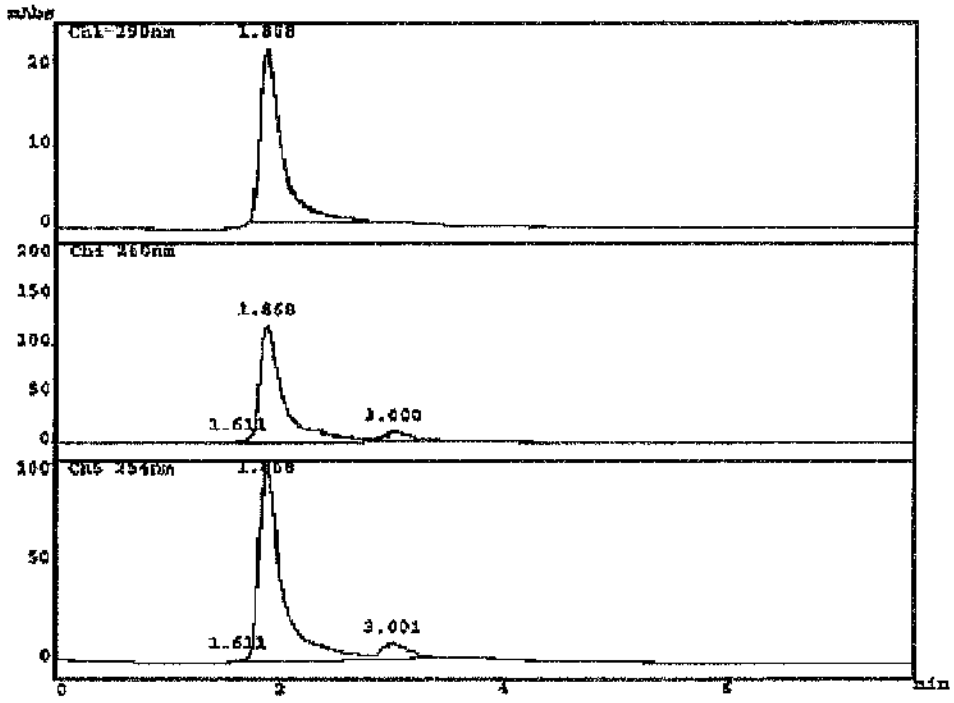


Figure 5.14 HPLC analysis of pure PR-ATP using the Partisil ODS3 column and 25mM potassium phosphate buffer

PR-ATP has a retention time of 1.8 minutes. The small amount of A260 absorbing material is due to the natural breakdown of PR-ATP to ATP and phosphoribose.

Time (min)	%B
0	0
8	0
9	100
14	100
15	0
20	0
20	stop

Buffer A is 25mM Potassium phosphate pH 6 and the flow rate was 2ml per minute. Buffer B is 100% methanol and is used to remove residual material bound to the column

possible to titrate the pH of the solution to 7.3. This pH is slightly high, but the compound appears to be relatively stable in these conditions. This buffer was used for the ion exchange chromatography on Q-Sepharose. The column was equilibrated with 50mM triethylamine bicarbonate pH 7.3. The biotransformation reaction mixture was diluted 20 fold with triethylamine buffer until it had the same conductivity as the equilibrating buffer. This was applied to the column and was eluted with a 1 litre linear gradient of 50mM to 1M triethylamine bicarbonate. 10ml fractions were collected. The PR-ATP eluted around 800mM triethylamine bicarbonate. This material was lyophilised overnight after which a white fluffy residue observed on the walls of the flask. This material became brown and sticky upon exposure to the moisture in the air. The buffer was almost completely removed during the first overnight drying. Subsequent drying was kept to a minimum to prevent breakdown of the PR-ATP. Analysis by NMR and HPLC showed the sample to be pure (Figure 5.13).

Chapter 6

Characterisation of the HisIE enzyme

6.1 Enzyme assay of HisIE

The *Arabidopsis* HisIE enzyme is a bifunctional enzyme which converts PR-ATP to Pro-FAR. The two reactions performed by the enzyme are the hydrolysis of the pyrophosphate group attached to the ATP part of the molecule and secondly the hydrolysis of the purine ring. When the enzyme was assayed as described in section 2.10.2 using a diode array spectrophotometer several changes were seen to occur during the course of the assay. The expected increase in absorbance at 290nm was observed and is due to the formation of the product Pro-FAR. A simultaneous decrease in the absorbance at 260nm is observed which occurs at the same rate as the increase in A_{290}. The hypothesis is that the decrease in A_{260} corresponds to the second step of the reaction which catalyses the hydrolysis of the purine ring. It should therefore be possible to monitor the two steps of the reaction by characterising these two changes independently. A third change is observed at 235nm. The increase in absorbance at 235nm may be due to the release of pyrophosphate by the pyrophosphatase activity of the enzyme. A typical profile from an enzyme assay is shown in Figure 6.1. The presence of similar spectral changes was reported by Davisson *et al.*, (1994). The paper from Davisson's laboratory is concerned with the synthesis and purification of Pro-FAR, which was generated from PRPP and ATP using a partially purified mixture of enzymes. This procedure was in effect a multiply coupled enzyme assay and was used to monitor the progress of the reaction over the course of 1.5 hours. In contrast, our objective was to develop a direct enzyme assay for the HisIE enzyme using purified PR-ATP as the substrate.

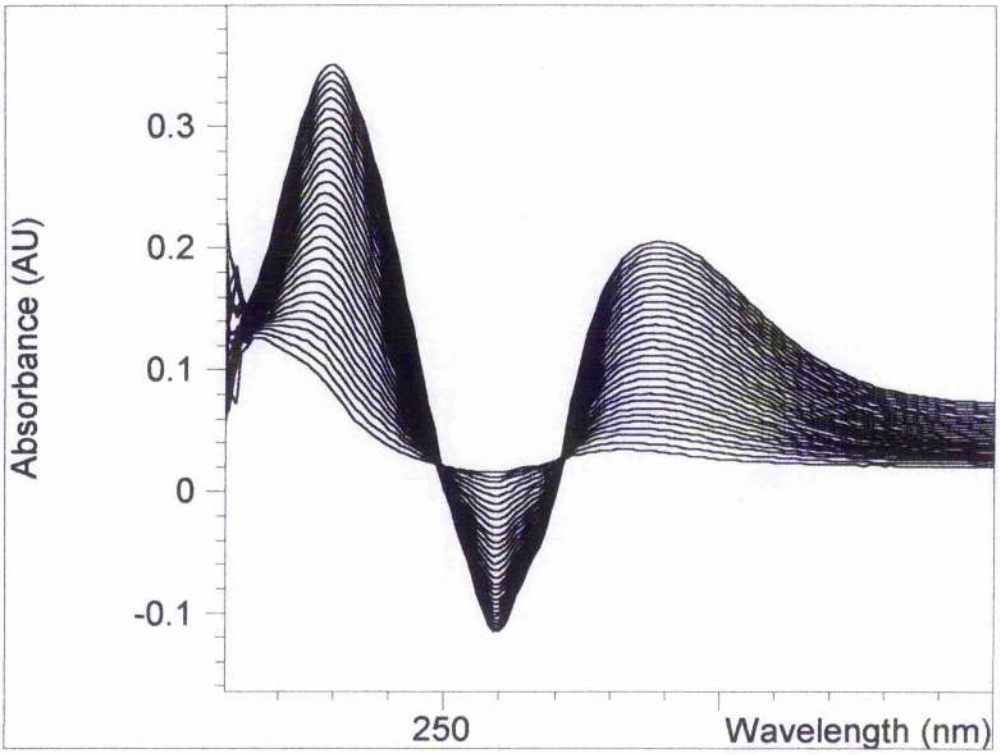


Figure 6.1 Profile observed during assay of the HisIE enzyme

The enzyme was assayed using a Hewlett Packard UV-visible spectrophotometer as described in section 2.10.2. This figure represents the normal pattern of absorbance changes that occur during the course of the enzyme assay. There are three changes occurring simultaneously. The increase at 290nm is due to the formation of the enzyme product Pro-FAR. The two other changes that occur are the decrease in the A_{260} and the increase in the A_{235}. The change in the A_{260} is thought to be due to the hydrolysis of the purine ring of PR-ATP, the third step of the pathway which is carried out by the N terminal domain of the bifunctional HisIE enzyme. The increase in absorbance at 235nm is presumed to correspond to the release of the pyrophosphate group due to the action of the C terminal region of the HisIE enzyme.

6.2 Characterisation of the HisIE enzyme

Following initial observations of enzymatic activity, a more detailed investigation of the kinetic and biochemical properties of this enzyme was necessary. The initial attempts to assay the HisIE enzyme used the same conditions as previously used for the phosphoribosyl ATP transferase enzyme. Certain parameters, such as pH and salt concentration were investigated in an attempt to find the most suitable assay conditions.

6.2.1 Investigation of the effect of pH on the activity of the HisIE enzyme

The enzyme was assayed in Tris buffer at a number of different pHs. The pH of the buffer solutions ranged from pH 6 to pH 9 in 0.5 increments. All other conditions in the assay were kept constant and the rate observed during the assay was plotted against pH (Figure 6.2). The results indicate that the enzyme prefers slightly alkaline conditions and in particular pH 8.5 seemed most favourable. Above pH 8.5 there is a decrease in the rate of enzyme activity and below pH 7 there is no detectable enzymatic activity. This lack of activity below pH 7 could be due to a histidine residue becoming protonated under acidic conditions. A deprotonated histidine residue may therefore be required for enzymatic activity. The histidine may accept a proton from a water molecule thereby generating a hydroxyl group which can then either hydrolyse the phosphodiester linkage to release pyrophosphate or hydrolyse the purine ring opening. The decrease in activity above pH 8.5 could be the result of the deprotonation of a cysteine, or a tyrosine or even the amino group of the N-terminal amino acid resulting in the loss of a crucial hydrogen bonding function in the active site.

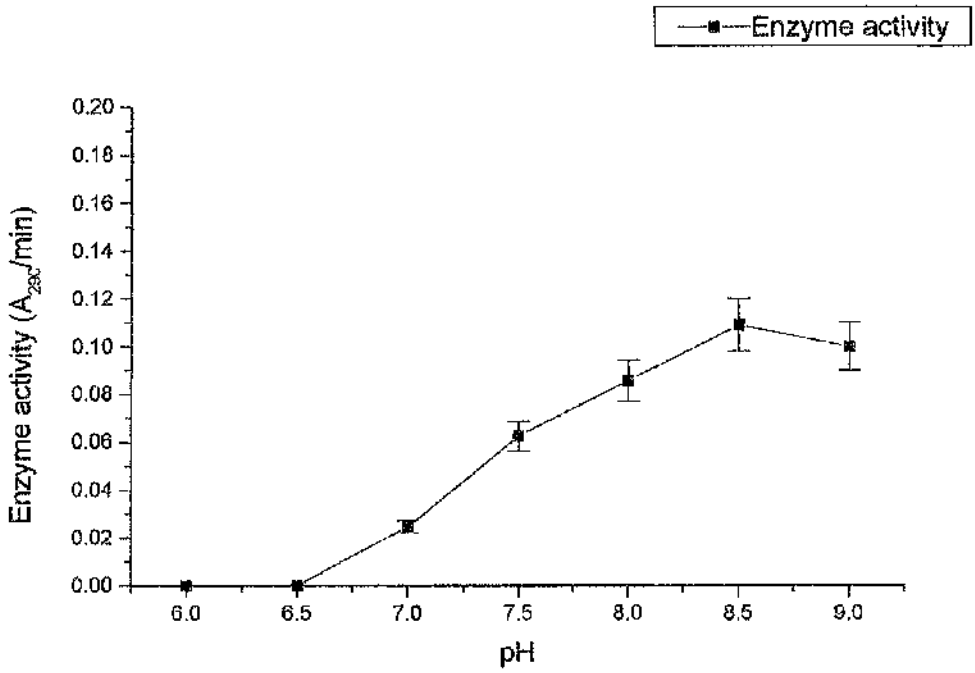


Figure 6.2 Variation in the activity of the *Arabidopsis* HisIE enzyme with pH

Assays were carried out in the presence of 100mM Tris HCl at a number of different pHs. Also present in the assay buffer were 10mM $MgCl_2$ and 150mM KCl. Assays were carried out in a final volume of 1ml and at 25°C. The enzyme activity was measured and plotted against pH. Below pH 7 it was not possible to detect any enzyme activity. The highest activity was observed at pH 8.5 and as the pH was further increased the activity began to decrease. Future assays were carried out at pH 8.5.

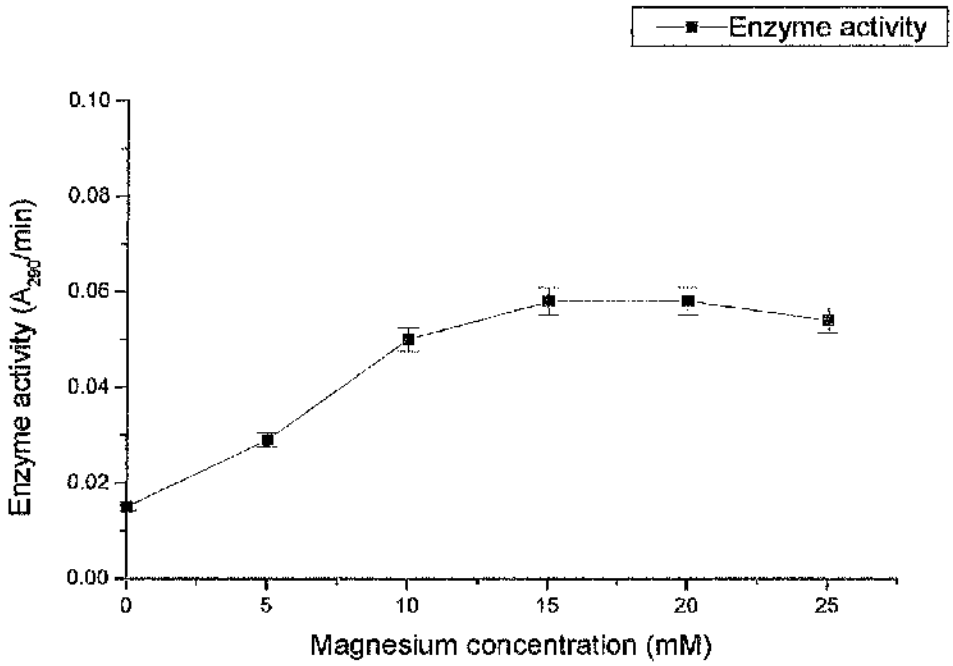


Figure 6.3 Activity of *Arabidopsis* HisIE enzyme plotted against magnesium concentration

The enzyme was assayed in the presence of 100mM Tris HCl pH 8.5 and 150mM KCl at 25°C in a final volume of 1ml. The concentration of $MgCl_2$ in the assay varied from 0-25mM. A plot of enzyme activity versus Mg^{2+} concentration indicates a requirement for free Mg^{2+} for enzyme activity. Enzyme activity increases with increasing Mg^{2+} concentration until 15mM Mg^{2+}, further increases in Mg^{2+} concentration appear to be inhibitory causing a decrease in enzyme activity. 10mM $MgCl_2$ was chosen for all subsequent assays.

6.2.2 Investigating the importance of Magnesium in the assay

It was assumed that the HisIE enzyme would require the presence of a divalent cation such as magnesium to aid binding of the substrate PR-ATP. The amount of magnesium required for enzymatic activity was investigated using concentrations of Mg^{2+} ranging from 0-25mM. In the absence of Mg^{2+} some residual enzyme activity was detected. This maybe the result of Mg^{2+} already associated with the active site of the enzyme. The presence of increasing amounts of Mg^{2+} resulted in an increase in enzymatic activity until around 15mM, further increases in Mg^{2+} concentration did not present any significant advantage to increasing enzyme activity. At higher Mg^{2+} concentrations the observed enzyme rate was seen to decrease. The presence of Mg^{2+} is therefore essential for the catalytic activity of the enzyme. A Mg^{2+}concentration of between 10 and 15mM gave the most consistent rates and therefore a concentration of 10mM $MgCl_2$ was chosen for all subsequent assays (Figure 6.3).

