UNIVERSITY OF SOUTHAMPTON

STUDIES ON RECOMBINANT UBIQUITOUS AND ERYTHROID HUMAN PORPHOBILINOGEN DEAMINASE AND MUTATIONAL ANALYSIS OF *E. COLI* PORPHOBILINOGEN DEAMINASE

By
Julie Elizabeth Mosley

A thesis submitted for the degree of
Doctor of Philosophy

Division of Biochemistry and Molecular Biology
School of Biological Sciences
Faculty of Science
December 2001

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE

DIVISION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Doctor of Philosophy

STUDIES ON RECOMBINANT UBIQUITOUS AND ERYTHROID HUMAN PORPHOBILINOGEN DEAMINASE AND MUTATIONAL ANALYSIS OF *E. COLI* PORPHOBILINOGEN DEAMINASE

By Julie Elizabeth Mosley

Porphobilinogen deaminase, the third enzyme of the haem biosynthesis pathway catalyses the stepwise tetramerisation of porphobilinogen to form the unstable hydroxymethylbilane, preuroporphyrinogen. In humans, non-erythroid (ubiquitous) and erythroid forms of porphobilinogen deaminase are both specified by a single 10 kilobase gene comprising 15 exons.

Experiments in this thesis are described in which the cDNA specifying human porphobilinogen deaminase was cloned so that both ubiquitous and erythroid forms could be over-expressed in *E. coli* BL21(DE3). After purification to homogeneity, both isoenzymes were characterized by physical and chemical techniques, including mass spectrometry, amino-terminal sequencing and kinetic studies. Crystals of ubiquitous and erythroid porphobilinogen deaminases were produced using the hanging drop vapour diffusion method, and X-ray crystallographic data were collected for the erythroid isoform to 3.0Å resolution. A three-dimensional X-ray model was built by molecular replacement using the known human ubiquitous porphobilinogen deaminase arginine 167 glutamine mutant structure.

Experiments were also carried out on two mutants of *E. coli* porphobilinogen deaminase, Gly 57 Ala and Gly 264 Cys. Gly 57 is located on a mobile active site strand in domain 1 and Gly 264 is a surface residue of domain 3 close to Cys 134 in domain 2. It was hoped that investigations on these two mutants would provide information about conformational changes in the enzyme during the catalytic cycle.

Acknowledgements

I would like to express my gratitude to my supervisor Professor Peter Shoolingin-Jordan for providing me with the opportunity to study in his research group and his continued support through difficult circumstances.

I am forever indebted to my mother for her love, encouragement and amazing strength of character and I dedicate this thesis to her. I thank my father for emphasing the importance of education and supporting me in my undergraduate studies.

I would like to thank my stepfather and brothers for taking good care of me over the years and being there when I needed them.

I am extremely grateful to my boyfriend Simon who kept my spirits up when I was ready to throw in the towel and showed the patience of a saint during the later days of writing up.

I couldn't forget my dogs Jo and Tootsie for being a loyal friends and fun companions.

Special thanks to Dr. Muhammad Sarwar and Dr. Darren Thompson for their scientific advice and work in cloning and crystallography respectively. Thanks to Danica Butler for her work on the enzyme purification and Dr. Maria Russell for help preparing several figures and Aberr AlDbass for her enthusiam and collaboration on the purification and crystallography of the human deaminase.

I am grateful to Dr. John Langley for his help with mass spectrometry and being a great friend. Thanks to Neville wright, Paul Skipp, Professor Steve Wood and all the crystallography group who helped tremendously. Thanks to all the members of the Jordan group, past and present, who ensured life at Southampton was rarely dull !

Finally, a big thanks to my colleagues at GlaxoSmithKline for their understanding and giving me time off to complete this manuscript.

CONTENTS

1 Introduction

2 Materials and Methods

3 Studies on human porphobilinogen deaminase: cloning, expression and purification of the two isoforms

ABBREVIATIONS AND NOMENCLATURE

Abbreviation	
A	Adenine
Abs	Absorbance
AIP	Acute intermittent porphyria
C	Cytosine
CRIM	Crossreactive immunoreactive material
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic acid
DTNB	5,5'-Dithiobis(2-nitrobenzoic acid)
ECPCR	Expression cassette polymerase chain reaction
EDTA	Ethylenediaminetetra-acetic acid
f.p.l.c.	Fast protein liquid chromatography
G	Guanine
kb	Kilobase
KDa	Kilodalton
NEM	*N*-ethylmaleimide
n.m.r.	Nuclear magnetic resonance spectroscopy
PAGE	Polyacrylamide gel electrophoresis
PBG	Porphobilinogen
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
SDS	Sodium dodecyl sulphate
T	Thymine
TEMED	N,N,N',N'- Tetramethylenediamine
Tris	Trishydroxymethylaminoethane
UV	Ultraviolet
U	Uracil

Chapter 1 Introduction

Tetrapyrroles biosynthesis takes place in the vast majority of living organisms. Modified tetrapyrroles are macromolecules that form prosthetic groups for proteins, many of which are involved in vital energy-generating reactions including respiration, photosynthesis and methanogenesis. In respiration the transport of oxygen to the muscles and the removal of waste carbon dioxide is mediated by haemoglobin, a tetrameric protein that contains the prosthetic group haem (figure 1.1). Sirohaem acts as a cofactor for the bacterial enzyme sulphite reductase and chlorophylls are the light harvesting pigments of plants responsible for the trapping of solar energy in photosynthesis (figure 1.1). The structurally complex tetrapyrrole, vitamin $B_{12,}$ is used as a cofactor in transformation reactions that use free radicals. Methanogenic bacteria utilise Factor F_{430} as a cofactor for the enzyme methyl coenzyme reductase involved in the production of methane. The versatility of the tetrapyrrole macrocycle is attributable to the variety of substituents on the outside of the macrocyclic ring and to the unique environment and metal ligand geometry provided by the protein.

1.1 Biosynthesis of tetrapyrroles

The tetrapyrrole biosynthetic pathway is broadly analogous in every living organism, with 5-aminolaevulinic acid being an essential precursor, eight molecules of which form uroporphyrinogen III from which all other tetrapyrroles are formed. The first common intermediate, 5-aminolaevulinic acid (5-ALA), is synthesised by one of two distinct routes according to the species. The next three enzymatic steps for the synthesis of uroporphyrinogen III are ubiquitous to all living organisms. In the first of these reactions, 5-ALA dehydratase catalyses the condensation of two molecules of 5-ALA forming porphobilinogen. In the second reaction, porphobilinogen deaminase catalyses the polymerisation of four molecules of porphobilinogen producing preuroporphyrinogen, a highly unstable molecule. Preuroporphyrinogen is rapidly converted into uroporphyrinogen III by uroporphyrinogen III synthase in a reaction involving dehydration, D ring rearrangement and cyclisation. Once formed uroporphyrinogen III can have

several possible fates due to branching of the tetrapyrrole biosynthetic pathway. It can be converted into protoporphyrin IX for the production of haem, chlorophylls and bacteriochlorophylls or methylated on route to sirohaem, vitamin B_{12} and factor F_{430}. Although no single organism utilises all these final products, an overview showing the possible branches of tetrapyrrole biosynthetic pathway is shown in figure 1.2. All these pathways have been extensively studied and detailed reviews published, to which the reader should refer for details beyond the scope of this thesis (Shoolingin-Jordan and Cheung, 1999).