6.2.3 Requirement of the *Arabidopsis* HisIE enzyme for Zinc

Following publication of the paper by D'Ordine *et al.*, (1999) which identified that the monofunctional HisI enzyme from *M. vannielli* was a zinc containing metallo-enzyme, the enzyme requirement for zinc was investigated. The enzyme was dialysed against a buffer solution containing 10mM 1,10-phenanthroline (a zinc chelating agent), 50mM Tris HCl pH 7.5 and 0.4mM DTT. Samples of enzyme were removed at time points and assayed for the presence of enzyme activity. The enzyme activity decreased to less than 10% of the starting activity following dialysis for 60 minutes in the presence of 1, 10-phenantholine. Efforts were made to add zinc back to the enzyme with the hope of recovering some of the lost activity. It was not possible to recover enzyme activity by addition of zinc either in the assay buffer or by direct addition of

$ZnCl_2$ to the enzyme sample. It would appear that the *Arabidopsis* bifunctional HisIE enzyme is a zinc containing enzyme in the same manner as the monofunctional HisI enzyme from *M. vannielli*. The requirement for Zn^{2+} in the enzyme assay was investigated but it was found that addition of Zn^{2+} to the assay buffer was inhibitory to the enzyme even at low levels. This was consistent with the data presented in the paper by D'Ordine *et al.*, (1999) where the presence of 3-5μM free zinc strongly inhibited the enzyme. The conclusion was drawn that there was enough endogenous zinc associated with the enzyme itself or present as a contaminant within the existing assay buffer to supply the Zn^{2+} requirement for the enzyme.

6.2.4 Investigating the concentration of salt present in the enzyme assay.

The assay for phosphoribosyl ATP transferase had been used as a model for the development of an assay for the HisIE enzyme. The assay for phosphoribosyl ATP transferase contained 150mM KCl. In order to investigate if the HisIE enzyme required the presence of such high quantities of salt the enzyme activity was determined in the presence of salt concentrations ranging from 0-150mM KCl in the presence of 10mM $MgCl_2$. The presence of KCl had a small but significant effect on the assay. The presence of increasing concentrations of KCl led to a decrease in the activity of the enzyme. The highest activity was detected in the absence of KCl (Figure 6.4). All subsequent assays were carried out in the absence of KCl.

6.2.5 Investigation of the buffer concentration required for enzyme activity

The optimal buffer concentration for the enzyme assay was investigated using concentrations of Tris-HCl buffer pH 8.5 varying from 0-100mM. In the absence of buffer no enzyme activity was detected. Enzyme activity increased sharply up to a

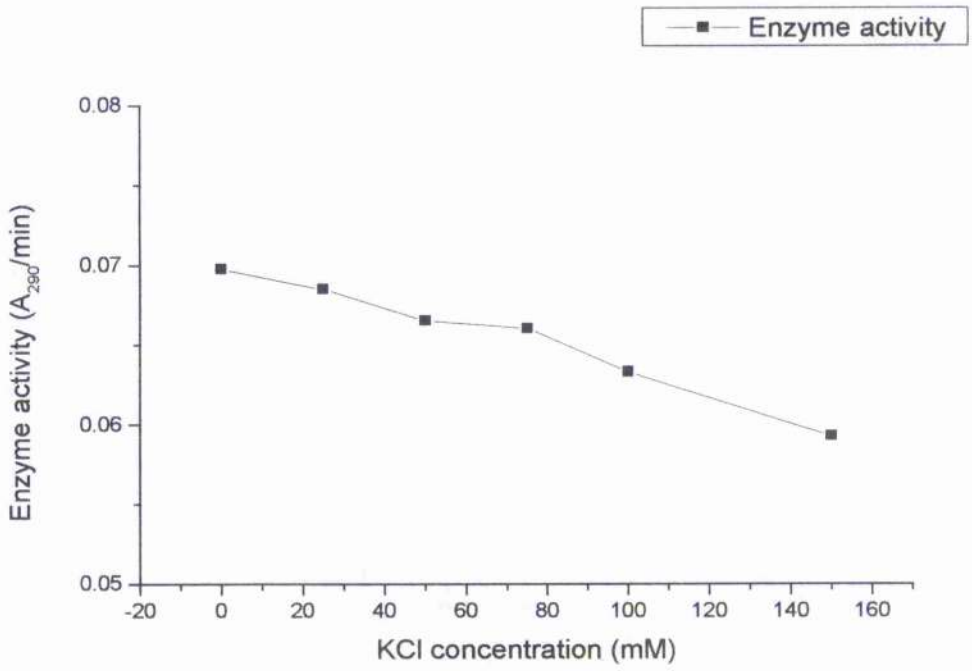


Figure 6.4 Investigation of the effect of increasing concentrations of KCl in the HisIE enzyme assay

The enzyme was assayed in the presence of 100mM Tris HCl pH 8.5 and 10mM $MgCl_2$ at 25°C in a final volume of 1ml. The KCl concentration was varied from 0-150mM KCl. The highest enzyme activity was observed in the absence of KCl in the assay buffer. Increasing concentrations of KCl appeared to have a slight inhibitory effect on the activity of the enzyme as the rate of activity decreased with increasing salt concentration. The enzyme was therefore assayed in the absence of KCl.

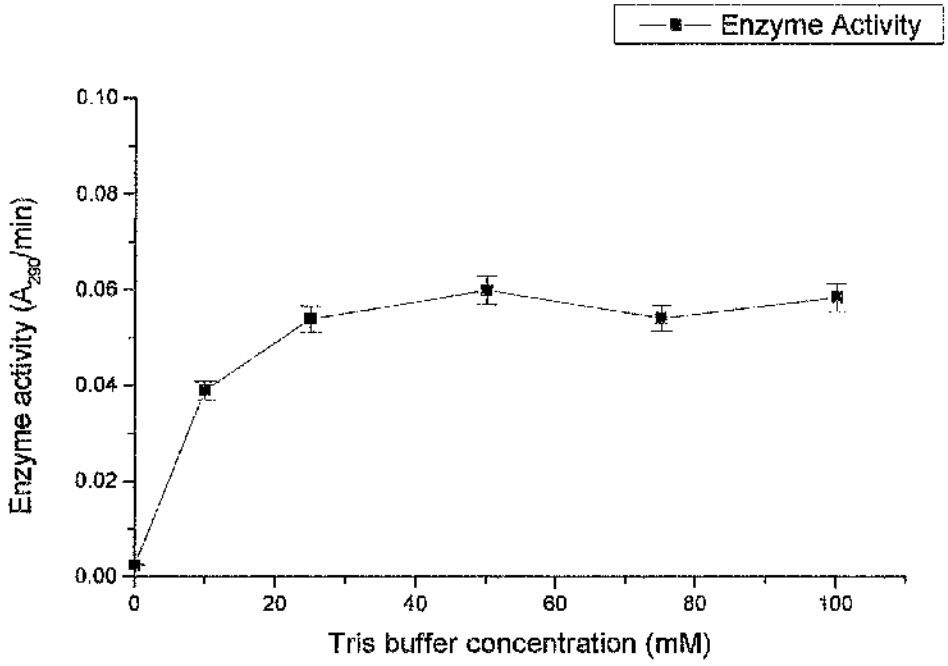


Figure 6.5 Determination of the concentration of buffer required for enzyme activity

The concentration of buffer in the experiment varied from 0-100mM Tris HCl pH 8.5. 10mM $MgCl_2$ was present in all assays which were carried out in a final volume of 1ml at 25°C. No activity was detected in the absence of buffer. Activity increased sharply until a concentration of 50mM Tris HCl pH 8.5. Further increases in buffer concentration did not cause further increases in enzyme activity. A concentration of 50mM Tris HCl pH 8.5 was chosen for all subsequent experiments.

concentration of 50mM Tris, further increases in buffer concentration did not produce an increase in the rate of enzyme activity. Therefore all subsequent assays were carried out in the presence of 50mM Tris HCl pH 8.5 (Figure 6.5).

6.2.6 Final assay conditions

The assay buffer used to detect the presence of HisIE activity was: 50mM Tris HCl pH 8.5, 10mM $MgCl_2$, PR-ATP was added to give the desired final substrate concentration.

The assays were carried out in a final volume of 1ml at 25°C.

6.3 Determining the K_m and k_{cat} of the *Arabidopsis* HisIE enzyme

6.3.1 Determining the K_m

The K_m (Michaelis constant) for the *Arabidopsis* HisIE enzyme was determined using the assay described above, using three concentrations of enzyme and concentrations of PR-ATP in the range 10-100μM. The K_m of an enzyme is equivalent to the substrate concentration at which half of the active sites contain substrate i.e. the rate of the enzyme catalysed reaction is half its maximal value (Stryer, 1988). The K_m for the *Arabidopsis* HisIE enzyme was15μM at pH 8.5 and 25°C (Figure 6.6).

6.3.2 Determining k_{cat}

The k_{cat} (turnover number) for the *Arabidopsis* HisIE enzyme is calculated using the following equation;

$$V_{max} = k_3 [E_T]$$

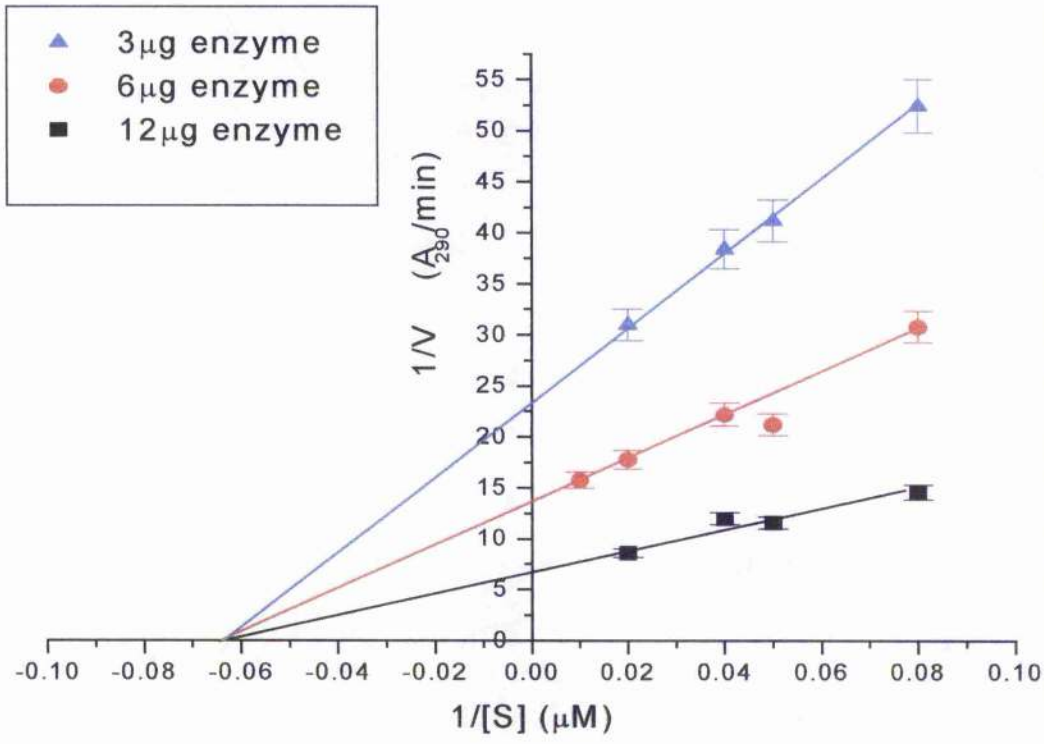


Figure 6.6 Determination of the K_m of the *Arabidopsis* HisIE enzyme at pH 8.5

Double reciprocal plots of initial velocity versus [PR-ATP] at three concentrations of enzyme

All velocity measurements relate to the change in absorbance at 290nm and are plotted as the change in absorbance per minute. Substrate concentration ranged from 12.5-100μM

To calculate rate constants the change in absorbance was converted to concentrations using the difference in the extinction coefficient between substrate and product ($4400M^{-1}s^{-1}$) see page 49.

where, V_{max} is the maximum velocity, k_3 (k_{cat}) the turnover number and $[E_T]$ the concentration of active sites. The enzyme concentration was determined using the Lowry method (Lowry *et al.*, 1951) and was carried out by Mr J. Greene. The turnover number is the number of substrate molecules converted into product by an enzyme molecule in a unit time where the enzyme is fully saturated with substrate (Stryer, 1988). The k_{cat} for the *Arabidopsis* enzyme was found to be 1.244 per second.

6.3.3 k_{cat} / K_m (catalytic efficiency)

The value for k_{cat} / K_m was 8.29×10^4 M^{-1} s^{-1}

6.4 Determination of the K_m in the presence of ATP

The set of enzyme assays used to determine the K_m of the HisIE enzyme were repeated in the presence of varying concentrations of ATP to determine if ATP had an inhibitory affect on the enzyme. Concentrations of 0-1mM ATP were used (Figure 6.7). The presence of ATP slightly altered the K_m of the HisIE enzyme. It appeared to be a weak inhibitor of the enzyme. The result was not as dramatic as would have been predicted considering that the enzyme substrate PR-ATP contains an ATP molecule as a significant part of its structure. It would be useful to repeat this experiment to double check the results of this experiment. The presence of 1mM ATP raised the K_m of the enzyme from 15 to 22μM.

6.4.1 Calculation of K_i for ATP

K_i was determined from a secondary plot of the gradients of each line in the presence of varying concentrations of inhibitor and was calculated as 1.65mM.

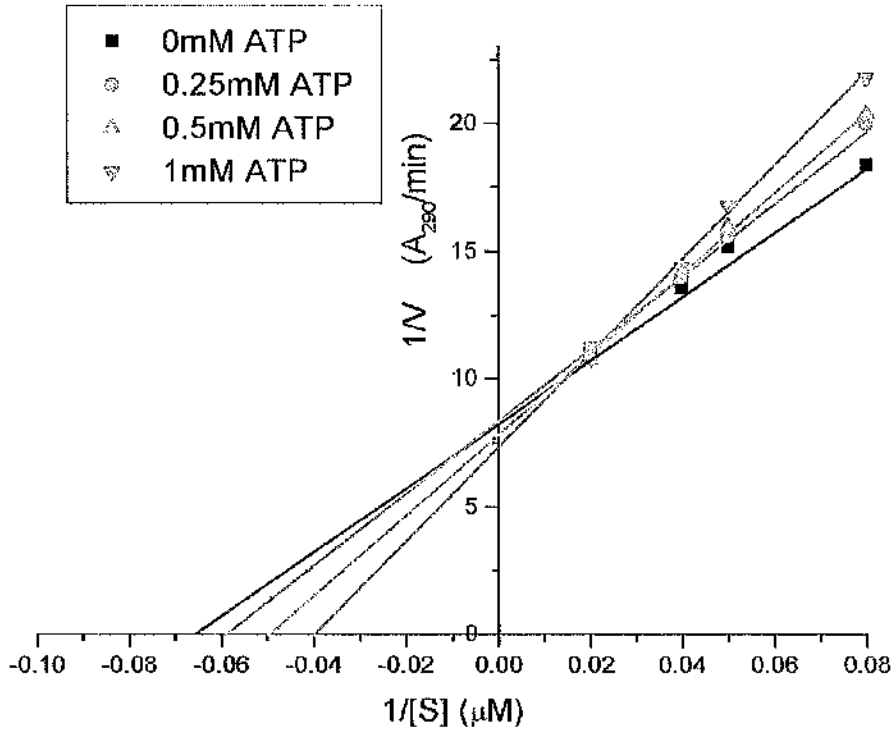


Figure 6.7 Determining the effect of ATP on the activity of the HisIE enzyme

Double reciprocal plot of initial velocity versus [PR-ATP] in the presence of differing concentrations of ATP

All velocity measurements relate to the change in absorbance at 290nm per minute. 6μg of HisIE enzyme was present in all assays. Substrate concentration in the enzyme assays ranged from 12.5 to 50μM PR-ATP.

6.5 Chemical modification of HisIE enzyme

Chemical modification is a technique used to identify residues of functional importance within a protein. There are a number of reagents which target specific groups or side chains of amino acids. Treatment of a protein with these reagents may result in the loss of enzymatic activity if the modified residue is involved in catalysis or binding. The loss of enzymatic activity may not always be due to modification of a residue involved with the active site, it is possible that steric hindrance of the active site may occur by the modification of a surface residue or through a conformational change brought about through the covalent attachment of the modifying group. Thus chemical modification can be used as a tool to identify residues which may be of potential importance within a protein. In order to further identify the function of a modified residue it is essential to use other techniques such as site directed mutagenesis to confirm the function of the residue.

6.5.1 Chemical modification with diethylpyrocarbonate (DEPC)

DEPC is reagent used for the selective modification of histidine residues (Lundblad and Noyes, 1984). The reaction between DEPC and histidine residues yields N-carbethoxyhistidine through the reaction with one of the imidazole nitrogens of histidine. The reaction can be reversed by the addition of hydroxylamine which results in the recovery of the histidine. The formation of carbethoxyhistidine can be monitored by the increase in absorbance at 237-242nm. The number of residues which have been modified can be calculated using the molar absorption difference ($\Delta\varepsilon 240 = 3200 M^{-1} cm^{-1}$; Miles, 1977).

One difficulty with using DEPC is the instability of the reagent in aqueous solution. The half life of DEPC under the normal assay conditions for the HisIE

enzyme is very short, such that within a matter of minutes the DEPC would be almost destroyed. In the presence of Tris-chloride at 25°C and pH 8.2 DEPC has a half life of 0.37 minutes. The half life of DEPC in 60mM sodium phosphate buffer pH 6 at 25°C is 24 minutes (Miles, 1977). The stability of the reagent DEPC is highly variable depending on the type of buffer as well as pH.

6.6 Treatment of the HisIE protein with DEPC

In order to investigate the effect of DEPC on the HisIE protein it was first essential to find a new buffer which could be used to assay the enzyme. The HisIE enzyme is normally assayed in the presence of Tris buffer. Unfortunately because of the reactivity of the primary amino group Tris buffers are not suitable for studying reactions involving DEPC. Therefore a number of different buffers were investigated to find a buffer which was compatible with the HisIE enzyme. The following buffers were investigated; 10mM potassium acetate pH 6.2, 10mM potassium phosphate pH 7.2 and 10mM sodium borate pH 8.5. The enzyme was incubated at 25°C in the presence of each buffer and the activity was monitored by removing aliquots of enzyme and assaying over a 10 minute period. A control containing the normal assay conditions for the HisIE enzyme was used for comparison. The potassium phosphate buffer produced a profile which was similar to that of the Tris buffer. The enzyme was inactivated more rapidly in the presence of potassium acetate buffer and very rapidly in the presence of the sodium borate buffer. All chemical modification experiments were therefore carried out using the following assay buffer;

10mM potassium phosphate pH 7.2, 10mM $MgCl_2$, PR-ATP

Assays were carried out in a final volume of 1ml at 25°C.

6.6.1 Kinetics of DEPC inactivation of the HisIE enzyme

Incubation of the HisIE enzyme with DEPC results in the rapid loss of enzymatic activity. The fraction of activity remaining was calculated as the percentage of activity remaining (A/A_0, where A is the activity at a specific time and A_0 is the initial activity). The percentage activity remaining was plotted against time (Figure 6.8). This was carried out for a number of different concentrations of DEPC. Concentrations as low as 50μM DEPC resulted in a 40% loss of activity within 4 minutes. The inactivation with DEPC shows pseudo first order kinetics over the first few minutes of the time course. The rate of enzyme inactivation is dependent on the DEPC concentration. It was possible to estimate the half time of inactivation at a number of different concentrations from the pseudo first order plot (Figure 6.8) see Table 6.1. The half life of inactivation is the time at which the enzyme activity has decreased to half the initial value. Thus;

$$t_{½} = \ln 2 / k_{obs}$$

The pseudo first order rate constant for inactivation (k_{obs}) can be calculated as follows;

$$k_{obs} = 0.693 / t_{½}$$

The second order rate constant of inactivation can be calculated by plotting the first order rate constant (k_{obs}/min) against concentration of DEPC [DEPC]. (Figure 6.9). The gradient of the plot is the second order rate constant. The second order rate constant of inactivation is $2500 M^{-1} min^{-1}$.

[DEPC] μM	$t_{½}$ (min)	k_{obs}/min
25	10.75	0.064
50	5.5	0.126
75	4.25	0.163
100	1.75	0.396

Table 6.1 Half life ($t_{½}$) and first-order inactivation rate constant (k_{obs}) for the inactivation of *Arabidopsis* HisIE at varying concentrations of DEPC

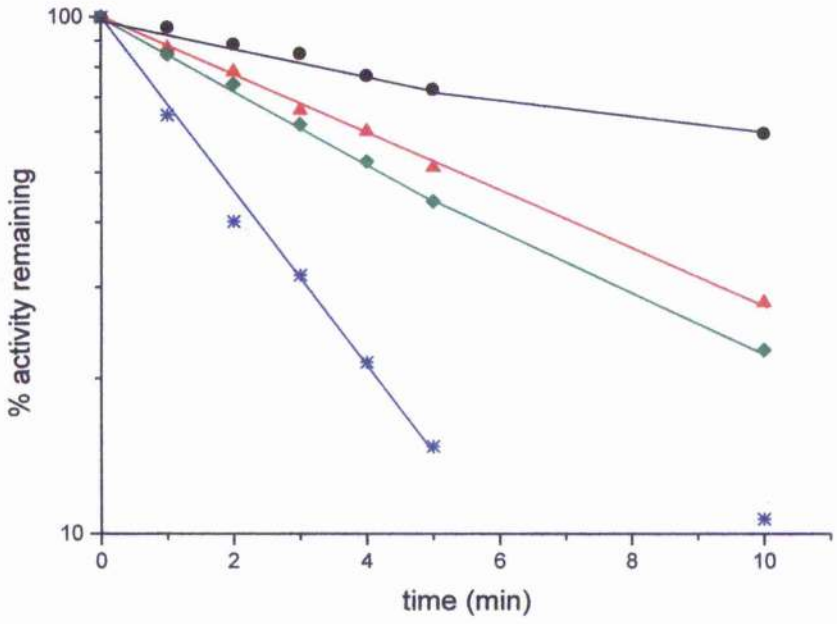


Figure 6.8 Inactivation of HisIE with DEPC

HisIE enzyme (13.4μM) was incubated with increasing concentrations of DEPC in 10mM potassium phosphate buffer pH 7.2 at 25°C. Aliquots were removed at various time points and the activity determined. The activity, adjusted for the observed activity changes in the control, is plotted as the percentage of A_{290} activity remaining on a logarithmic scale. The concentrations of DEPC used were,

● 25μM DEPC, ▲ 50μM DEPC, ◆ 75μM DEPC, ✳ 100μM DEPC

The same profile of inactivation is observed if the A_{260} values are plotted.

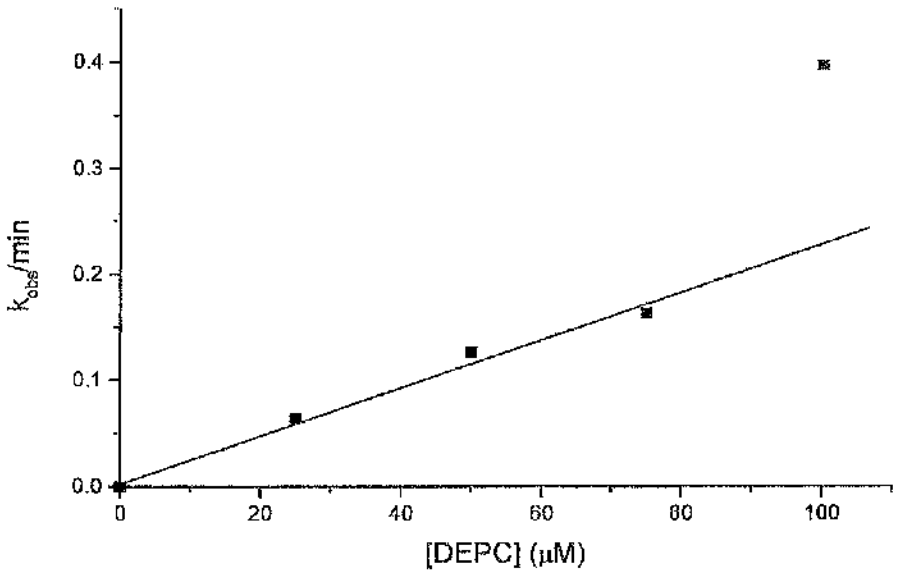


Figure 6.9 Determination of the second order rate constant of inactivation

The gradient of this line, calculated to be 2500M^{-1} min^{-1} is the second order rate constant (k) of inactivation.

The reaction in the presence of 100μM DEPC is very fast and it is difficult to determine the slope and the half life of the reaction, and so this data point cannot be plotted with the other data.

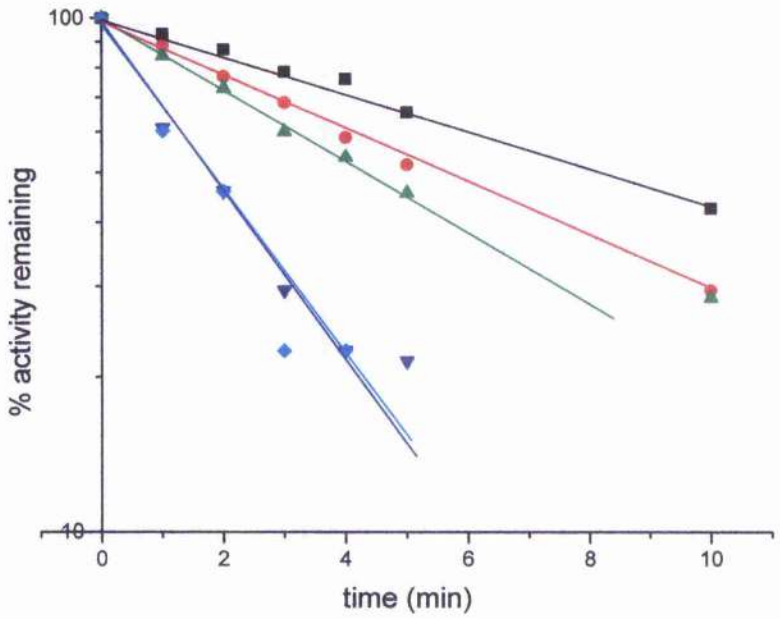


Figure 6.10 Substrate protection of HisIE against inactivation by DEPC- pseudo first order plot

HisIE enzyme (13.4μM) was incubated with 100μM DEPC and increasing concentrations of substrate to determine if protection against inactivation could be achieved. Enzyme was incubated in 10mM potassium phosphate buffer pH 7.2 at 25°C. Aliquots were removed at various time points and the activity determined. The activity, adjusted for the observed activity changes in the control, is plotted as the percentage of A_{290} activity remaining on a logarithmic scale. The concentrations of substrate used were,

■ 400μM PR-ATP, ● 200μM PR-ATP, ▲ 100μM PR-ATP, ▼ 50μM PR-ATP, ◆ 0μM PR-ATP

A similar profile of inactivation is observed if the A_{260} values are plotted.

6.7 Substrate protection against DEPC inactivation

The presence of enzyme substrate may provide a means of protecting an enzyme from inactivation. If inactivation is the result of modification of an active site residue, the presence of substrate in the active site may prevent or slow this inactivation. As DEPC proved to be a very good reagent for inactivation the experiments were repeated in the presence of varying concentrations of PR-ATP to test if substrate was able to provide protection. The substrate concentrations used in protection experiments ranged from 50-400μM PR-ATP; the concentration of DEPC used in each case was 100μM (Figure 6.10). The presence of PR-ATP did provide some protection against inactivation but only in the presence of relatively high amounts. The percentage protection offered by the presence of substrate was calculated and is shown in Table 6.2. The K_s value was determined as 175μM at pH 7.2 (Figure 6.11). The K_s is approximately 10 times the K_m for the enzyme and may be the result of hydrolysis of substrate to product, allowing the product to provide protection.

[Substrate]μM	$t_{1/2}$ (min)	k_{obs}/min	% protection
0	1.75	0.396	none
50	1.75	0.396	none
100	4.4	0.1575	60
200	5.6	0.123	69
400	8.1	0.0855	78.4

Table 6.2 Effect of substrate concentration on inactivation of HisIE by DEPC

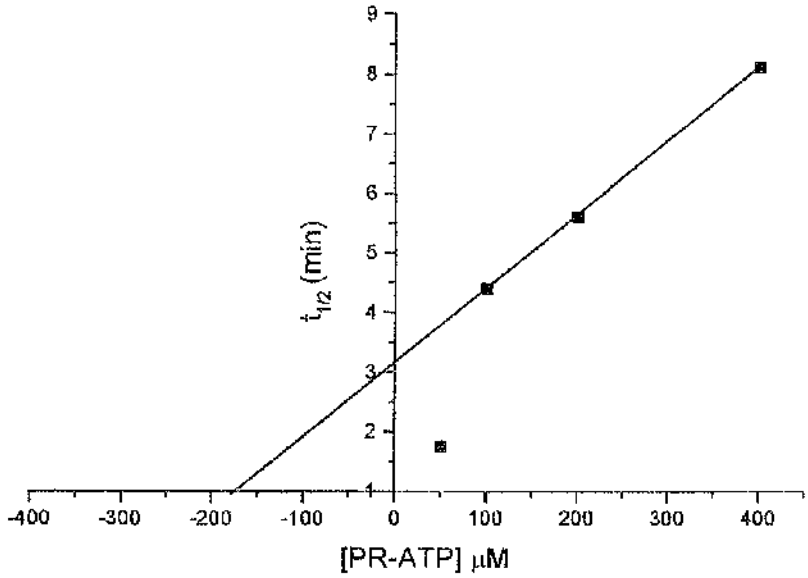


Figure 6.11 Determination of K_s for PR-ATP

The half life of the enzyme as calculated in Figure 6.10 was plotted against the concentration of PR-ATP. From the intercept the K_s value is 175μM.

The K_s value is 10 times the calculated K_m of the enzyme. The hydrolysis of substrate to product may be occurring and it is the available product which is providing protection against inactivation.

6.8 Identification of the modified histidine residues

There are two histidine residues in the *Arabidopsis* HisIE protein. This work demonstrates that the histidine-directed reagent DEPC inactivates the enzyme rapidly and suggests that, because substrate affords some protection against inactivation, one or both of the histidines may have a role in the enzyme activity. The difference spectrum of DEPC modified HisIE and untreated enzyme shows an absorbtion maximum at 239nm, which is characteristic of the carbethoxylation of histidine residues (Miles, 1977). Due to the rapid inactivation (more than 10 times that observed for the type I dehydroquinases, Deka *et al.*, 1992, Moore *et al.*, 1993) it was difficult to establish the stochiometry of the reaction by direct observation of the change in absorbance at 240nm. Indications are that one molecule of DEPC reacts with one molecule of protein as calculated using the extinction coefficient for the reaction ($\Delta\varepsilon 240 = 3200 M^{-1}cm^{-1}$; Miles, 1977). It was also not possible to use mass spectrometry to determine whether one or both histidine residues were being modified because the bond between DEPC and histidine residues is very labile and as a result it is not possible to determine the sites of modification after proteolytic cleavage. It has also been shown that modification with DEPC is not completely specific for histidine residues for example Krell *et al.*, 1998 showed that DEPC modification of Shikimate dehydrogenase resulted in more modified species than there were histidine residues in the protein. Both residues are well conserved across a number of different species, Histidine 146 in domain 1 (the HisI domain) is absolutely conserved across all species. Histidine 245 in domain 2 (the HisE domain) is conserved apart from a few of the fungal species, where in this case this residue is substituted with a phenylalanine residue (refer to Figure 3.2). To confirm whether the histidine residues were essential for activity site directed mutagenesis was carried out (see Chapter 7).

6.9 Chemical modification of the HisIE protein with TNBS

TNBS is a reagent used for the selective modification of lysine residues (Lundblad and Noyes, 1984). Modification with TNBS can be followed by spectral analysis at 420 or 367nm (Lundblad and Noyes, 1984).

6.9.1 Treatment of the HisIE enzyme with TNBS

The HisIE enzyme was incubated in the presence of 10mM potassium phosphate pH 7.2 , 10mM $MgCl_2$ and PR-ATP in a final volume of 1ml at 25°C for the modification experiments with TNBS.