1.1.1 The synthesis of 5-aminolaevulinic acid by the Shemin pathway

In mammals, eukaryotes and photosynthetic bacteria (5-ALA) is synthesised from glycine and succinyl-CoA by the 'Shemin' pathway (Shemin and Russell, 1953). 5-Aminolaevulinic acid synthase (5-ALA synthase), a pyridoxal 5'-phosphate requiring enzyme, catalyses the condensation of glycine and succinyl-CoA. Reaction products include CoA-SH and carbon dioxide as well as 5-ALA (figure 1.3). The cofactor pyridoxal 5'-phosphate is bound to the enzyme through a schiff base with an invariant lysine residue. In mice this invariant residue has been identified as lysine 313. Mutation of lysine 313 drastically reduces the activity of the enzyme, although the cofactor remains non-covalently bound (Ferreira *et al.*, 1995). Further work implicated lysine 313 as a catalytic residue involved in the removal of the *pro*R hydrogen atom from glycine and final transfer to the 5'-position of 5-aminolaevulinic acid (Hunter and Ferreira, 1999;Shoolingin-Jordan, 1998.

1.1.2 Regulation of 5-aminolaevulinic acid synthase

5-ALA synthase, the key regulatory enzyme, is very tightly regulated, both pre- and post- transcriptionally. This is essential not only for controlling the cellular haem levels, but also for preventing the build-up of pathway intermediates which have neurological and cytotoxic photodynamic properties. In the humans the regulatory mechanisms are complicated because of the existence of two isoforms

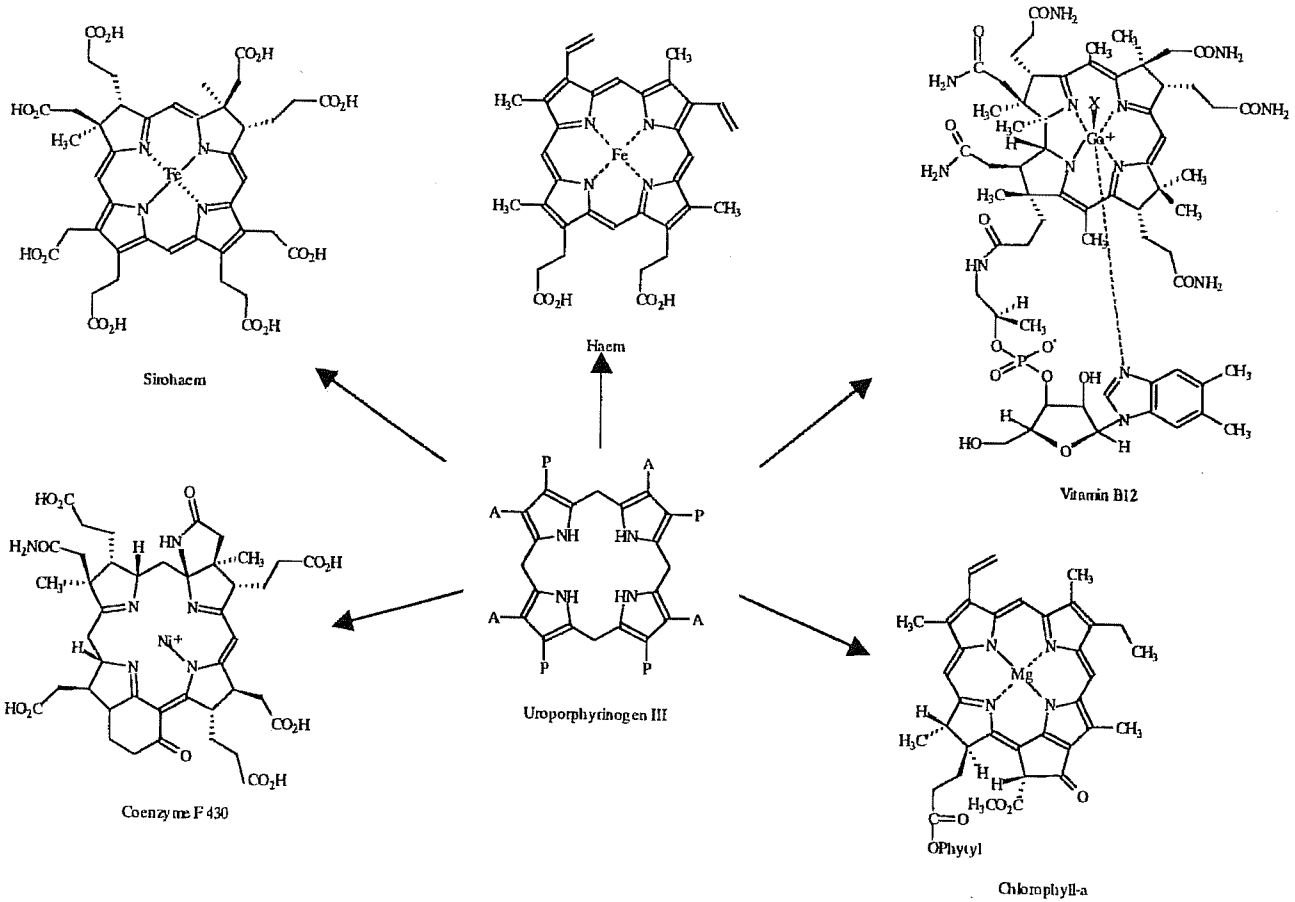


Figure 1.1 The structures of the modified tetrapyrroles haem, sirohaem, chlorophyll-a, vitamin B_{12} and Factor F_{430}.