6.9.2 Kinetics of TNBS inactivation of the HisIE enzyme

The incubation of the HisIE enzyme with TNBS resulted in the rapid loss of enzymatic activity. The concentration of TNBS used in the experiments ranged from 0.1-1mM. The percentage activity remaining was calculated and was plotted against time (Figure 6.12). Incubation with 0.5mM TNBS resulted in the loss of 50% of the starting activity within 6.5 minutes. The inactivation with TNBS shows pseudo first order kinetics over the first few minutes of the time course. The rate of inactivation was dependent on the TNBS concentration (Figure 6.12)and the half times of inactivation at a number of different concentrations were estimated from the pseudo first order plot (Figure 6.12, Table 6.3). The second order rate constant was calculated by plotting the first order rate constant (k_{obs}/min) against concentration of TNBS. A value of $210M^{-1}min^{-1}$ was obtained.

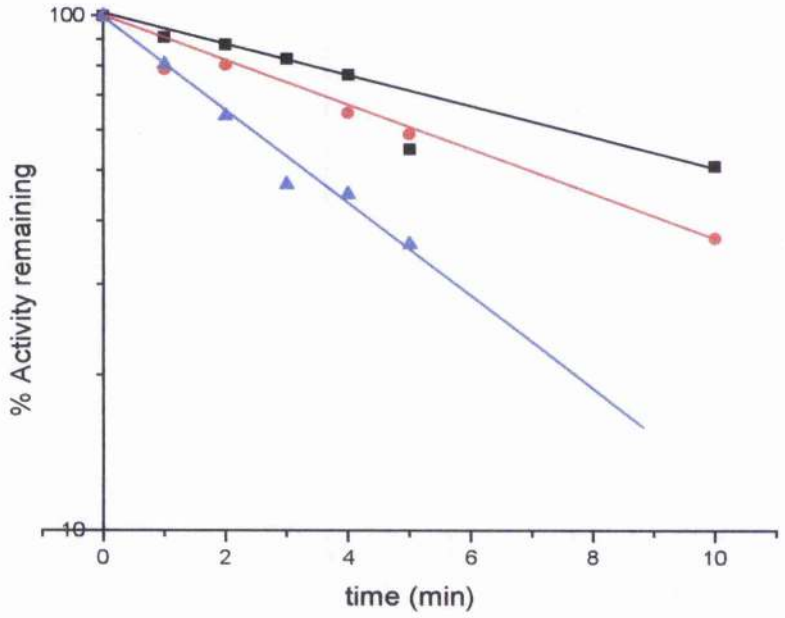


Figure 6.12 Inactivation of HisIE with TNBS

HisIE enzyme was incubated with increasing concentrations of TNBS in 10mM potassium phosphate buffer pH 7.2 at 25C. Aliquots were removed at various time points and the activity determined. The activity is plotted as the percentage activity remaining on a logarithmic scale. The concentrations of TNBS used in the experiments were,

■ 0.1mM TNBS, ● 0.5mM TNBS, ▲ 1mM TNBS.

The lines have been corrected for the observed activity changes in the control.

[TNBS] mM	$t_{1/2}$ (min)	k_{obs}/min
0.1	10.4	0.066
0.5	6.5	0.11
1	3.3	0.21

Table 6.3 Half life ($t_{1/2}$) and first order inactivation rate constant k_{obs}/min for the inactivation of the *Arabidopsis* HisIE enzyme at varying concentrations of TNBS

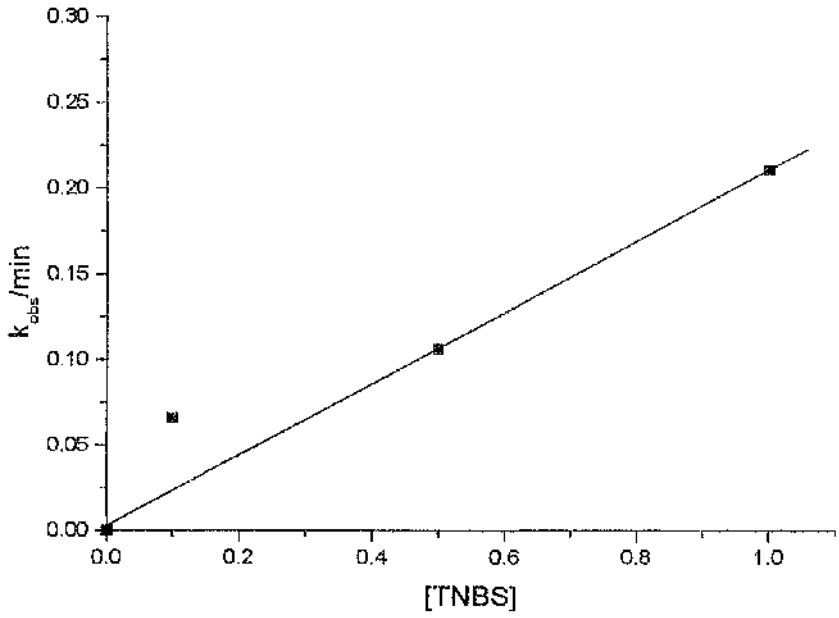


Figure 6.13 Determination of the second order rate constant of inactivation

The gradient of this line, calculated to be $210M^{-1}min^{-1}$ is the second order rate constant (k) of inactivation.

6.9.3 Substrate protection against TNBS inactivation

The TNBS modification experiments were repeated in the presence of varying concentrations of PR-ATP to test if substrate was able to provide protection against inactivation. The substrate concentrations in the protection experiments ranged from 0-400μM PR-ATP; the concentration of TNBS in each case was 1mM. The presence of PR-ATP did provide protection against inactivation. It was observed however that increasing the concentration of PR-ATP in the incubation mixture resulted in a more rapid inactivation of the HisIE enzyme (Figure 6.14). The implication of this result is that substrate binding makes the reactive side chains on the enzyme more susceptible to chemical modification. Also that the key site of modification that leads to inactivation is not within the substrate binding pocket.

6.9.4 AMP protection against TNBS inactivation

Since PR-ATP was able to provide protection against inactivation but only at lower concentrations, attempts were made to identify other molecules which might protect the enzyme from inactivation. The product of the HisIE enzyme is not commercially available and it was not an objective of this project to synthesise and purify this molecule. The product of the HisIE enzyme contains an AMP molecule as part of its structure, therefore it seemed logical to test if AMP was able to provide protection against inactivation. The concentrations of AMP in the experiments ranged from 0-2mM. TNBS was used at a concentration of 1mM for all experiments (Figure 6.15) The presence of AMP was able to provide protection against inactivation with TNBS. The percentage protection offered by the presence of AMP was calculated and is shown in Table 6.4. The Ks value for AMP was determined as 0.37mM at pH 7.2 and 25°C (Figure 6.16)

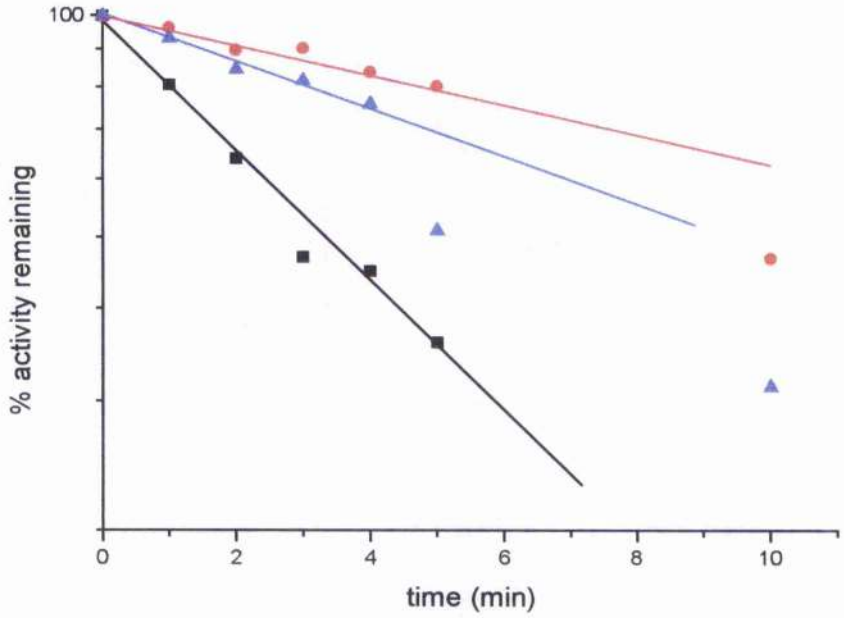


Figure 6.14 Substrate protection of HisIE against inactivation by TNBS

HisIE enzyme was incubated with 1mM TNBS and increasing concentrations of PR-ATP to determine if protection against inactivation could be achieved. The enzyme was incubated in 10mM potassium phosphate buffer pH 7.2 at 25°C. Aliquots were removed at various time points and the activity determined. The activity is plotted as percentage of activity remaining on a logarithmic scale. The concentrations of PR-ATP used were,

■ 0μM PR-ATP, ● 200μM PR-ATP, ▲ 400μM PR-ATP

Adjusted for the observed activity changes in the control.

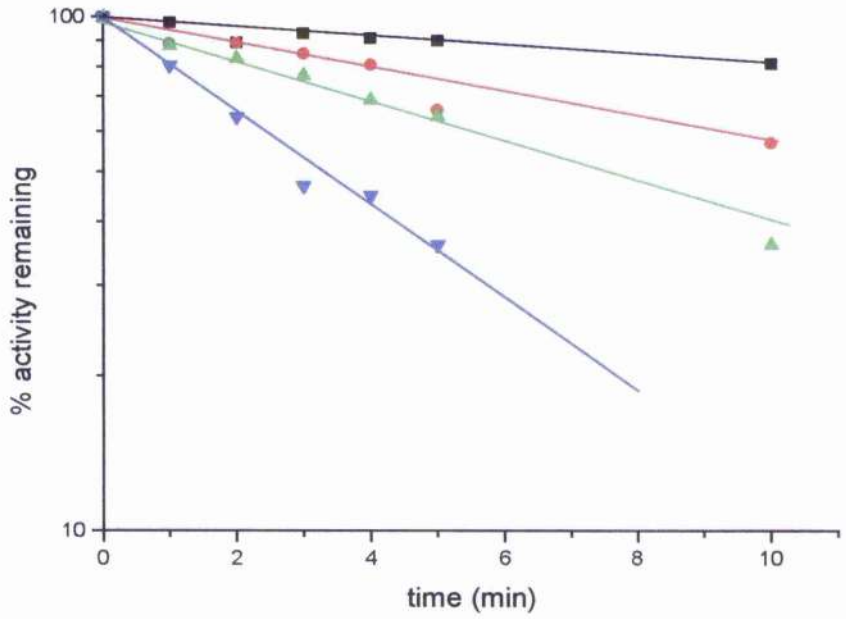


Figure 6.15 AMP protection of HisIE against inactivation by TNBS

HisIE enzyme was incubated with 1mM TNBS and increasing concentrations of AMP to determine if protection against inactivation could be achieved. The enzyme was incubated in 10mM potassium phosphate buffer pH 7.2 at 25°C. Aliquots were removed at various time points and the activity determined. The activity is plotted as percentage of activity remaining on a logarithmic scale. The concentrations of AMP used were,

■ 2mM AMP, ● 1mM AMP, ▲ 0.5mM AMP, ▼ 0mM AMP

Adjusted for the observed activity changes in the control.

[AMP]	$t_{1/2}$	k_{obs}/min	% protection
0	3.3	0.21	none
0.5	7.8	0.089	58
1	12.4	0.056	73
2	31	0.022	90

Table 6.4 Effect of AMP concentration on inactivation of HisIE by TNBS

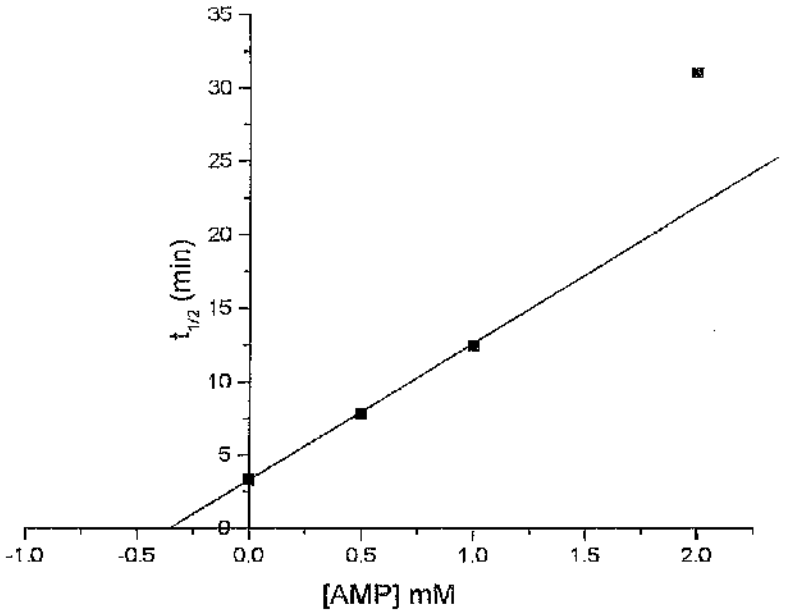


Figure 6.16 Determination of K_s for AMP

The half life of the enzyme as calculated in Figure 6.15 was plotted against the concentration of AMP. From the intercept the K_s value is 0.37mM.

Chapter 7

Site directed mutagenesis of two conserved histidine residues and kinetic analysis of both mutants

7.1 Introduction

The work described in Chapter 6 indicated that one or more histidine residues may be involved in the catalytic activity of the HisIE enzyme. There are two histidine residues within the *Arabidopsis* HisIE protein, Histidine 146 and Histidine 245 (numbered according to the full length *Arabidopsis* hisIE sequence, Figure 3.1). His146 is absolutely conserved across all known species, whereas His245 is conserved in most species with the exception of a few of the fungal species. At the time of carrying out the chemical modification experiments it was impossible to determine which histidine residue was modified by DEPC. A site directed mutagenesis strategy was applied in order to mutate each of these residues to allow the two histidine residues to be studied independently and to identify the key residue involved in enzyme activity.

7.2 Site directed mutagenesis strategy

A PCR based approach was used to engineer the mutations into the hisIE cDNA. Oligonucleotides were designed containing the desired codon changes to allow the substitution of an alanine residue for both histidine residues. The oligonucleotides used for mutagenesis are given in Table 7.1. The PCR strategy was performed in two steps. The first step was to generate two fragments of the hisIE cDNA each containing the mutation. The second step of the PCR strategy was to join these two fragments together to generate a full length cDNA (Figure 7.1).

7.2.1 First round of PCR

Using the oligonucleotides described in Table 7.1 the first round of PCR was carried out as described in section 2.6.2. Both mutants were made independently at

Table 7.1 Oligonucleotides used for the site directed mutagenesis of the *Arabidopsis* hisIE cDNA

Oligonucleotide	Sequence	Comments
ForH 1 Ala	GAT GGA CCT ACC TGT GCC ACA GGG	Generation of the 3' fragment of the H146A mutant (Histidine 146 to Alanine) along with the BamIE primer
Rev H1 Ala	GTC TCT TCC CCT GTG GCA CAG GTA GG	Generation of the 5' fragment of the H146A mutant (Histidine 146 to Alanine) along with the Nde2 primer
For H2 Ala	GCT GAT GTT TTA TAC GCC GCA ATG GTG	Generation of the 3' fragment of the H245A mutant (Histidine 245 to Alanine) along with the BamIE primer
Rev H2 Ala	GCA CCA TTG CGG CGT ATA AAA CAT CAG	Generation of the 5' fragment of the H245A mutant (Histidine 245 to Alanine) along with the Nde2 primer
Nde2	GTA TTC GCG TGC AAT GAT CAT ATG AAC AAT G	Used for the generation of the 5' end and contains the restriction site to permit cloning into the vector pTB361
BamIE	CAA TGA ACA ACA GGA TCC GTT CCG AG	Used for the generation of the 3' end and contains the restriction site to permit cloning into the vector pTB361

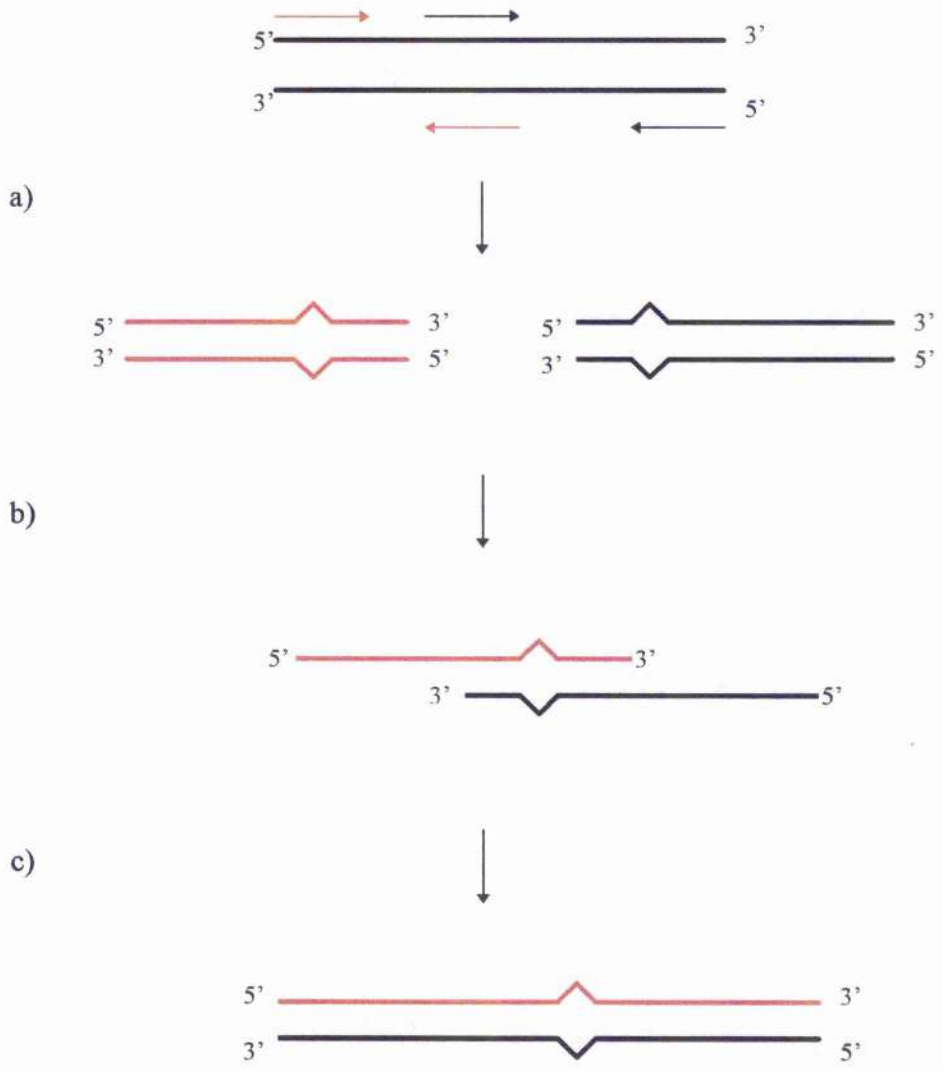


Figure 7.1 Strategy for site directed mutagenesis of hisIE cDNA.

a) First round of PCR generates 2 fragments containing the desired mutation

b) Use of the two smaller PCR fragments as the template for the second round of PCR

c) Second round of PCR generating full length PCR product containing the required mutation

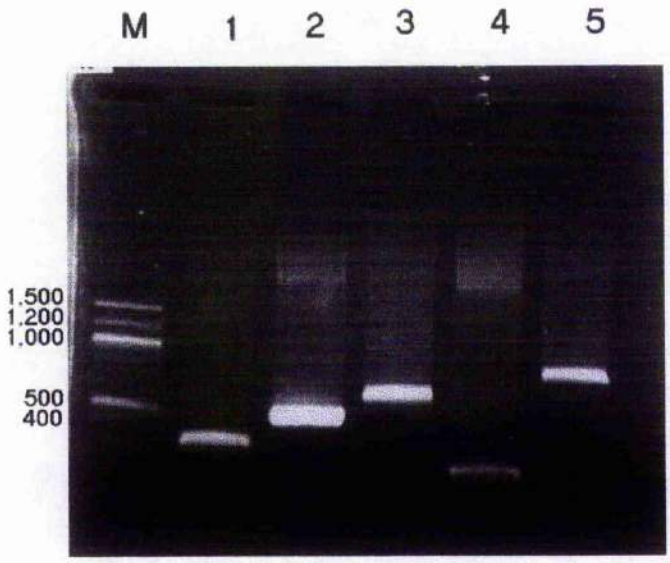


Figure 7.2 Agarose gel showing the analysis of DNA fragments produced from the first round of site directed mutagenesis

Markers (Fragment sizes are given in bp)

Lane 1 H146A 5' fragment

Lane 2 H146A 3' fragment

Lane 3 H245A 5' fragment

Lane 4 H245A 3' fragment

Lane 5 Control (short HisIE construct)

1% agarose gel in TBE buffer containing 0.5μg/ml ethidium bromide.

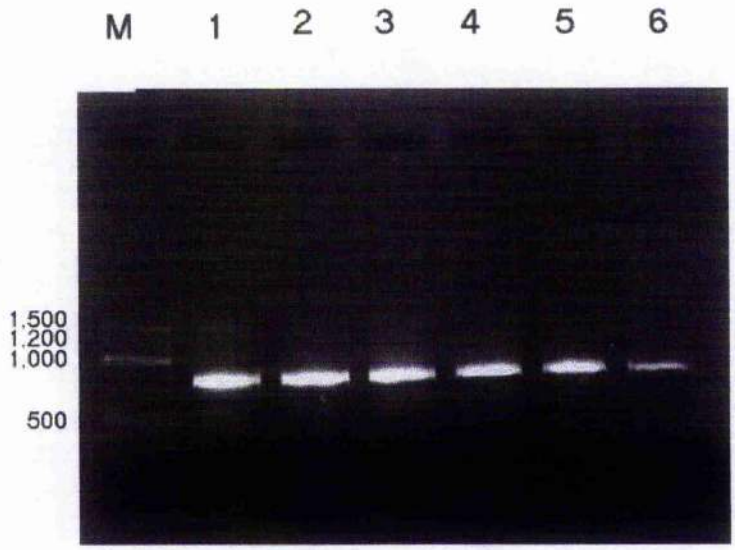


Figure 7.3 Agarose gel showing the fragments of DNA produced during the second round of PCR

Markers (Fragment sizes are given in bp)

Lane 1 H146A 5' and 3' fragments

Lane 2 H146A 5' and 3' fragments

Lane 3 H245A 5' and 3' fragments

Lane 4 H245A 5' and 3' fragments

Lane 5 Control (short HisIE construct)

Lane 6 Control fragment from first round of PCR

1% agaorose gel in TBE buffer containing 0.5g/ml ethidium bromide.

the same time. The products from the PCR reaction were analysed on a 1% agarose gel (containing ethidium bromide). Fragments of the appropriate size were visible when viewed under UV light (Figure 7.2). The two smaller fragments for each mutant were purified and used as the template for the second round of PCR.

7.2.2 Second round of PCR

The second round of PCR was again carried out as described in section 2.6.2 using the two smaller fragments as templates. Again the PCR products were analysed by gel electrophoresis and bands of the appropriate size were observed under UV light (Figure 7.3). Each full length PCR product was purified using the Wizard™ PCR clean kit (sections 2.6.3 and Chapter2, Appendix IV). The purified DNA was then ligated into an expression vector.

7.3 Cloning of the mutant hisIE cDNAs

The mutant cDNAs were cloned using the strategy described in Chapter 3 for the cloning of the wild type cDNA. The fragments of DNA were cloned into the T7 expression vector pTB361 (Brockbank & Barth, 1993). Following transformation of competent cells with the ligation reactions a number of colonies were observed on the plates. Plasmid DNA was prepared from several of these clones and was subjected to restriction analysis to identify the presence of an insert of the appropriate size. Several clones were found to contain an insert of the appropriate size. The data from the cloning experiments is described in Table 7.2. Those clones containing an insert of the appropriate size were sequenced to ensure that the required mutations had been incorporated and to ensure that no other mutations had been incorporated during the PCR mutagenesis.

Clone	Insert	Complements	Expression	Sequence
His1-4	yes	yes	yes	yes- OK
His1-5	yes			
His1-6	yes	no	yes	no
His1-7	yes			
His1-8	yes			
His1-9	yes	yes	yes	no-frameshift
His1-11	yes	yes	yes	no-frameshift
His2-3	yes	yes	yes	no
His2-4	no			
His2-5	yes	yes	yes	no-frameshift
His2-6	yes	yes	yes	
His2-7	no			
His2-8	yes	yes	yes	yes-OK
His2-9	yes			
His2-10	yes			

Table 7.2 hisIE mutant clones

The clones His1-4 and His2-8 were used for all future experiments and from now on will be referred to as H146A and H245A.

7.4 Overexpression studies of HisIE mutants

Those clones containing an insert were transformed into the expression host BL21(DE3)pLysS (Moffatt & Studier, 1987) to determine if they were able to express a mutant protein. The small scale expression analysis was carried out as described in section 2.7.1. Analysis by SDS-PAGE revealed that all of the chosen mutants were able to express a band of protein of the appropriate size when induced with IPTG. Although there was approximately 30% expression of the wild type construct in this host strain, the mutants did not overexpress the mutant protein to the same extent as the wild type. Best estimates are that the protein has been overexpressed to between 15 and 20% of total cell protein. Samples were removed for further analysis by SDS PAGE and for assays to determine if there was HisIE activity present. It was not possible to express either of the mutants in the *E. coli hisIE* mutant strain UTH903 (Goldschmidt *et al;* 1970) since it did not contain a natural T7 RNA polymerase. It was not possible to introduce the T7 gene using a bacteriophage vector as the mutant strain failed to express the maltose binding receptor thus making the strain λ resistant. For this reason and in order to ensure continuity between the mutants and the wild type both mutants were expressed and purified from *E. coli* BL21(DE3)pLysS.

7.5 Purification of the HisIE mutant proteins

Both mutants were purified as described in Chapter 4 using the DEAE Sephacel column followed by a Phenyl Sepharose column and finally by gel filtration on an S-200 column.

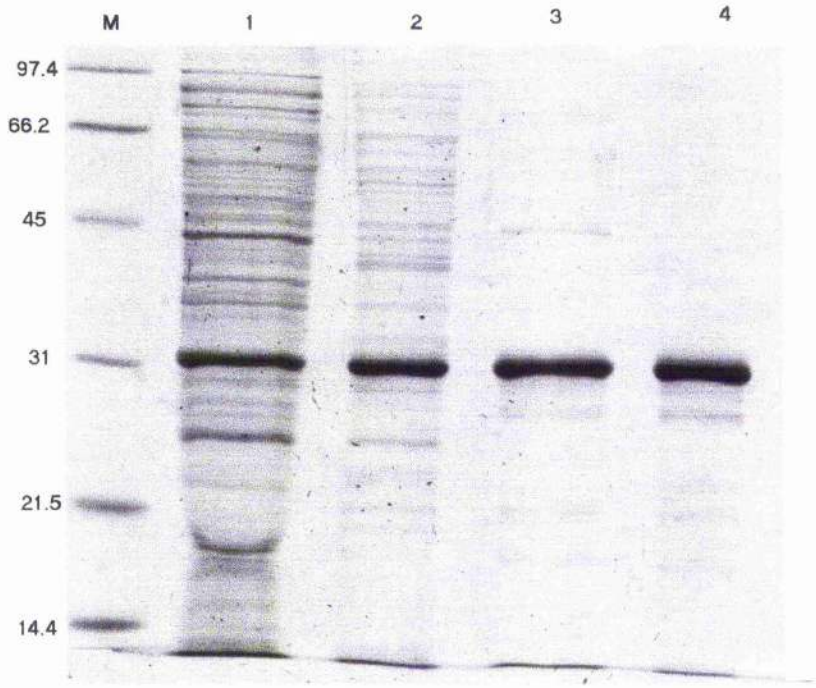


Figure 7.4 SDS Page gel showing the purification of the HisIE H146A mutant protein

Markers (molecular weights are given alongside the gel photograph in kDa)

Lane 1 Crude extract (20μg total protein)

Lane 2 DEAE pool (10μg total protein)

Lane 3 Phenyl Sepharose pool (5μg total protein)

Lane 4 S-200 pool (5μg total protein)

The faint bands present below the main band at 31kDa are believed to be degradation/ proteolytic fragments of the HisIE protein.

A 15% acrylamide gel was run at room temperature and stained with Coomassie blue (section 2.7.3).

Sample	Volume (ml)	Protein (mg/ml)	Protein (mg)	Activity (units/ml)	Total Units	% recovery	Specific activity	Purification
Crude extract	86	38	3268	9.8	842.2	100	0.2578	1
DEAE pool	90	10	900	8.55	770	91.4	0.855	3.31
Phenyl sepharose	82	5	410	4.16	341.8	40.5	0.8336	3.23
S-200	9.8	7.3	71.5	25.5	249.9	29.6	3.49	13.5

Table 7.3 Purification table showing the purification of the *Arabidopsis* HisIE H146A mutant from the overexpression strain BL21(DE3)pLysS pSACIEH146A

Sample	Volume (ml)	Protein (mg/ml)	Protein (mg)	Activity (units/ml)	Total Units	% recovery	Specific activity	Purification
Crude extract	80	45	3600	66.7	5337	100	1.48	1
DEAE pool	112	9	1008	31.3	3505	65.6	3.47	2.34
Phenyl sepharose	100	6.3	630	25.78	2578	48.3	4.09	2.76
S-200	10.8	7.7	83.16	228.7	2470	46.2	29.7	20

Table 7.4 Purification table showing the purification of the *Arabidopsis* HisIE H245A mutant from the overexpression strain BL21(DE3)pLysS pSACIEH245A

7.6 Characterisation of the HisIE mutant proteins

Following purification of the mutant HisIE proteins the K_m and the k_{cat} of each mutant was determined. These are detailed in Table 7.5.

Enzyme	K_m (μM)	k_{cat} (per second)	k_{cat}/K_m ($M^{-1} s^{-1}$)
Wild type	15	1.24	8.3×10^4
H146A mutant	7.69	0.43	5.59×10^4
H245A mutant	15	2.75	1.83×10^5

Table 7.5 Comparison of k_{cat} and K_m of wild type and histidine to alanine mutants of the *Arabidopsis* HisIE protein

7.6.1 Comparison of the H146A mutant with wild type enzyme

The K_m of the mutant enzyme, H146A, is approximately half that of the wild type enzyme (Figure 7.5, Table 7.5). This would indicate that the mutant enzyme has a stronger affinity for the substrate PR-ATP than the wild type enzyme. The mutation of the His146 to Alanine may have removed some sort of steric hindrance that now enables the PR-ATP to fit more tightly into the active site. The k_{cat} of the H146A mutant enzyme is approximately 3 fold lower than the wild type enzyme indicating that the efficiency of catalysis has been impaired. Since the effect on k_{cat} of removing the imidazole side chain of H146 is relatively small it is very unlikely that this residue is directly involved in the catalytic mechanism. More likely H146 may have a role in

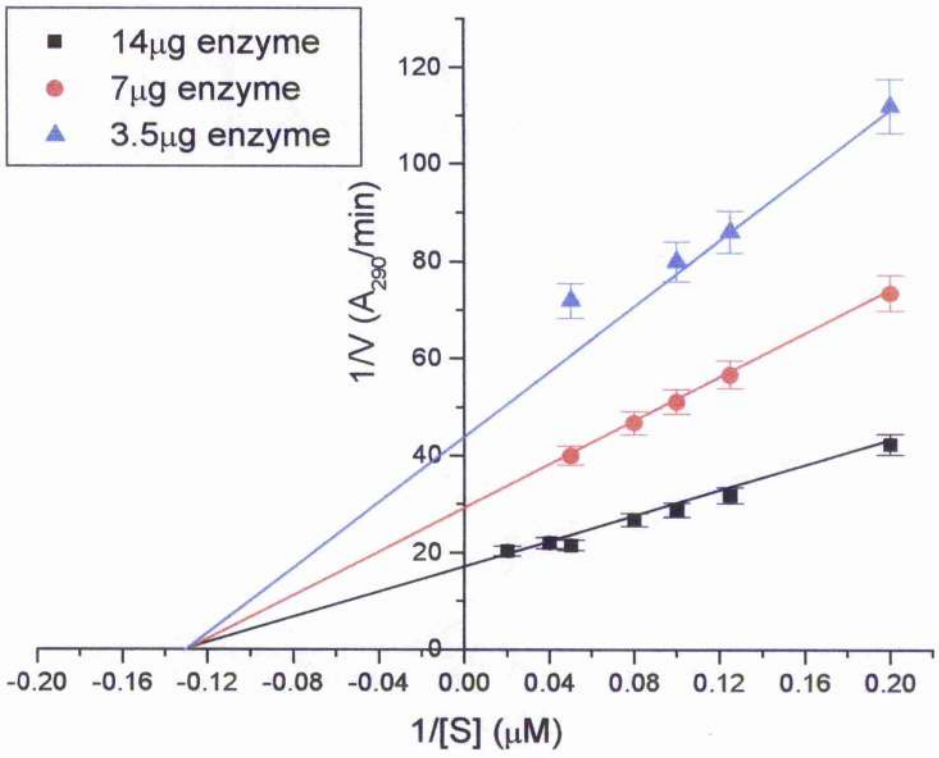


Figure 7.5 Determination of the K_m of the H146A mutant of the HisIE protein

Double reciprocal plot of initial velocity versus [PR-ATP] at three concentrations of enzyme

All measurements relate to the change in absorbance at 290nm. Velocity is plotted as change in absorbance per minute. Substrate concentration in the enzyme assays ranged from 5 to 50μM.