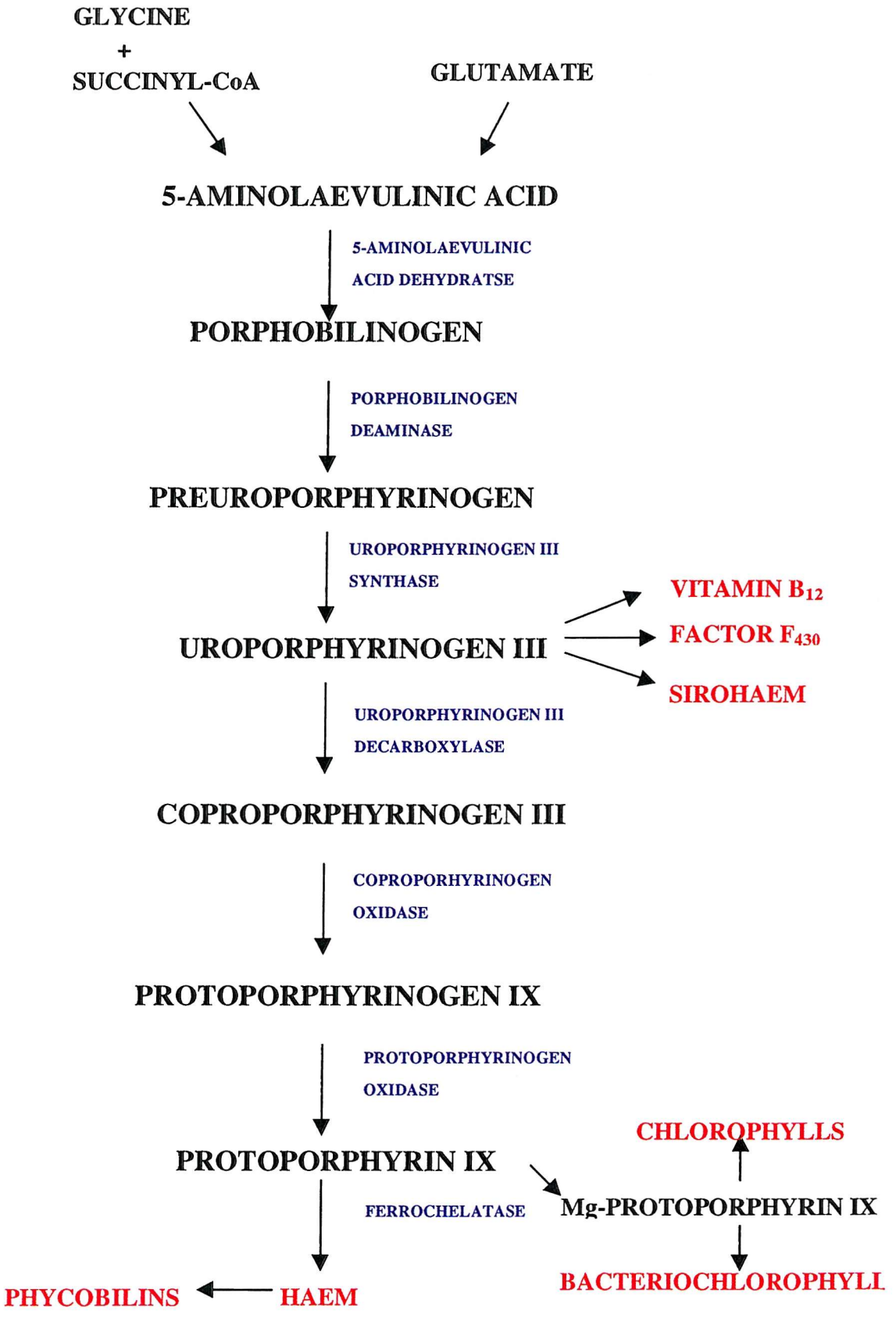


Figure 1.2 A summary of all the possible tetrapyrrole biosynthetic pathways.
Uroporphrinogen III, the universal precursor, is converted into protoporphyrin IX for the production of haem, chlorophylls and bacteriochlorophylls or methylated on route to sirohaem, vitamin B_{12} and factor.

of 5-ALA synthase a ubiquitous form and an erythroid form, encoded by the ALAS-1 and ALAS-2 genes respectively. The ubiquitous enzyme is present in all tissues, but the erythroid enzyme is expressed only in erythrocytes. The two isoenzymes are expressed from different genes and have considerable homology at the C-terminus, which contains the core region. The N-terminal region is much more variable and is involved in mitochondrial import and regulatory control mechanisms.

Erythroid 5-ALA synthase is regulated both at the level of transcription and translation. In its promoter region the ALAS-2 gene contains sequences that are involved in transcriptional regulation in the differentiating erythrocyte. At the level of translation, erythroid ALA-synthase is regulated by iron levels through interaction with an iron-responsive element also found in mRNAs specifying ferritin and transferrin, two proteins involved in iron metabolism (May *et al.*, 1990). Thus regulation of ALAS-2 is linked to availability of iron, an essential mechanism in the developing erythrocyte where iron is in high demand for incorporation into haemoglobin. Ubiquitous 5-ALA synthase in chicken liver cells is inhibited directly by haem, a good example of end-product regulation. Also the translocation of the enzyme from the cytosol into the mitochondrion is inhibited in the same cells (May *et al.*, 1990).

1.1.3 The synthesis of 5-aminolaevulinic acid by the C_5 pathway

In higher plants and many prokaryotic systems, synthesis proceeds by the 'C_5 pathway' that utilises the carbon skeleton of glutamate (Beale and Castelfranco, 1973;Beale *et al.*, 1975)(figure 1.3). This pathway is of great interest since it is one of the few that utilises a tRNA molecule, as glutamyl-tRNA, in a process other than protein synthesis (Schon *et al.*, 1986). This observation lead to the speculation that control of protein synthesis and tetrapyrrole biosynthesis are linked. Glutamyl tRNA synthase catalyses the attachment of glutamate to $tRNA^{glu}$ in a reaction which requires ATP and magnesium ions. The glutamyl-$tRNA^{glu}$ is reduced to glutamate-1-semialdehyde in a NADPH requiring reaction catalysed by glutamyl-tRNA reductase. Glutamate-1-semialdehyde, a highly

reactive α-aminoaldehyde, is readily isomerised yielding 5-ALA, the final product of the C5 pathway (Kannangara *et al.*, 1988).

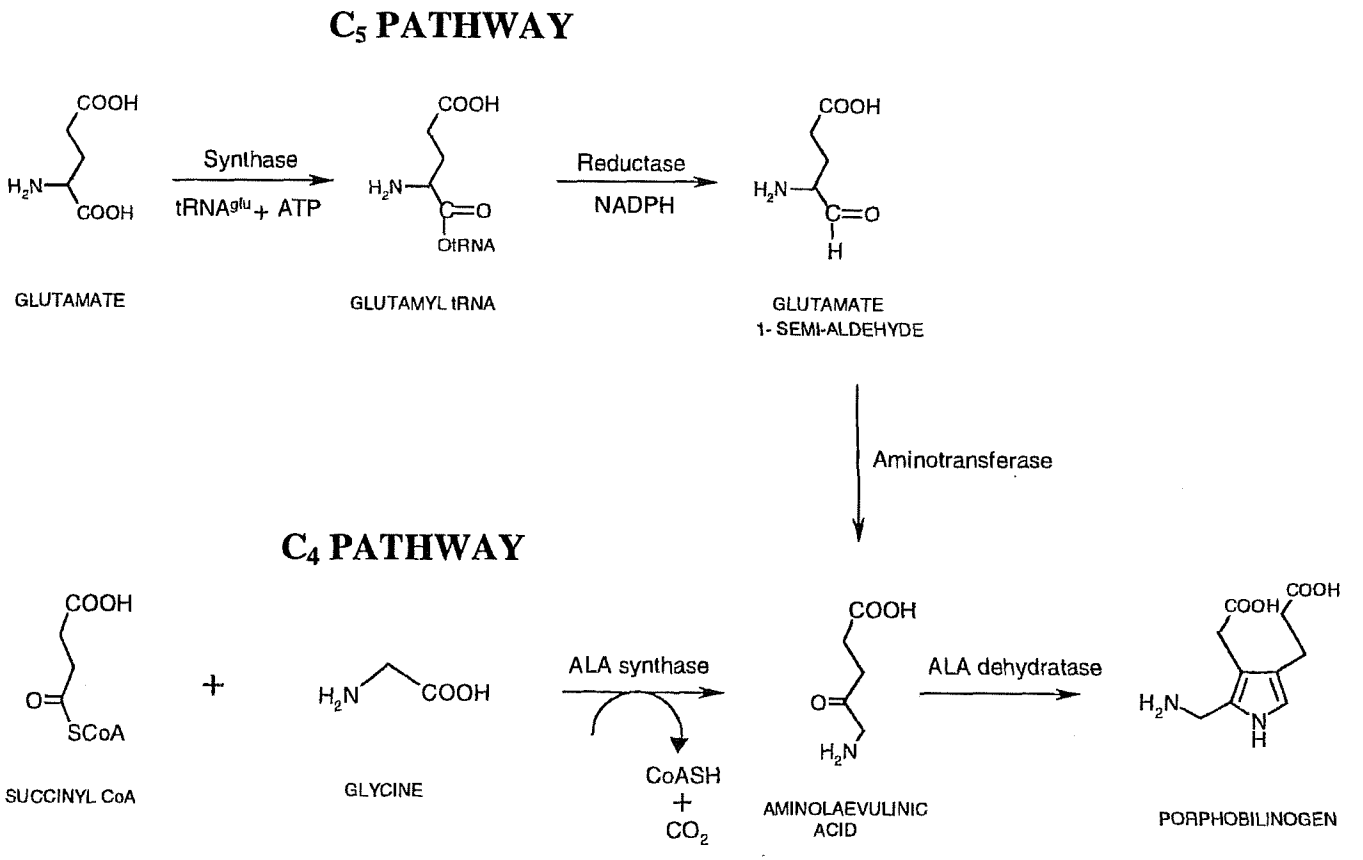


Figure 1.3 The alternative pathways used for 5-ALA synthesis. Higher plants and many prokaryotic systems synthesis ALA by the C5 pathway, which utilises the carbon skeleton of glutamate. The C4 or Shemin pathway occurs in mammals and eukaryotes and synthesises ALA from succinyl-CoA and glycine.