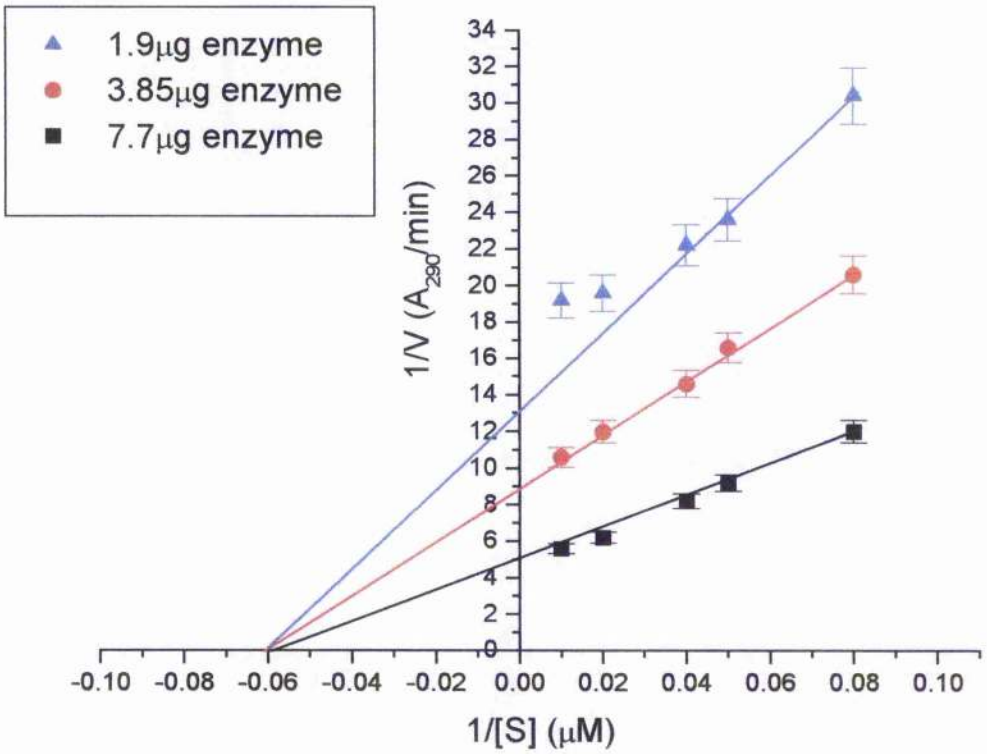


Figure 7.6 Determination of the K_m of the H245A mutant of the HisIE protein

Double reciprocal plot of initial velocity versus [PR-ATP] at three concentrations of enzyme

All measurements relate to the change in absorbance at 290nm. Velocity is plotted as change in absorbance per minute. Substrate concentration in the enzyme assays ranged from 12.5 to 100μM PR-ATP.

maintaining the structural integrity of the active site, for example through the formation of a hydrogen bond.

7.6.2 Comparison of the H245A mutant with wild type enzyme

The K_m of the mutant enzyme, H245A, is the same as the wild type enzyme while the k_{cat} is twice that of the wild type enzyme (Figure 7.6, Table 7.5). Thus the catalytic activity of the HisIE enzyme has been slightly improved by removing His245 and substituting it with an alanine residue. This residue is present within the second domain of the HisIE protein. This domain is responsible for the hydrolysis of the pyrophosphate moiety. H245 is not absolutely conserved, for example in fungal species it is substituted by a phenylalanine residue. These data strongly suggest that H245 is not essential for the catalytic activity of the enzyme.

One possible explanation for the increase in activity of the H245 mutant is that H245 is in some way blocking the active site and by substituting it with a smaller residue such as alanine this steric hindrance is reduced and the substrate can enter the active site more readily. The fact that this residue is clearly conserved across most species indicates that it must have a role of some sort perhaps an ancient function which has now become obsolete.

One theory is that the histidine residue may have a role in controlling access to the active site. Experiments to test if ATP acted in a competitive manner against both mutants indicated that the ATP is a stronger competitive inhibitor for the H245A mutant than for the wild type enzyme. The K_m of the H245A mutant in the presence of 1mM ATP is 38μM compared to 22μM for the wild type enzyme. The H245 may partially block the active site of the enzyme restricting access by the substrate but also reducing access to nucleotides such as ATP. A mechanism such as this could prove to

be important due to the relatively high concentrations of nucleotides within the cell. As the metabolic substrates for the histidine biosynthetic enzymes such as HisIE are relatively scarce the enzyme has to maximise the potential of substrate binding and prevent blocking of the active site with ATP etc. This explanation is also consistent with the chemical modification results. Addition of a DEPC group to the histidine side chain would very substantially block the active site of the enzyme. This steric effect, rather than the modification of the catalytic properties of the histidine side chain, may account for the inactivation observed with DEPC.

7.7 Determination of the molecular weight of the H146A and H245A mutants under native and high salt conditions

The molecular weight of both mutants was determined under low and high salt conditions using the S-200 gel filtration column. The H245A mutant was present as a dimer under both low and high salt conditions with a calculated molecular weight of 63kDa. The H146A was present as a dimer under native low salt conditions and as a monomer under high salt conditions suggesting that H146 is involved in the interaction of the two subunits.

7.8 Mass spectroscopy analysis of HisIE mutant proteins

Electrospray mass spectrometry of the purified H146A and H245A proteins was carried out by Dr K. Lilley, University of Leicester. Both mutant enzymes had molecular weights lower than the wild type enzyme (see p101). The predicted molecular weight was 26,792 and the observed molecular weights were 26,814 (H146A) and 26815 (H245A). As with the measurements on the wild type enzyme

there were difficulties with the instrument calibration and further experiments will be necessary to confirm the molecular weight.

Chapter 8

Discussion

8.1 Discussion of Results

The cDNA encoding the bifunctional second enzyme of histidine biosynthesis was the only cDNA which was cloned by complementation of the *E. coli* auxotrophic mutants (Chapter 3). Although complementation of the *hisG* and *hisA* mutants was observed on the first round of screening, subsequent investigation of the isolated plasmid DNA revealed that the insert in these clones contained the identical sequence to that isolated for the hisIE cDNA. One possible explanation is that since the histidine auxotrophic mutants have not been fully characterised, the possibility remains that both the *hisG* and *hisA* mutants are *hisIE* mutants. This may be the reason why complementation was so readily achieved with the hisIE cDNA. It was only possible to show that the *E. coli* mutants were auxotrophic for histidine by growing on minimal media in the presence or absence of histidine. If the other intermediates in the pathway had been available it may have been possible to confirm the auxotrophy for each of the mutants by growing in the presence or absence of these compounds. Once the other intermediates of the pathway have been isolated it would be sensible to characterise the auxotrophy of each *E. coli* mutant. The cDNA for the third enzyme of the pathway was isolated by complementation of the *hisA* mutant by Fujimori *et al.*, (1998) which suggests that this was a valid method for isolation and cloning of cDNA's from *Arabidopsis* and that the *hisA* mutant was auxotrophic for the correct gene.

Analysis of the hisIE cDNA from this work (Chapter 3) as well as the work of Fujimori and Ohta, (1998[a]) suggests that the HisIE protein is a bifunctional protein, containing two separate domains which are responsible for catalysing the second and third steps of histidine biosynthesis. The presence of an N-terminal chloroplastic targeting sequence indicates that the pathway takes place within the chloroplast. This is in agreement with studies on other histidine biosynthetic enzymes isolated from plants

which suggest that the enzymes are also targeted to the chloroplast (Tada *et al.*, 1994, Nagai *et al.*, 1992[b], El Maliki *et al.*, 1998).

An overexpression construct expressing a truncated HisIE protein from which the chloroplast targeting sequence had been removed was generated and permitted the overexpression of active HisIE protein to approximately 30% total cell protein. This provided large quantities of material for the purification of homogeneous HisIE protein by a novel method (Chapter 4). Homogeneous protein was obtained using three chromatographic steps (ion exchange, hydrophobic interaction and gel filtration chromatography).

In order to characterise the bifunctional plant enzyme, a key part of this project was the generation of the enzyme substrate PR-ATP (Chapter 5). The chemical nature of the PR-ATP molecule provided a challenge for the generation and purification of material suitable for kinetic analysis. The molecule is intrinsically unstable under certain conditions. Extreme conditions such as temperature or pH cause the molecule to breakdown to phosphoribose and ATP. This presented a significant problem when attempting to isolate pure material for kinetic analysis. Establishing the correct buffer conditions for chromatography and freeze drying was essential for keeping the procedure for generating PR-ATP as short as possible in order to minimise any breakdown. A simplified protocol for the generation and purification of PR-ATP has been developed. Significant quantities of PR-ATP can be purified using an ion exchange step followed by freeze drying to remove the residual buffer leaving behind pure PR-ATP.

Initial characterisation of the HisIE enzyme revealed a K_m of 15μM and a k_{cat} of 1.244 per second. Previous experiments on the histidinol dehydrogenase from cabbage determined a K_m of 11μM for histidinol (Nagai et al 1992[a]). These K_m values

are typical of low abundance biosynthetic enzymes. The substrates for biosynthetic enzymes are likely to be present in tiny amounts within the cell and therefore the enzymes need to be highly efficient at binding any available substrate. ATP acts as a weak inhibitor of the HisIE enzyme which is surprising, considering that ATP forms a large part of the substrate molecule. The K_i for ATP was calculated as 1.65mM. The ratio of ATP to PR-ATP within the plant cell will be very high and it would be logical to suggest that the enzyme must have a way of reducing the inhibitory effect of ATP in order that it can bind any available PR-ATP. Inactivation experiments with DEPC suggested that a histidine residue was important for enzyme activity. Enzyme activity was lost very rapidly in the presence of micromolar concentrations of DEPC. Analysis of the change in absorbance at 240nm suggests that one DEPC molecule is incorporated per molecule of protein although it is not possible to determine which residue is modified. The presence of relatively high concentrations of PR-ATP was able to provide some protection against inactivation which suggested that the modification was affecting the active site or the surrounding area (binding pocket). Site directed mutagenesis of the two highly conserved histidine residues present within the HisIE protein was carried out and the mutant proteins overexpressed and purified in the same manner as for the wild type enzyme (Chapter 7). Characterisation of the two mutant proteins and comparison of the K_m and k_{cat} values (Table 7.5) would suggest that neither of the histidine residues are catalytic. There is only a two or three fold difference in the values for K_m and k_{cat} for the wild type and mutant proteins. If a key catalytic residue had been affected we would expect a difference in k_{cat} of greater than a thousand fold. Clearly these residues are important to the activity of the HisIE protein due to their high degree of conservation across a number of species as well as the effect they have on enzyme activity when they are substituted by another residue.

The histidine residues may have role in maintaining the structural integrity of the protein molecule by forming a key interaction such as a salt bridge or hydrogen bond. Evidence to substantiate this could be the lack of enzyme activity observed below pH 7 (section 6.2.1 and Figure 6.2) which may be the result of a histidine residue becoming protonated under acidic conditions thus resulting in the loss of hydrogen bonding potential and the destruction of a key hydrogen bond.

The HislE mutant H245A has a k_{cat} which is two fold higher than the wild type. This would indicate that the enzyme is more active and one reason for this is that by substituting histidine with alanine, we have removed steric hindrance around the active site allowing the substrate greater access. This data correlates with the DEPC modification experiments in which the rapid loss of enzyme activity may be the result of DEPC modifying a histidine residue close to the active site and merely obstructing the active site rather than affecting a catalytic residue. Experiments to determine if ATP was acting in a competitive manner against both histidine mutant proteins, indicated that ATP is a stronger inhibitor of the H245A mutant compared to the wild type enzyme. The presence of H245 at the entrance to the active site may be a mechanism for preventing ATP and other nucleotide molecules from binding to the active site. The HislE enzyme has to maximise the binding of the relatively scarce substrate PR-ATP and prevent blocking of the active site with ATP etc. This would be a very important mechanism due to the high concentration of ATP within the chloroplast.

The molecular weights of the wild type and mutant proteins were determined by gel filtration chromatography on an S-200 column. The wild type HislE protein appears to be a dimer under both native and high salt conditions, suggesting that there is a strong interaction between the two subunits. This correlates with work by

D'Ordine *et al.*, (1999) on the monofunctional HisI protein from *M. vannielli* which indicates that this protein is also a dimer. The H245 mutant HisIE protein is also a dimer under native and high salt conditions. The H146 mutant protein however is a dimer under native gel filtration conditions but is present as a monomer in the presence of high salt. This would suggest that H146 is involved in the interaction of the two subunits.

8.2 Further work on the *Arabidopsis* HisIE enzyme

Following initial characterisation of the enzyme, there are a number of other experiments which are essential to furthering the knowledge of this enzyme. The chemical modification experiments with TNBS appear to indicate that a lysine or a cysteine residue may be important for the activity of the enzyme, most probably at a structural level rather than a mechanistic one. There are 15 lysines and 5 cysteine residues within the HisIE protein, with 3 cysteines and 3 lysines conserved across all species. Further experiments to identify which residue is modified would be a good idea before starting site directed mutagenesis in order to limit the number of mutants to be made and characterised. Identification of active site residues will be important, in order to determine the mechanism of the enzyme.

As the enzyme is bifunctional, this may pose a few problems in determining the mechanism of the enzyme. The first step will be to determine which reaction occurs first; removal of the pyrophosphate group or the hydrolysis of the purine ring. It may be possible to use rapid reaction techniques to study the enzyme but it is likely that the generation of protein that has either activity completely knocked out, will be required for this process. This requires identification of key residues involved in the mechanism which could then be mutated or expression of the two individual protein domains.

A key step towards identifying the important catalytic residues is the determination of the crystal structure of the enzyme. Small crystals of the HisIE enzyme have been obtained by Mr B. Lohkamp, Division of Biochemistry and Molecular Biology, Glasgow University. Further work will be required in order to obtain larger crystals that can be used for X-ray diffraction and data collection. A crystal structure would allow identification of key residues which are likely to be involved in substrate binding as well as those residues that are likely to be involved in catalysis. This will aid the full characterisation of the enzyme and ultimately determination of the mechanism.

By using structure based drug design, it may be possible to synthesise potential inhibitors of this enzyme which may act as herbicides/ antibiotics.

8.3 Future Research

A key area of importance for future research will be to understand how the expression of the genes for the histidine biosynthetic enzymes is regulated. In bacterial systems, the genes encoding the enzymes are present in an operon which can be regulated by a number of different factors including the concentration of histidyl tRNA synthetase. In yeast and other fungal systems, the genes for the enzymes of histidine biosynthesis are known to be regulated by a process called general control. This general control is brought about by a number of different proteins, the most well characterised being the GCN4 protein. This is a transcriptional activator (Hinnesbusch, 1997) and a member of a family of DNA binding proteins that contain the b-ZIP (basic region leucine zipper) DNA binding motif (Vinson *et al.*, 1989). This protein binds to a common hexanucleotide motif present in one or more copies in the 5' non-coding region of all genes subject to general control (Vogt *et al.*, 1987, Bohmann *et al.*,

1987). In *S. cerevisiae* in response to amino acid starvation, the de-repression of the GCN4 protein leads to the increased expression of more than 30 different genes in at least 10 different pathways. These elements are found upstream of a number of different genes involved in several different amino acid biosynthetic pathways. Although the majority of work has been carried out in *S. cerevisiae,* a number of other fungi appear to show a GCN4 type response. *Neurospora crassa* (Paluk *et al.*, 1988) and *Candida albicans* (Pereira and Livi, 1995) are two other organisms which experience general control. The CPC-1 protein from *Neurospora crassa* has been identified as a homologue of GCN4 (Paluk *et al.*, 1988) and has been shown to bind to the same DNA sequence as the GCN4 protein from *S. cerevisiae* (Ebbole, 1991). There is evidence of a GCN4 homologue in *Aspergillus nidulans* which may be involved in the process of general control (Piotrowska, 1980).