1.1.4 5-Aminolaevulinic acid dehydratase

ALA dehydratase catalyses the Knorr-type condensation of two molecules of 5-ALA yielding porphobilinogen, the fundamental pyrrole building block. The substrate, 5-ALA, is intrinsically reactive and can readily dimerise in a non-enzymic reaction forming dihydropyrazine instead of the desired pyrrole product PBG (figure 1.4). To overcome this problem the enzyme has two substrate binding sites, known as the 'A-site' the 'P-site', each of which binds a single molecule of 5-ALA (Spencer and Jordan, 1995). The enzymic reaction proceeds by formation of a Schiff base between ALA bound at the P-site and a nearby lysine residue. Binding of the second molecule of 5-ALA at the A-site results in a condensation reaction producing porphobilinogen, with the elimination of two water molecules.

1.1.5 Porphobilinogen deaminase

The next enzyme in the pathway, porphobilinogen deaminase (PBGD; other names include uroporphyrinogen I synthase and 1-hydroxymethylbilane synthase) catalyses the tetrapolymerisation of four molecules of the monopyrole, porphobilinogen, to produce the highly unstable intermediate preuroporphyrinogen (figure.1.5). Porphobilinogen deaminase, the enzyme investigated in this thesis, is discussed in detail in Section 1.2.

1.1.6 Uroporphyrinogen III synthase

The final step in the production of uroporpyrinogen III is catalysed by uroporphyrinogen III synthase (also called cosynthase). This enzyme catalyses the cyclisation of the linear substrate, preuroporphyrinogen, with the reversal of the D pyrrole ring, to yield uroporphyrinogen III, the template for all tetrapyrroles. It is thought that D ring inversion is achieved through the formation of a *spiro*-intermediate. This hypothesis was initially derived over 40 years ago (Mathewson and Corwin, 1963). Four possible isomers of uroporphyrinogen exist, but only uroporphyrinogen III is of biological

significance. In the absence of uroporphyrinogen III synthase, preuroporphyrinogen is cyclised non-enzymically forming uroporphyrinogen I (figure 1.5).

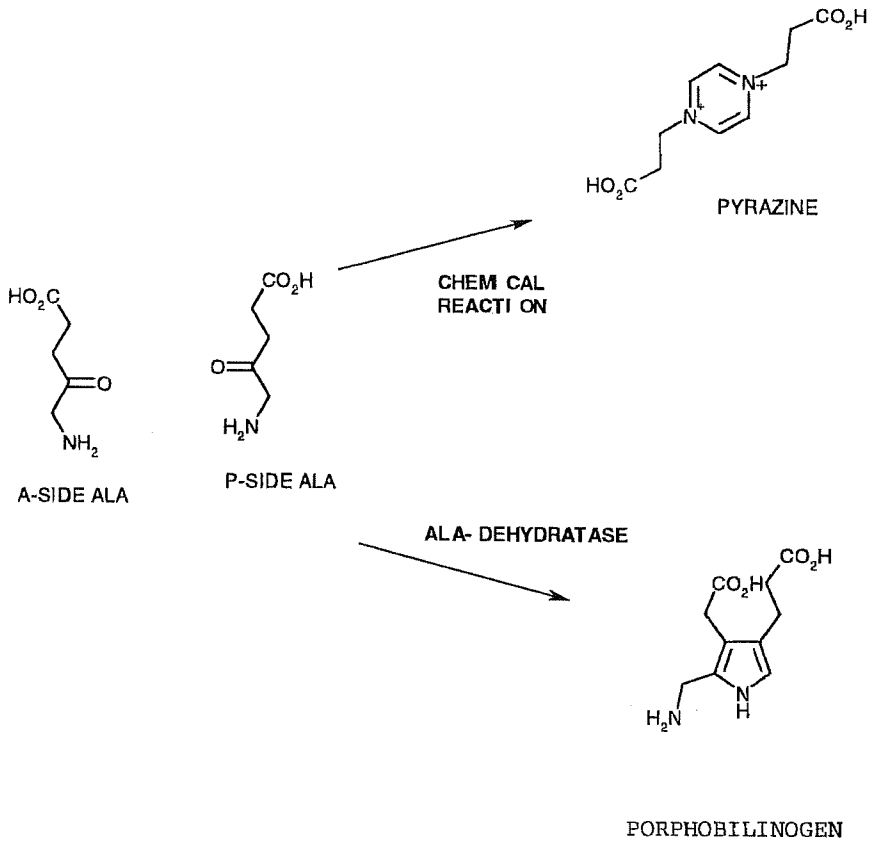


Figure 1.4 The dimerisation reaction catalysed by ALA-dehydratase. The A-side and P-side ALA become the acetate and proprionate side chains of PBG respectively. Dihydropyrazine is formed from the non-enzymic reaction between two molecules of ALA and is converted to pyrazine.

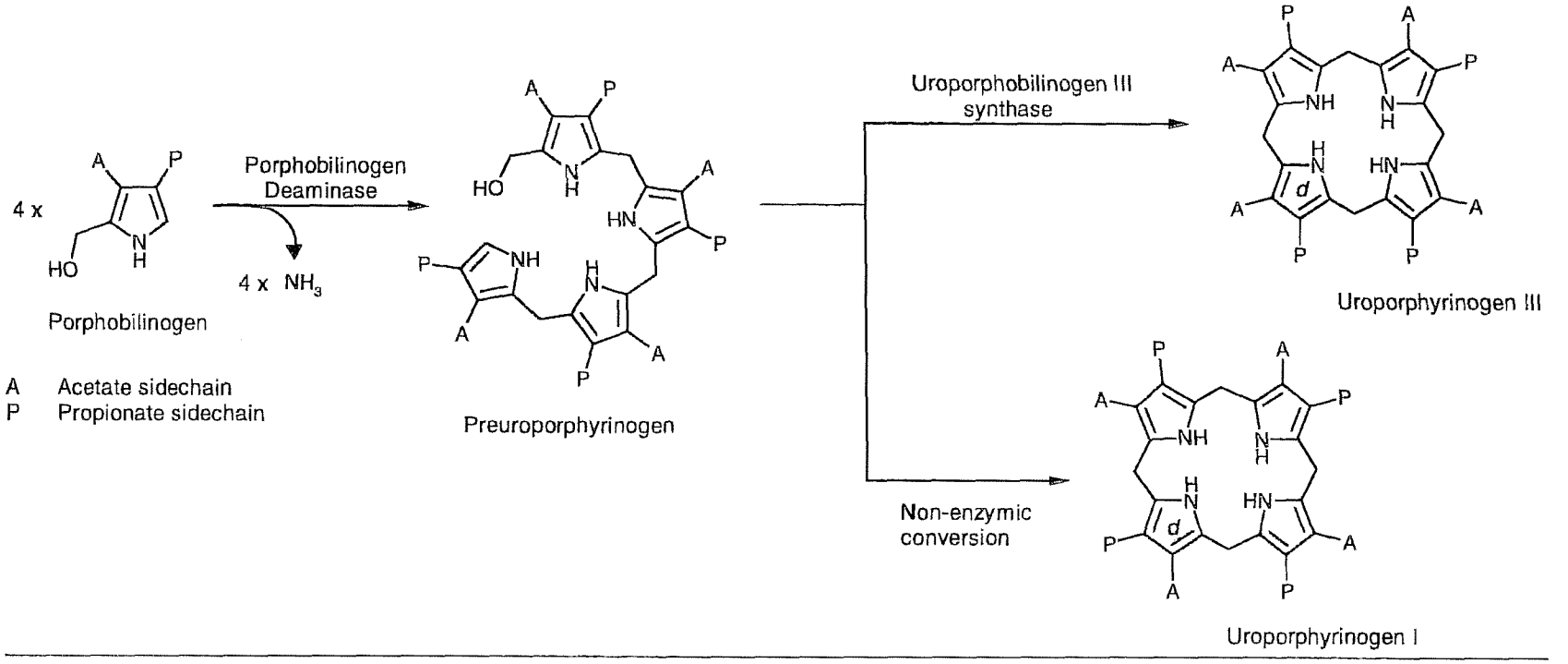


Figure 1.5 The sequential reactions catalysed by porphobilinogen deaminase and uroporphyrinogen III synthase. Porphobilinogen deaminase catalyses the tetrapolymerisation of porphobilinogen yielding the hydroxymethlybilane, preuroporphyrinogen. The next enzyme,uroporphyrinogen III synthase, catalyses the cyclisation of the linear substrate, preuroporphyrinogen, with the reversal of the D pyrrole ring, to yield uroporphyrinogen III. In the absence of uroporphyrinogen III synthase, preuroporphyrinogen is converted into the product uroporphyrinogen I, which has no known biological function.

1.1.7 The production of modified tetrapyrroles from uroporphyrinogen III

The extensive biological functions of the cyclic tetrapyrroles are only manifested by products subsequent to uroporphyrinogen III, the major branch point in the tetrapyrrole pathway. Decarboxylation commits the route towards the biosynthesis of protoporphyrin IX and on to haems and chlorophylls. Alternatively, C-methylation of uroporphyrinogen III directs the pathway towards the synthesis of sirohaem, factor F_{430} and vitamin B_{12} (figure 1.6).

1.1.8 The pathway from uroporphyrinogen III to haem and chlorophyll

Uroporphyrinogen III decarboxylase catalyses the conversion of four acetic acid chains of uroporphyrinogen III to methyl groups, producing coproporphyrinogen III. The conversion of coproporphyrinogen III into protoporphyrinogen IX is achieved by the oxidative decarboxylation of the a and b ring propionate side chains to give vinyl groups (figure 1.6). This reaction is catalysed by coproporphrinogen III oxidase (Sano and Granick, 1961). Protoporphyrinogen IX is then oxidised to protoporphyrin IX by the enzyme protoporphyrinogen oxidase. All haems with the exception of sirohaem and haem d_1, are derived from protoporphyrin IX. The enzyme ferrochelatase catalyses the insertion of iron (Fe^{2+}) into protoporphyrin IX in a mechanism where the ring undergoes 'bending' to promote the binding of the metal ion (Dailey *et al.*, 1989). The product of this reaction is protohaem IX, also termed haem *b*, from which all haems (a, c, d, o) and derived products are formed.
In chlorophyll biosynthesis, protoporphyrin IX incorporates a magnesium ion forming Mg-protoporphyrin. The propionic acid side chain of the c ring is subsequently methylated forming Mg-protopophyrin IX monomethlyester. Further reactions results in isocyclic ring formation and the production of protochlorophyllide. This compound provides the branch point for production of chlorophyllide a and chlorophyll c. Chlorophyllide a is the precursor to all other chlorophylls (a and b) and the bacteriochlorophylls (a, b, c, d, e and g).

1.1.9 The pathway from uroporphyrinogen III to vitamin B_{12}, sirohaem and factor F_{430}

Uroporphyrinogen III is converted into dihydrosirohydrochlorin by a single enzymic step. This product is the common intermediate for the biosynthesis of vitamin B_{12}, sirohaem and factor F_{430}. Vitamin B_{12} is one of the most complex biological molecules known, and its synthesis involves around 30 enzyme steps (Raux *et al.*, 2000). Dihydrosirohydrochlorin undergoes a series of methylations using S-adenosyl methionine, followed by decarboxylation, ring contraction and cobalt insertion to form cobyrinic acid via several precorrin intermediates. Cobyrinic acid is then converted via cobyric acid and cobinamide, to cobalamin and finally vitamin B_{12}. Sirohaem is the prosthetic group of assimilatory nitrile and sulphate reductases in bacteria and plants. Sirohaem is also synthesised from uroporphyrinogen III. In the anaerobic bacterium *S. typhimurium*, this is produced for vitamin B_{12} biosynthesis. The production of factor F_{430} from dihydrosirohydrochlorin occurs in methanobacteria. The latter compound is converted to factor F_{430} by insertion of nickel, amidation of the a and b rings acetate groups, reduction of two double bonds, cyclisation of the ring b acetamide and the ring d propionic acid group. However, the exact mechanism of conversion and the number of steps are to be elucidated.

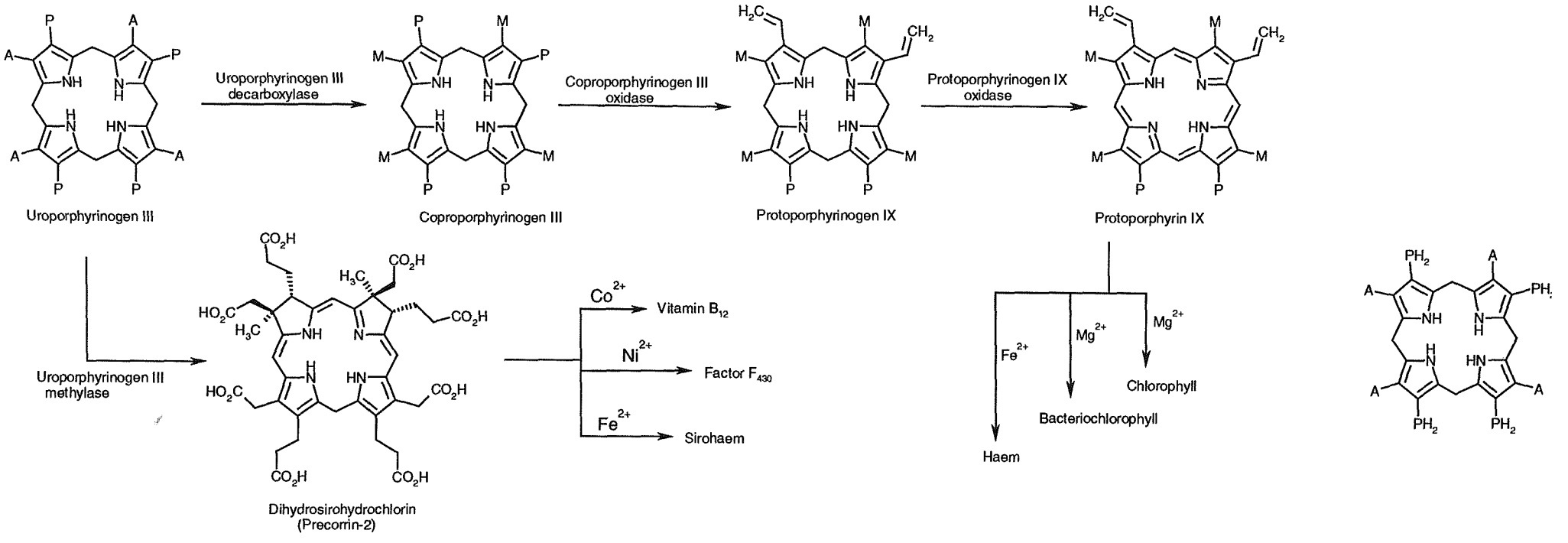


Figure 1.6 The production of modified tetrapyrroles from uroporphyrinogen III. Decarboxylation to coporporphyrinogen III commits the route towards the synthesis of haems and chlorophylls. Alternatively C-methylation to precorrin-2 directs the pathway towards the synthesis of sirohaem, factor F_{430} and vitamin B_{12}.