In plants there is gathering evidence for the presence of a general control mechanism analogous to that in yeast. In *Arabidopsis thaliana,* blocking histidine biosynthesis by inhibition of IGPD resulted in the increased expression of eight genes involved in the biosynthesis of histidine, lysine and purines (Guyer *et al.*, 1995). Addition of histidine terminated the gene regulating effects of the inhibitor, demonstrating that the changes in gene expression resulted from the inhibition of histidine biosynthesis. It provides some evidence that plants are capable of cross pathway regulation/ general control (Guyer *et al.*, 1995).

Investigation of the published sequence data in the *Arabidopsis* data-base has allowed the identification of the chromosomes on which the histidine biosynthetic genes are found. The *hisG* gene has been identified on chromosome 1, *hisA* can be mapped to chromosome 2, *hisHF* to chromosome 4, *hisB* to chromosomes 3 and 4 (evidence from Tada *et al.*, 1994 indicates that there are two copies of this gene in the

Arabidopsis genome) and *hisD* can be mapped to chromosome 5. Despite recent searches of the genome sequence data, it has not been possible to assign the *hisIE* and *hisC* genes to a specific chromosome. It is not clear if the *hisC* gene is present within the *Arabidopsis* genome although it has been isolated from *Nicotiana tabacum* (El Malki *et al.*, 1998) and therefore one would assume that it will be present.

Clearly, as these genes are scattered throughout the genome of *Arabidopsis*, a mechanism for the co-ordinated expression of these genes would be required. A GCN4 type element has been identified upstream of the *Arabidopsis hisIE* gene (Fujimori and Ohta., 1998[a]) and the *Arabidopsis* aspartate kinase homoserine dehydrogenase (Ghislain *et al.*, 1994). From analysis of the region of chromosome 1 encoding the *hisG* gene, a GCN4 element can be observed upstream of the *hisG* gene. It has also been observed upstream of the *hisB* gene from *Arabidopsis*.

An *Arabidopsis* nucleoside diphosphate kinase (NDPK Ia) has recently been isolated by complementation of a yeast *gcn4* mutant with an *Arabidopsis* cDNA library (Zimmermann *et al.*, 1999). The NDPK Ia is responsive to UV light and induces genes involved in histidine biosynthesis in a similar manner to the natural GCN4 protein (Zimmermann *et al.*, 1999). Although NDPK Ia acts as a transcriptional activator in yeast it not known which plant genes are the target of NDPK Ia. It is possible that this protein may target the genes for amino acid biosynthesis, particularly the genes for histidine biosynthesis. Expression of NDPK Ia in *Arabidopsis* seedlings is strongly expressed by UV irradiation further suggesting that there is strong evidence for analogous systems in yeast and plants (Zimmermann *et al.*, 1999).

There is a great deal of interest in understanding more about the amino acid biosynthetic pathways as interesting targets for herbicides and antimicrobials. For this reason, ongoing work to isolate and characterise the enzymes of these pathways is

essential. Understanding how these pathways are regulated may provide an insight into how plants can be manipulated for various purposes. The use of plants, in which we can express higher concentrations of histidine, which are tolerant to heavy metals could be used as a method of decontaminating toxic waste land. Plants designed to be tolerant to heavy metals would provide an inexpensive method of decontaminating waste land (Salt *et al.*, 1998).

The improvement of the nutritional quality of food for both human and animal consumption would be a significant step forward to improving the quality of life especially in the third world. Many of the basic crop plants which are used as the staple food sources in many diets are relatively poor in some of the more essential amino acids. If it were possible to manipulate the levels of amino acids produced in plants it could reduce the need for supplementation of the diet with expensive animal produce. This also has application for farming in the first world as it would provide an alternative to substituting herbivore feed with animal material.

Chapter 9

References

9.1 References

Adams E. (1954) The enzymatic synthesis of histidine from histidinol. *J.Biol.Chem.* **209**, 829-846.

Albritton W.L. and Levin A.P. (1970) Some comparative kinetic data on the enzyme imidazoleacetolphosphate aminotransferase derived from mutant strains of *Salmonella typhimurium*. *J.Biol.Chem.* **245**, 2525-2528.

Alifano P., Fani R., Lio P., Lazcano A., Bazzicalupo M., Carlomagno M.S. and Bruni C.B. (1996) Histidine Biosynthetic Pathway and Genes: Structure, Regulation and Evolution. *Microbiological Reviews* **60**, 4-69.

Ames B.N. and Mitchell H.K. (1955) The biosynthesis of histidine: imidazoleglycerol phosphate, imidazoleacetol phosphate, and histidinol phosphate. *J.Biol.Chem.* **212**, 687-697.

Ames B.N., Martin R.G. and Garry B.J. (1961) The first step of histidine biosynthesis. *J.Biol.Chem.* **236**, 2019-2026.

Ames B.N., Tsang T.H., Buck M. and Christman M.F. (1983) The leader mRNA of the histidine attenuator region resembles $tRNA^{His}$: Possible general regulatory implications. *Proc.Natl.Acad.Sci.USA.* **80**, 5240-5242.

Beckler G.S. and Reeve J.N. (1986) Conservation of primary structure in the *hisI* gene of the archaebacterium, *Methanococcus vannielli*, the eubacterium *Escherichia coli* and the eukaryote *Saccharomyces cerevisiae*. *Mol.Gen.Genet.* **204**, 133-140.

Bell R.M., Parsons S.M., Dubravac S.A., Redfield A.G. and Koshland Jr.D.E. (1974) Characterisation of Slowly Interconvertible States of Phosphoribosyladenosine Triphosphate Synthetase Dependent on Temperature, Substrates, and Histidine. *J.Biol.Chem.* **249**, 4110-4118.

Bohmann D., Bos T.J., Admon A., Nishimura T., Vogt P.K. and Tijian R. (1987) Human proto-oncogene *c-jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* **238**, 1386-1392.

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

Brenner M. and Ames B.N. (1971) The histidine operon and its regulation. In: Vogel H.J., (Ed.) *Metabolic Pathways*, pp. 349-387. New York: Academic Press Inc

Brockbank S.M. and Barth P. (1993) Cloning, sequencing and expression of the DNA gyrase genes from *Staphylococcus aureus*. *J.Bacteriol.* **175**, 3269-3277.

Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A. and Gocayne J.D. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058-1073.

Bürger E. and Görish H. (1981) Patterns of product inhibition of a bifunctional dehydrogenase, L-Histidinol:NAD^+ oxidoreductase. *Eur.J.Biochem.* **116**, 137-142.

Carlomagno M.S., Chiarotti L., Alifano P., Nappo A.G. and Bruni C.B. (1988) Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. *J.Mol.Biol* **203**, 585-606.

Chamberlin M. and Ring J. (1973) Characterisation of T7 specific ribonucleic acid polymerase I general properties of the enzymatic reaction and the template specificity of the enzyme. *J.Biol.Chem.* **248**, 2235-2244.

Chiarotti L., Nappo A.G., Carlomagno M.S. and Bruni C.B. (1986) Gene structure in the histidine operon of *Eschericia coli;* identification and nucleotide sequence of the *hisB* gene. *Mol.Gen.Genet.* **202**, 42-47.

Chong S., Shao Y., Paulus H., Benner J., Perler F.B. and Xu M.-Q (1996) Protein splicing involving the *Saccharomyces cerevisiae* VMA intein: the steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J.Biol.Chem.* **271**, 22159-22168.

Chong S., Mersha F.B., Comb D.G., Scott M.E., Landry D., Vence L.M., Perler F.B., Benner J., Kucera R.B., Hirovonen C.A., Pelletier J.J., Paulus H. and Xu M.-Q (1997) Single column purification of free recombinant proteins using a self cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 277-281.

Clark A.J., Marguilies A.D. and Chamberlin M. (1969) Production of a merodiploid strain from a double male strain of *Escherichia coli* K12. *Mol.Gen.Genet.* **105**, 1-15.

Cole D.J. (1994) Molecular Mechanisms to confer herbicide resistance. In: Marshall G. and Walters D., (Eds.) *Molecular Biology in Crop Protection*, pp. 146-177. London: Chapman and Hall

Cox J.M., Hawkes T.R., Bellini P., Ellis R.M., Barrett R., Swanborough J.J., Russell S.E., Walker P.A., Barnes N.J., Knee A.J., Lewis T. and Davies P.R. (1997) The Design and synthesis of Inhibitors of Imidazoleglycerol Phosphate Dehydratase as potential herbicides. *Pestic.Sci.* **50**, 297-311.

Crane D.I. and Gould S.J. (1994) The *Pichia pastoris HIS4* gene: nucleotide sequence, creation of a non-reverting *his4* deletion mutant, and development of *HIS4*-based replicating and integrating plasmids. *Curr.Genet.* **26**, 443-450.

D'Alessio J.M., Bebee R., Hartley J.L., Noon M.C. and Polayes D. (1992) Automatic subcloning of cDNA. *Focus* **14**, 76-79.

D'Ordine R.L., Klem T.J. and Davisson V.J. (1999) N1-(5'-phosphoribosyl)adenosine-5'-monophosphate cyclohydrolase: purification and characterisation of a unique metalloenzyme. *Biochemistry* **38**, 1537-1546.

Dall-Larson T. and Klungsoyr I. (1976) The binding of specific ligands to adenosine triphosphate phosphoribosyl transferase. *Eur.J.Biochem.* **69**, 195-201.

Dall-Larson T. (1988) Regulation of the first step of Histidine biosynthesis in *Escherichia coli*. *Int.J.Biochem.* **20**, 231-235.

Davisson V.J., Deras I.L., Hamilton S.E. and Moore L.L. (1994) A plasmid based approach for the synthesis of a histidine biosynthetic intermediate. *J.Org.Chem* **59**, 137-143.

Deka R., 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.

Delorme C., Ehrlich S.D. and Renault P. (1992) Histidine biosynthesis in *Lactococcus lactis* subsp *lactis*. *J.Bacteriol.* **174**, 6571-6579.

Deutscher M.P. (1990) *Guide to protein purification*, Methods in Enzymology, Vol 182, Academic press.

Donahue T.F., Farabaugh P.J. and Fink G.R. (1982) The nucleotide sequence of the *HIS4* region of yeast. *Gene* **18**, 47-59.

Ebbole D.J., Paluk J.L., Plamann M., Sachs M.S. and Yanofsky C. (1991) *cpc-1*, the General Regulatory Gene for Genes of Amino Acid Biosynthesis in *Neurospora crassa*, is differentially expressed during the asexual life cycle. *Molecular and Cellular Biology* **11**, 928-934.

El Malki F., Frankard V. and Jacobs M. (1998) Molecular cloning and expression of a cDNA sequence encoding histidinol phosphate aminotransferase from *Nicotiana tabacum*. *Plant Molecular Biology* **37**, 1013-1022.

Elledge S.J., Mulligan J.T., Ramer S.W., Spottswood M. and Davies R.W. (1991) λYES: A multifunctional cDNA expression vector for the isolation of genes by

complementation of yeast and *Escherichia coli* mutations. *Proc.Natl.Acad.Sci.USA.* **88**, 1731-1735.

Fani R., Alifano P., Allotta G., Bazzicalupo M., Carlomagno M.S., Gallori E., Rivellini F. and Polsinelli M. (1993) The histidine operon of *Azospirillum brasilense*: organisation, nucleotide sequence and functional analysis. *Res.Microbiol.* **144**, 187-200.

Fani R., Lio P., Chiarelli I. and Bazzicalupo M. (1994) The evolution of the histidine biosynthetic genes in prokaryotes: a common ancestor for the *hisA* and *hisF* genes. *J.Mol.Evol.* **38**, 489-495.

Fani R., Lio P. and Lazcano A. (1995) Molecular Evolution of the Histidine Biosynthetic Pathway. *J.Mol.Evol.* **41**, 760-774.

Fani R., Tamburini E., Mori E., Lazcano A., Lio P., Barberio C., Casalone E., Cavalieri D., Perito B. and Polsinelli M. (1997) Paralogous histidine biosynthetic genes: evolutionary analysis of the *Saccharomyces cerevisiae HIS6* and *HIS7* genes. *Gene* **197**, 9-17.

Filho S., Chaumont F., Leterme S. and Boutry M. (1996) Mitochondrial and chloroplast targeting sequences in tandem modify protein import specificity in plant organelles. *Plant Molecular Biology* **30**, 769-780.

Fujimori K. and Ohta D. (1998[a]) Isolation and characterisation of a Histidine biosynthetic gene in *Arabidopsis* encoding a polypeptide with two separate domains for Phosphoribosyl-ATP pyrophosphohydrolase and phosphoribosyl-AMP cyclohydrolase. *Plant Physiol.* **118**, 275-283.

Fujimori K. and Ohta D. (1998[b]) An *Arabidopsis* cDNA encoding a bifunctional glutamine amidotransferase/cyclase suppresses the histidine auxotrophy of a *Saccharomyces cerevisiae IIIS7* mutant. *FEBS Letters* **428**, 229-234.

Fujimori K., Tada S., Kanai S. and Ohta D. (1998) Molecular cloning and characterisation of the gene encoding N'[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase from *Arabidopsis thaliana*. *Mol.Gen.Genet.* **259**, 216-223.

Ghislain M., Frankard V., Vandenbossch D., Matthews B.F. and Jacobs M. (1994) Molecular analysis of the aspartate kinase homoserine dehydrogenase gene from *Arabidopsis thaliana*. *Plant Molecular Biology* **24**, 835-851.

Goldschmidt E.P., Cater M.S., Matney T.S., Butler M.A. and Greene A. (1970) Genetic analysis of the histidine operon in *Escherichia coli* K12. *Genetics* **66**, 219-229.

Golomb M. and Chamberlin M. (1974) Characterisation of T7 specific ribonuclic acid polymerase IV. Resolution of the major *in vitro* transcripts by gel electrophoresis. *J.Biol.Chem.* **249**, 2858-2863.

Görish H. and Hölke W. (1985) Binding of histidinal to histidinol dehydrogenase. *Eur.J.Biochem.* **150**, 305-308.

Gross A., Abril O., Lewis J.M., Geresh S. and Whitesides G.M. (1983) Practical synthesis of 5-Phospho-D-ribosyl α-1-Pyrophosphate (PRPP): Enzymatic routes from Ribose-5-Phosphate or Ribose. *J.Am.Chem.Soc.* **105**, 7428-7435.

Guyer D., Patton D. and Ward E. (1995) Evidence for cross pathway regulation of metabolic gene expression in plants. *Proc.Natl.Acad.Sci.USA.* **92**, 4997-5000.

Hanahan D. (1985) *DNA Cloning*, Oxford: IRL Press Ltd.

Hawkes T.R., Cox J.M., Barnes N.J. and *et al.*, (1993) Imidazole glycerol phosphate dehydratase: a herbicidal target. In: *Proceedings of the Brighton Crop Protection Conference- Weeds*, pp. 739-744. Farnham, UK: British Crop Protection Council

Hilton J.L., Kearney P.C. and Ames B.N. (1965) Mode of action of the herbicide, 3-amino-1,2,4-triazole (amitrole): inhibition of an enzyme of histidine biosynthesis. *Arch.Biochem.Biophys.* **112**, 544

Hinnesbusch A.G. (1988) Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. *Microbiological Reviews* **52**, 248-373.

Hinnesbusch A.G. (1997) Translational regulation of yeast GCN4. *J.Biol.Chem.* **272**, 21661-21664.

Holmes D.S. and Quigley M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Anal.Biochem.* **114**, 193

Huynh T., Young R.A. and Davies R. (1985) *DNA Cloning, A Practical Approach* (Glover D.M. Ed.) Oxford: IRL Press Ltd.