1.2 Studies on Porphobilinogen deaminase

Porphobilinogen deaminases have been purified to homogeneity and characterised from many sources, including the recombinant *Escherichia coli* enzyme (table 1.1). In general, the deaminases exist as monomeric proteins with M_r ranging from 34k - 44kDa. They have similar properties including heat stability, isoelectric points between pH 4 and 4.5 and pH optima between 8.0-8.5. The K_m values for the substrate porphobilinogen are in the low μmolar range and the turnover of the enzyme is low at 0.5 sec^{-1}.

Source	**Reference**
Rhodobacter sphaeroides	Jordan and Shemin, 1973 Davies and Neuberger, 1973
Spinach	Higuchi and Bogorad, 1975
Human erythrocytes	Anderson and Desnick, 1980 Smythe and Williams, 1988
Chlorella regularis	(Shioi *et al.*, 1980)
Euglena gracilis	(Williams *et al.*, 1981)
Rat liver	(Williams, 1984)
Escherichia coli	Thomas and Jordan, 1986 (Hart *et al.*, 1986)
Pig liver	(Braithwaite and Russell, 1987)
Yeast	(Correa, Sr. *et al.*, 1991)
Pisum sativum	(Spano and Timko, 1991)
Arabidopsis thaliana	(Jones and Jordan, 1994)
Human hepatic	(Mazzetti and Tomio, 1990)
Scenedesmus obliques	(Juknat *et al.*, 1994)

Table 1.1 A summary of the various sources of purified deaminases.

1.2.1 Early experiments into porphobilinogen deaminase

The enzymatic production of uroporphyrinogen III from porphobilinogen was first demonstrated in spinach extracts by Bogorad and Granick (Bogorad and Granick, 1953). If these extracts were heated, before incubation with porphobilinogen, only uroporphyrinogen I was formed. This indicated that a heat sensitive enzyme was responsible for the isomerisation reaction. Porphobilinogen deaminase (PBGD) was later isolated from spinach leaf tissue and was shown to catalyse the formation of uroporphyrinogen I from four molecules of porphobilinogen. During these investigations, uroporphyrinogen III synthase (cosynthase) was isolated from wheat germ. When the cosynthase enzyme was added to the deaminase, uroporphyrinogen III was formed. However, uroporphyrinogen III synthase was neither able to catalyse the isomerisation of uroporphyrinogen I into uroporphyrinogen III, nor to act solely on porphobilinogen. Consequently, in 1963, Bogorad proposed that a polypyrrole intermediate was formed by the action of the deaminase and that this was the substrate for uroporphyrinogen III synthase. However, it was not known whether the two enzymes acted separately or together as a multifunctional enzyme complex (Higuchi and Bogorad, 1975). The latter hypothesis was supported through the observation that PBGD and uroporphyrinogen III synthase copurified in many sources through several stages of their isolation process.

1.2.2 Discovery of the preuroporphyrinogen product

After the proposal of Bogorad in 1963, he showed that the formation of porphyrinogens by PBGD, but not the rate of porphobilinogen consumption, was inhibited by the action of ammonia. The porphobilinogen deaminase was shown to produce an uncyclised linear tetrapyrrole (Radmer and Bogorad, 1972) and it was thought that this tetrapyrrole was a linear 1-aminomethylbilane ($NH_2CH_2APAPAPAP$), where A=acetate and P=propionate. Although this is not in fact the natural intermediate in the reaction, the aminomethylbilane was used to establish that the D ring of uroporphyrinogen III is rearranged at the linear bilane stage rather than at the di or tripyrrole stage (Battersby, 1978). The

discovery of preuroporphyrinogen as the true intermediate was first made from the observation of NMR signals from samples of PBGD incubated with [11-^{13}C] porphobilinogen. This observation eliminated the 1-aminomethylbilane as a possible intermediate, and the analogous 1-hydroxymethlybilane, preuroporphyrinogen was proposed as the alternative. Preuroporphyrinogen was generated by deaminase and shown to spontaneously form uroporphyrinogen I with a $t_{1/2}$ of 4.5 minutes at pH 8.0. It was shown that rapid formation of uroporphyrinogen III occurs upon incubation of preuroporphyrinogen with purified uroporphyrinogen III synthase, in the absence of the PBGD. From this work, preuroporphyrinogen was established as the true substrate for the synthase enzyme (Jordan and Seehra, 1979).

Further work showed that there were no significant differences observed in the rate of formation of uroporphyrinogen III in the presence of varying ratios of PBGD and uroporphyrinogen III synthase. It was therefore considered highly unlikely that the two enzymes have any specific interaction between each other *in vivo,* but act independently and sequentially in the overall production of uroporphyrinogen III from porphobilinogen.

1.2.3 Mechanism of porphobilinogen deaminase

The discovery that the deaminase product is a linear tetrapyrrole stimulated investigations to determine the order of pyrrole assembly in the tetrapolymerisation reaction. Single turnover reactions were undertaken using two approaches: a radiochemical approach by (Jordan and Seehra, 1979), and a ^{13}C NMR approach (Battersby *et al.*, 1979). The radiochemical approach produced labelled ^{14}C-preuroporphyrinogen using ^{14}C-porphobilinogen. The labelled product was converted to protoporphyrin IX that could be degraded into two compounds: ethlymethylmalemide (from pyrrole rings a and b) and haematinic acid (from pyrrole rings c and d). The radiolabel was found associated mainly with the ethylmethlymalemide, suggesting that the mechanism proceeded by the addition of the rings in the order a, b, c and d (figure 1.7). These findings were confirmed by the ^{13}C NMR approach (Battersby *et al.*, 1979).

1.2.4 Isolation of the enzyme-intermediate complexes

Incubation of PBGD with tritiated porphobilinogen (Anderson and Desnick, 1980) followed by subjection of the enzyme to gel electrophoresis led to the formation of labelled protein bands. This suggested intermediates had been formed with one, two, three and four pyrrole units linked to the enzyme. These enzyme-substrate intermediates are referred to as ES, ES_2, ES_3 and ES_4, respectively. The complexes have been found with deaminases from both *R. sphaeroides* (Berry *et al.*, 1981) and *E. coli* (Warren and Jordan, 1988), although the bacterial ES_4 complex was found too unstable to isolate. Sodium dodecylsulphate treatment of the enzyme-substrate complexes did not cause the bound substrate to be released providing strong evidence that a covalent link had been formed between the enzyme and the bound substrate molecules.
Further work has shown that the α-terminal pyrrole of each complex is exchangeable and that the enzyme intermediate complexes are interconvertible (Warren and Jordan, 1988). In addition, the highly labile ES_4 complex may be visualised by electrospray mass spectrometry.

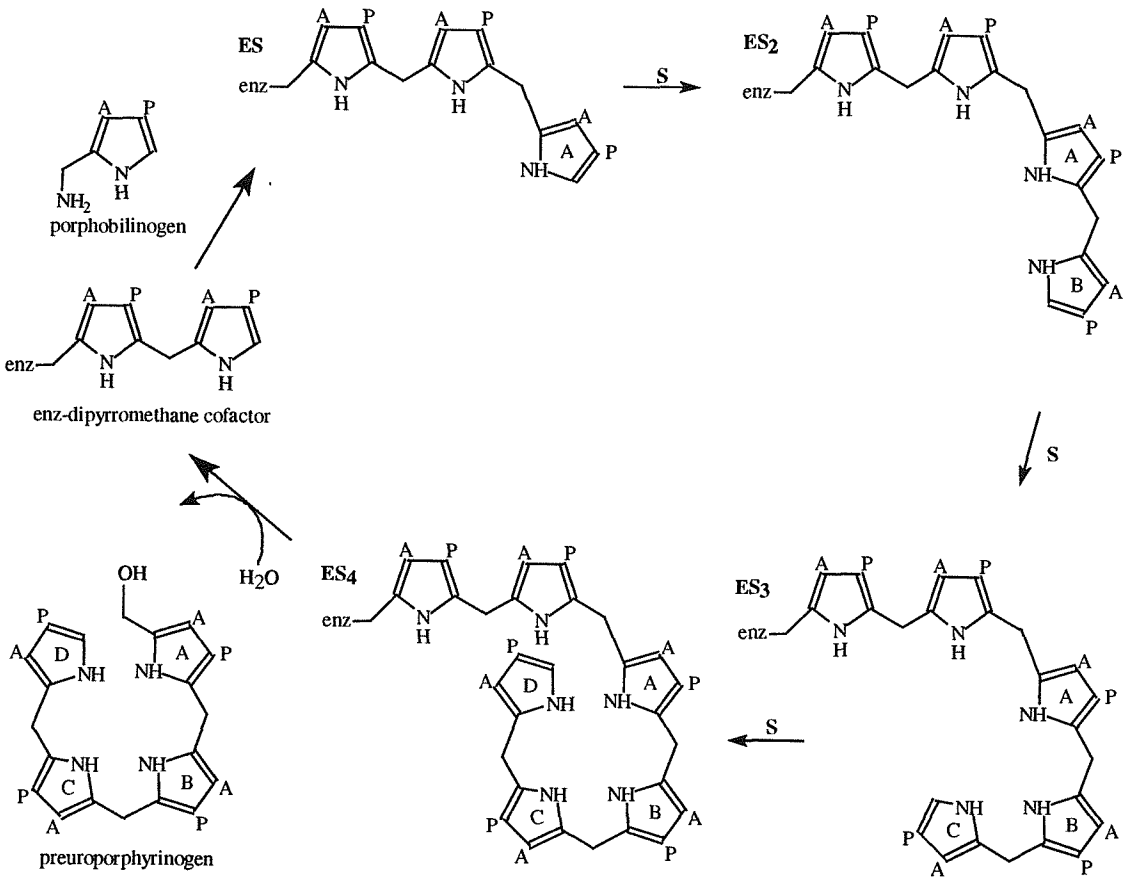
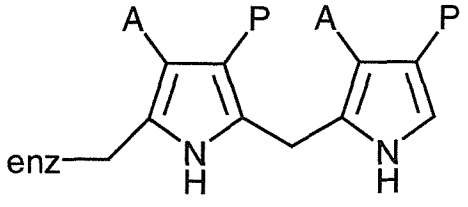


Figure 1.7 The stepwise addition of porphobilinogen substrate to the enzyme. The rings are labelled A, B, C, and D in order of assembly.

1.2.5 Establishment of the dipyrromethane cofactor as the substrate binding site

The enzymic group that bound the first substrate molecule to deaminase was of interest to many researchers. Jordan and Berry showed that the group was not positively charged (Jordan and Berry, 1981). Meanwhile, the use of site directed modifying agents by Russell and co-workers had suggested the involvement of a cysteine residue (Russell and Rockwell, 1980). N.M.R. studies using PBGD purified from *R. sphaeroides* and tritiated porphobilinogen also suggested the involvement of a cysteine residue in the covalent linkage to the substrate (Evans *et al.*, 1986).

A breakthrough in the study of PBGD resulted from the identification, cloning and overexpression of the *E. coli* deaminase gene (Thomas and Jordan, 1986). This allowed the production of milligram quantities of pure, recombinant enzyme, which facilitated future work. The treatment of purified recombinant *E. coli* PBGD with Ehrlich's reagent gave a reaction typical of a dipyrromethane (Pluscec and Bogorad, 1970). More conclusive evidence was provided with the reaction of the isolated ES_2 complex with Ehrlich's reagent that gave a reaction typical of a tetrapyrrole. This suggested that two molecules of substrate were covalently linked to two rings of an enzyme-bound dipyrromethane. From radiochemical studies, it was concluded that the dipyrromethane is resident at the active site and provides the attachment site for the covalent binding of the substrate. This linked dipyrromethane was therefore termed the dipyrromethane cofactor (Jordan and Warren, 1987) (figure 1.8).



enz-dipyrromethane cofactor

Figure 1.8 The enzyme dipyrromethane cofactor of PBGD.

Once the cofactor structure had been established, the nature of the linkage between the enzyme and cofactor was examined. One group of researchers treated the *E. coli* PBGD with the thiol reagent DTNB under denaturing conditions and found three reactive thiol groups. However, a study of the *E. coli* gene derived amino acid sequence showed that four thiol groups are present, suggesting that the fourth was binding the dipyrromethane cofactor. Peptide cleavage experiments, combined with knowledge of the conserved residues, suggested cysteine 242 as the residue responsible for binding the cofactor. Further evidence came from site-directed mutagenesis studies where each of the conserved cysteine residues were separately mutated to serine (Jordan *et al.*, 1988b;Scott *et al.*, 1988) showing that the C242S mutant had a greatly reduced enzyme activity.

1.2.6 Mutagenesis studies on *E. coli* porphobilinogen deaminase

A number of mutants have been generated (table 1.2), providing essential information about cofactor assembly, substrate binding and catalysis (Jordan and Woodcock, 1991). As previously mentioned, mutation of cysteine 242 had a dramatic effect on cofactor assembly. Other mutations also provided important information about other conserved residues. Mutagenesis of arginines 131 and 132 to histidine or leucine prevented cofactor assembly and resulted in apodeaminases lacking catalytic activity. Mutagenesis of arginines 11 and 155 resulted in the inability of the enzyme to bind substrate, while substitution of arginines 149, 176 and 232 affected the polymerisation mechanism (Jordan and Woodcock, 1991;Lander *et al.*, 1991).

The availability of the X-ray structure has provided important information and pointed the way to the construction of additional mutants to investigate other structure/function relationships of the enzyme. One such study is the mutation of the proposed catalytic residue aspartate 84. A glutamate 84 mutant possesses less than one percent of the wild-type activity with a cofactor that is more sensitive to oxidation (Woodcock and Jordan, 1994). The reasons for this are evident from the X-ray structure, which shows the loss of a hydrogen bond, normally found between aspartate 84 and the NH group of the C2 pyrrole ring.

Residue	Structural environment/role of side chain
R11	Salt bridges to the propionate group of substrate porphobilinogen and D46.
Q19	Line the active site cleft, hydrogen bonds to S81 and R176 side chains.
G45	Adjacent to F62 at interface between two portions of the lid region.
D76	Salt bridges to both R7 and R206, possibly involved in structural stability.
V79	Packed in hydrophobic core.
H80	Lines floor of the cleft, and forms a salt bridges to E204.
S81	Located below the cofactor, hydrogen bonds the C2 cofactor ring and Q19.
K83	Salt bridges to both cofactor acetates.
D84	Hydrogen bonds to the cofactor pyrrole nitrogens, critical for catalysis.
P86	Terminates short helical segment, role in protein folding.
R132	Salt bridges C1 cofactor ring
Q135	Hydrogen bonds to 193 N and 131 and 193 O, role in stabilising protein fold.
R149	Salt bridges substrate porphobilinogen acetate.
G150	Hydrogen bonds C2 ring acetate.
N151	Hydrogen bonds R11.
T154	Surface residue.
R155	Salt bridges C1 ring propionate and C2 ring acetate.
L169	Lines rear of the active site cleft, towards periphery of the hydrophobic core.
A170	Located on one face of domain 2's third helix, on the wall of the cleft, these small
A172	side chains may be necessary to accommodate the growing pyrrole chain.
R176	At bottom of the cleft, interacts with C2 ring propionate via bridging water molecule, and with Q19 and 199 O.
A195	Lines rear of active site cleft.
Q198	Hydrogen bonds to D106, L83, R131, 104 O, stabilises the bonding network.
G199	Third residue of β-turn, in region of limited space at bottom rear of cleft.
R206	Salt bridges D76, D209, hydrogen bonds to 70 and 75 O, stabilises fold.
E231	Salt bridges R101, hydrogen bonds backbone nitrogens of 197 and 250.
R232	Hydrogen bonds 85 and 86 O, forming stabilising interdomain interactions.
C242	COFACTOR ATTACHMENT RESIDUE, via thioester link.
P245	Packed in hydrophobic cluster of side chains.
G268	Stabilises protein fold.