Kane P.M., Yamashiro C.T., Wolczyk D.F., Neff N., Goebl M. and Stevens T.H. (1990) Protein splicing converts the yeast TFP1 gene product to the 69kD subunit of the vacuolar H^+-adenosine triphosphatase. *Science* **250**, 651-657.

Kirschner K. and Bisswanger H. (1976) Multifunctional proteins. *Ann.Rev.Biochem.* **45**, 143-166.

Kishore G.N. and Shah D.M. (1988) Amino acid biosynthesis inhibitors as herbicides. *Ann.Rev.Biochem.* **57**, 627-663.

Klem T.J. and Davisson V.J. (1993) Imidazole Glycerol Phosphate Synthase: The Glutamine Amidotransferase in Histidine Biosynthesis. *Biochemistry* **32**, 5177-5186.

Kleschick W.A. and Gerwick B.C. (1989) Copping L.G., Dalziel J. and Dodge A.D., (Eds.) *Prospects for Amino Acid Biosynthesis Inhibitors in Crop Protection and Pharmaceutical Chemistry*, pp. 139 British Crop Protection Council Monograph

Klungsoyr I. and Kryvi H. (1971) Sedimentation behaviour of phosphoribosyl adenosine triphosphate synthetase: effect of substrate and modifiers. *Biochim.Biophys.Acta* **227**, 327-336.

Kramer U., Cotter-Howells J.D., Charnock J.M., Baker A.J.M. and Smith J.A.C. (1996) Free histidine as a metal chelator in plants that accumulate nickel. *Nature* **379**, 635-638.

Krell T., Chackrewarthy S., Pitt A.R., Elwell A. and Coggins J.R. (1998) Chemical modification monitored by electrospray mass spectrometry: a rapid and simple method for identifying and studying functional residues in enzymes. *J.Peptide Res.* **51**, 201-209.

Kuenzler M., Balmelli T., Egli C.M., Paravicini G. and Braus G.H. (1993) Cloning, primary structure and regulation of the *HIS7* gene encoding a bifunctional glutamine amidotransferase: cyclase from *Saccharomyces cerevisiae*. *J.Bacteriol.* **175**, 5548-5558.

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

Landick R., Turnbourgh J.C.L. and Yanofsky C. (1996) Transcription Attenuation. In: Neidhart F.C., Ingraham J.L., Low K.B., Magasanik B., Schaechter M. and Umbarger H.E., (Eds.) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 1263-1286. Washington D.C.: Am. Soc. Microbiol.

LaRossa R.A. and Falco S.C. (1984) Amino acid biosynthetic enzymes as targets of herbicide action. *Trends in Biotechnology* **2**, 158-161.

LaRossa R.A. and Schloss J.V. (1984) The Sulphonylurea herbicide Sulfometuron Methyl is an extremely potent and selective inhibitor of acetolactate synthase in *Salmonella typhimurium*. *J.Biol.Chem.* **259**, 8753-8757.

Lazcano A., Fox G.E. and Oto J. (1992) Life before DNA: the origin and evolution of early Archean cell. In: Mortlock R.P., (Ed.) *The evolution of metabolic function*, pp. 237-339. CRC Press

Le Coq D., Fillinger S. and Aymerich S. (1999) Histidinol Phosphate Phosphatase, Catalysing the Penultimate Step of the Histidine Biosynthesis Pathway, is encoded by *ytvP* (*hisJ*) in *Bacillus subtilis*. *J.Bacteriol.* **181**, 3277-3280.

Legerton T.L. and Yanofsky C. (1985) Cloning and characterisation of the multifunctional *his-3* gene of *Neurospora crassa*. *Gene* **39**, 129-140.

Lin J., Jesse J. and Bloom F.R. (1992) Rapid isolation of genes from cDNA libraries. *Focus* **14**, 98

Low B. (1968) Formation of Merodiploids in matings with a class of REC⁻ recipient strains of *Escherichia coli* K12. *Proc.Natl.Acad.Sci.USA.* **60**, 160-167.

Lowry O.H., Rosebrough N.J., Farr A.L. and Randall R.J. (1951) Protein measurement with the Folin phenol reagent. *J.Biol.Chem.* **193**, 265-275.

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

Malone R.F., Kim S., Bullard S.A., Lundquist S., Hutchings-Crow L., Cramton S., Lutfiyya L. and Lee J. (1994) Analysis of a recombination hotspot for gene conversion occurring at the *HIS2* gene of *Saccharomyces cerevisiae*. *Genetics* **137**, 5-18.

Mano J., Hatano M., Koizumi S., Tada S., Hashimoto M. and Scheidegger A. (1993) Purification and properties of a monofunctional imidazoleglycerol phosphate dehydratase from wheat. *Plant Physiol.* **103**, 733-739.

Margolis M.N. and Goldberger R.F. (1966) Isolation of the fourth enzyme (isomerase) of histidine biosynthesis from *Salmonella typhimurium*. *J.Biol.Chem.* **241**, 3262-3269.

Margolis M.N. and Goldberger R.F. (1967) Physical and chemical characterisation of the isomerase of histidine biosynthesis in *Salmonella typhimurium*. *J.Biol.Chem.* **242**, 256-264.

Martin R.G. (1963) The First Enzyme in Histidine Biosynthesis: The nature of Feedback Inhibition by Histidine. *J.Biol.Chem.* **238**, 257-268.

Martin R.G. (1970) Imidazoleacetol-phosphate aminotransferase: mechanism of action. *Arch.Biochem.Biophys.* **138**, 239-244.

Martin R.G., Berberich M.A., Ames B.N., Davies W.W. , Goldberger R.F. and Yourno J.D. (1971) Enzymes and Intermediates of Histidine Biosynthesis in *Salmonella typhimurium*. *Methods Enzymol.* **17B**, 3-44.

Matney T.S., Goldschmidt E.P., Erwin N.S. and Scroggs R.A. (1964) A preliminary map of genomic cloning sites for F-attachment in *Escherichia coli* K12. *Biochem.Biophys.Res.Commun.* **17**, 278-281.

Merck Index (1976) *An encyclopedia of Chemicals and Drugs*, ninth edn. Rahway N.J.: Merck and Co. Inc.

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

Millay R.H. and Houston L.L. (1973) Purification and properties of yeast histidinol phosphate phosphatase. *Biochemistry* **12**, 2591-2596.

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

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

Mori I., Fonne-Pfister R., Matsunaga S., Tada S., Kimura Y., Iwasaki G., Mano J., Hatano M., Nakano T., Koizumi S., Scheidegger A., Hayakawa K. and Ohta D (1995) A Novel Class of Herbicides: specific inhibitors of imidazoleglycerol phosphate dehydratase. *Plant Physiol.* **107**, 719-723.

Mousdale D.M. and Coggins J.R. (1991) Amino acid synthesis. In: Kirkwood R.C., (Ed.) *Target sites for herbicidal action*, pp. 29-56. New York: Plenum Publishing Corporation

Nagai A., Suzuki K., Ward E., Moyer M., Hashimoto M., Mano J., Ohta D and Scheidegger A. (1992[a]) Overexpression of plant histidinol dehydrogenase using a baculovirus vector system. *Arch.Biochem.Biophys.* **295**, 235-239.

Nagai A., Suzuki K., Ward E., Moyer M., Mano J., Beck J., Tada S., Hashimoto M., Chang J-Y. and Ryals J. (1992[b]) Histidinol dehydrogenase in higher plants. purification, cloning and expression. In: *Research in Photosynthesis*, Vol IV edn. pp. 95-98. Dordrecht, The Netherlands: Kluwer Academic Publishers

Oriol E., Mendez-Alvarez S., Barbe J. and Gibert I. (1996) Cloning of the *Rhodobacter sphaeroides hisI* gene: unifunctionality of the encoded protein and lack of linkage to other *his* genes. *Microbiology* **142**, 2071-2078.

Paluk J.L., Orbach M.J., Legerton T.L. and Yanofsky C. (1988) The cross pathway control gene of *Neurospora crassa, cpc-1*, encodes a protein similar to GCN4 of yeast and the DNA binding domain of the oncogene *v-jun*-encoded protein. *Proc.Natl.Acad.Sci.USA*. **85**, 3728-3732.

Parsons S.M. and Koshland Jr.D.E. (1974) Multiple Aggregation States of Phosphoribosyladenosine triphosphate Synthetase. *J.Biol.Chem.* **249**, 4119-4126.

Parsons S.M. and Koshland Jr.D.E. (1974) A rapid isolation of Phosphoribosyladenosine triphosphate synthetase and comparison to native enzyme. *J.Biol.Chem.* **249**, 4104-4109.

Pereira S.A. and Livi G.P. (1995) A GCN-like response in *Candida albicans*. *Cell Biology International* **19**, 65-69.

Perler F.B., Davies E.O., Dean G.E., Gimble F.S., Jack W.E., Neff N., Noren C.J., Thorner J. and Belfort M. (1994) Protein splicing elements: inteins and exteins- a definition of terms and recommended nomenclature. *Nucleic Acids Res* **22**, 1125-1127.

Perrin D.D. and Dempsey B. (1974) *Buffers for pH and metal ion control*, London: Chapman and Hall.

Pillmoor J.B., Lindell S.D., Briggs G.G. and Wright K. (1995) The influences of Molecular Mechanisms of Action on Herbicide Design. In: *American Chemical Society*, pp. 292-303.

Piotrowska M. (1980) Cross pathway regulation of ornithine carbamoyltransferase synthesis in *Aspergillus nidulans*. *J.Gen.Microbiol.* **116**, 335-339.

Rost B. and Sander C. (1993) Prediction of protein structure at better than 70% accuracy. *J.Mol.Biol.* **232**, 584-599.

Rost B. and Sander C. (1993) Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc.Natl.Acad.Sci.USA.* **90**, 7558-7562.

Salt D.E., Smith R.D. and Raskin I. (1998) Phytoremediation. *Ann.Rev.Plant Physiology and Molecular Biology* **49**, 643-668.

Sambrook J., Fritsch E.E. and Maniatis T. (1989) *Molecular Cloning: a Laboratory Manual*, 2 edn. New York: Cold Spring Harbor.

Shanner D.L., Anderson P.C. and Stidham M.A. (1984) Imidazolinones. *Plant Physiol.* **76**, 545-546.

Shedlovsky A.E. and Magasanik B. (1962) A defect in histidine biosynthesis causing an adenine deficiency. *J.Biol.Chem.* **237**, 3725-3730.

Sissler M., Delorme C., Bond J., Ehrlich D., Renault P. and Francklyn C. (1999) An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc.Natl.Acad.Sci.USA.* **96**, 8985-8990.

Smith D.W.E. and Ames B.N. (1964) Intermediates in the early steps of histidine biosynthesis. *J.Biol.Chem.* **239**, 1848-1855.

Smith D.W.E. and Ames B.N. (1965) Phosphoribosyladenosine Monophosphate, an intermediate in histidine biosynthesis. *J.Biol.Chem.* **240**, 3056-3063.

Steinrucken H.C. and Amrhein N. (1980) The herbicide Glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic-3-phosphate synthase. *Biochem.Biophys.Res.Commun.* **94**, 1207-1212.

Stryer L. (1988) Introduction to Enzymes. In: *Biochemistry*, 3rd edn. New York: Freeman and Company

Studier F.W. and Moffatt B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *J.Mol.Biol.* **189**, 113-130.

Tada S., Volrath S., Guyer D., Scheidegger A., Ryals J., Ohta D and Ward E. (1994) Isolation and characterisation of cDNAs encoding Imidazoleglycerol phosphate dehydratase from *Arabidopsis thaliana*. *Plant Physiol.* **105**, 579-583.

Tada S., Hatano M., Nakayama Y., Volrath S., Guyer D., Ward E. and Ohta D. (1995) Insect cell expression of recombinant imidazoleglycerol phosphate dehydratase of *Arabidopsis* and wheat and inhibition by triazole herbicides. *Plant Physiol.* **109**, 153-159.

Tebar A.R., Fernandez V.M., Martindelrio R. and Ballesteros A.O. (1973) Studies on the quaternary structure of the first enzyme of histidine biosynthesis. *Experimentia* **29**, 1477-1479.

Tebar A.R., Fernandez V.M., Martindelrio R. and Ballesteros A.O. (1975) Fluorescence studies of phosphoribosyladenosine triphosphate synthetase of *Escherichia coli*. *FEBS Letters* **50**, 239-242.

Thoma R., Obmolova G., Lang D.A., Schwander M., Jeno P., Sterner R. and Wilmanns M. (1999) Efficient expression, purification and crystallisation of two hyperthermostable enzymes of histidine biosynthesis. *FEBS Letters* **454**, 1-6.

Umbarger H.E. (1987) Biosynthesis of branch chain amino acids. In: Neidhardt F.C., Ingraham J.L., Low K.B., Magasanik B., Schaechter M. and Umbarger H.E., (Eds.) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 352-367. Washington D.C.: Am. Soc. Microbiol.

Vinson C.R., Sigler P.B. and McKnight S.L. (1989) Scissors grip model for DNA recognition by a family of leucine zipper proteins. *Science* **246**, 911-916.

Vogt P.K., Bos T.J. and Doolittle R.F. (1987) Homology between the DNA binding domain of the GCN4 regulatory protein of yeast and the carboxyl terminal region of a protein coded for by the oncogene *jun*. *Proc.Natl.Acad.Sci.USA.* **84** , 3316-3319.

Voll M.J., Appella E. and Martin R.G. (1967) Purification and composition studies of Phosphoribosyl-adenosine triphosphate:Pyrophosphate Phosphoribosyltransferase, the first enzyme of histidine biosynthesis. *J.Biol.Chem.* **242**, 1760-1767.

White H.B. (1976) Coenzymes as fossils of an earlier metabolic state. *J.Mol.Evol.* **7**, 101-117.

Wiater A., Krajewska-Grynkiewicz K. and Klopotowski T. (1971) Histidine biosynthesis and its regulation in plants. *Acta Biochim Pol* **18**, 299-307.

Winkler M.E. (1988) Biosynthesis of Histidine. In: Neidhardt F.C., Ingraham J.L., Low K.B., Magasanik B., Schaechter M. and Umbarger H.E., (Eds.) *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*, pp. 395-411. Washington: American Society for Microbiology

Yanish-Peron C., Viera J. and Messing J. (1982) Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Zimmermann S., Baumann A., Jaekek K., Marbach I., Engelberg D. and Frohnmeyers H. (1999) UV responsive genes of *Arabidopsis* revealed by similarity to the Gcn-4 mediated UV response in yeast. *J.Biol.Chem.* 274, 17017-17024.

Appendix VII Purification of *E. coli* Phosphoribosyl ATP transferase

The overexpressing strain *E. coli* B834 (DE3)pLysS-EpHisG2 (constructed by Andrew Elwell) is grown overnight at 37°C in 500ml LB containing chloramphenicol and tetracycline. The following morning nine 500ml flasks containing both antibiotics are inoculated with 25ml of the overnight culture. Cells are grown to an A_{600} of 0.6 before inducing with 0.4mM IPTG. Cells are grown for a further 4 hours before harvesting and storing at -20°C until required.

The cells are broken by two passages through the French Press at 1000psi. The extract is clarified by centrifuging for one hour at 4°C and 18,000rpm. The extract is then loaded onto a DEAE Sephacel column which has been pre-equilibrated with Buffer A (50mM Tris-HCl pH 7.5, 0.4mM DTT and 1 protease inhibitor tablet per litre of buffer). The protein is eluted from the column with a linear gradient of 0-500mM NaCl, at approximately 350mM salt. The fractions are assayed for activity as described in section 2.10.1. Active fractions are pooled and dialysed against Buffer A overnight. The dialysed protein is loaded onto a Reactive Green 19 column (Sigma), which has been pre-equilibrated with Buffer A. The protein is washed for several hours with Buffer A before eluting with 200mM KCl in Buffer A. Again the fractions are assayed for enzyme activity and are dialysed overnight against Buffer A. The dialysed protein is then concentrated in an Amicon concentrator to approximately 2ml and is loaded onto a Sephacel S-200 column pre-equilibrated in Buffer B (500mM KCl, 50mM Tris-HCl pH 7.5, 0.4mM DTT). The column is run at 10ml per hour and 5ml fractions are collected. The fractions containing enzyme activity are determined and are pooled and the protein dialysed into Buffer C (50% glycerol, 50mM Tris-HCl pH 7.5, 0.4mM DTT) and is then stored at -20°C.