Table 1.2 The invariant residues of *E. coli* porphobilinogen deaminase and their roles in the crystal structure (adapted from Louie *et al.*, 1996).

1.2.7 Structure of *E. coli* porphobilinogen deaminase

The structure of *E. coli* PBGD in the oxidised form was initially solved by multiple isomorphous replacement to 1.9 Å resolution and subsequently refined to 1.76Å resolution (Louie *et al.*, 1996). The polypeptide chain consists of 313 amino acids folded into three domains of approximately equal size (figure 1.9). Domains 1 (residues 4-99 and 200-217) and two (residues 105-193) have a similar five-stranded β-sheet arrangement with intervening α-helices packed against each face. Domain 3 (residues 222-307) is an open-faced three-stranded antiparallel β-sheet with three α-helices covering one face. Two hinge segments (residues 100-104 and 194-199) connect domains 1 and 2 and domain 1 is connected to domain 3 by a short polypeptide chain (residues 218-221). Analysis of the porphobilinogen deaminase structure reveals that the three domains have relatively few other direct interactions and that the inter-connecting segments can allow substantial flexibility between the domains. This flexibility is thought to be essential for allowing observed conformational changes to occur during the catalytic cycle to accommodate the growing polypyrrole chain (Warren *et al.*, 1995). Interestingly, domains 1 and 2 have a similar topology to the transferrins and periplasmic binding proteins (Louie, 1993). These proteins are known to undergo extensive conformational change on ligand binding. Further evidence for protein flexibility is shown by several amino acids that have poorly defined positions in the X-ray structure. Residues 48-58 form part of a loop that is thought to be located directly in front of the active-site cleft. The poorly resolved electron density for this loop region suggests a highly mobile region which possibly functions as a flexible 'lid'. This 'lid' may function to form additional binding interactions with the substrate and to transiently seal off the active site during catalysis. Further analysis of this loop region also shows adjacent sites of protease sensitivity and susceptibility of nearby lysine residues to modification, suggesting a solvent-accessible position.

1.2.7.1 Detail of the active site

The active site cleft has approximate dimensions of 15 x 13 x 12 Å, occupying nearly half the width and depth of the structure. The floor and ceiling of the cleft are bordered by the carboxyl-termini of the β-sheets, the amino-termini of the α-helices and the connecting segments of domains 1 and 2. The dipyrromethane cofactor is linked to cysteine 242 on a loop from domain three and projects into the catalytic cleft formed at the interface between domains 1 and 2 (figure1.9). The acetate and propionate side chains of the cofactor form extensive salt bridges and hydrogen bonds with conserved residues lining the catalytic cleft. These include a number of invariant arginine residues (11, 131, 132, 149, and 155), that have been investigated by site-directed mutagenesis (Jordan and Woodcock, 1991). The purified mutant enzymes were either unable to bind the cofactor (apoenzyme) or substrate, or a stage of the tetrapolymersation process was affected. Another invariant residue in the active-site cleft, aspartate 84, which interacts with the pyrrole ring NH groups of the cofactor, has been proposed as a key catalytic residue. Substitution of aspartate 84 by site-directed mutagenesis resulted in inactive enzyme, providing supporting evidence for its role. It is postulated that the cofactor contributes to the stability of the holoenzyme when compared with the apoenzyme (Scott *et al.*, 1989) by partially neutralising the electropositive residues lining the active-site cleft. Electron density maps suggest that the cofactor can adopt two different conformations depending on the oxidation state of the enzyme (figure 1.10). The reduced conformation has previously been observed as the minor conformer (20% occupancy) in the oxidised form (Louie *et al.*, 1992) and at 100% occupancy in the selenomethionyl variant crystallised under reducing conditions (Hadener *et al.*, 1993). The interactions of ring C1 remain the same in both states because there is little change in its conformation. However, the second pyrrole ring (C2), in the reduced form of the enzyme, occupies a more internal position in the active site-cleft. In this position the C2 ring interacts with the side-chains of serine-81, aspartate-84, arginine-132 and arginine-176. In the oxidized conformation ring C2 interacts with arginine 11, serine 13, aspartate 84, arginine 149 and arginine 155. In the functional form of the enzyme the cofactor adopts the reduced conformation (Louie *et al.*, 1996).

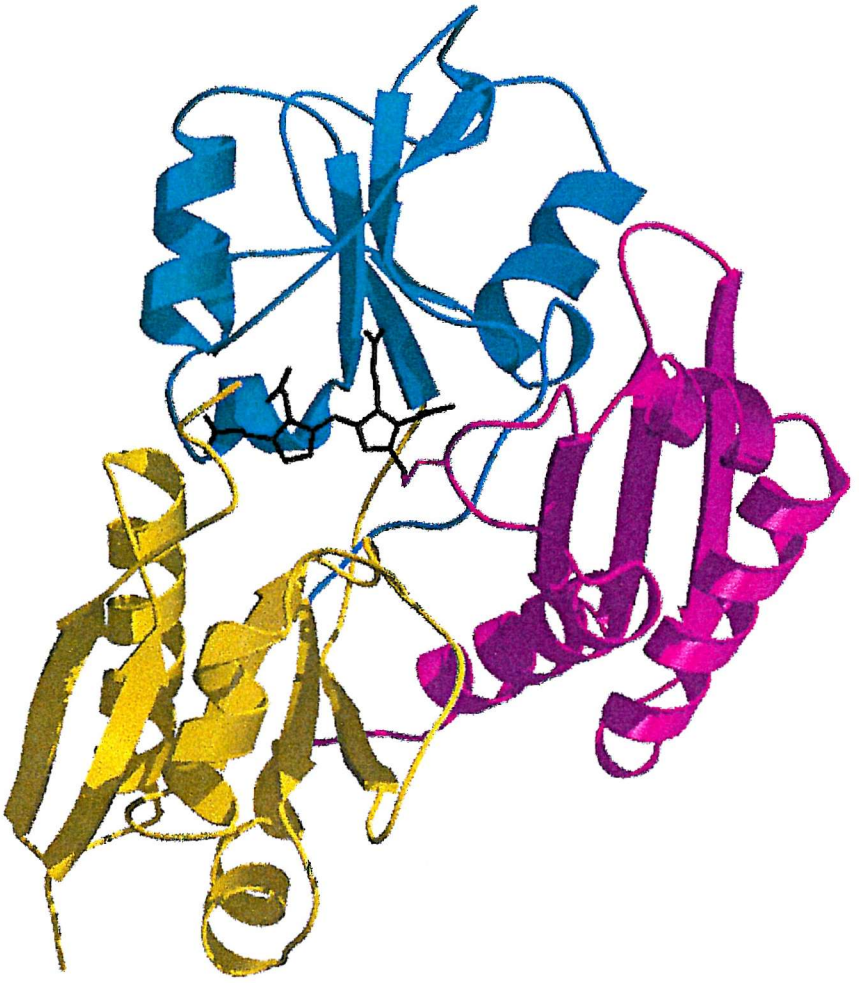


Figure 1.9 The three-dimensional structure of *E. coli* porphobilinogen deaminase. Domain 1 is in yellow, domain 2 in cyan and domain 3 in purple. Domains 1 and 2 have a similar overall topology consisting of five-stranded β-sheet with intervening α-helices packed against each face The dipyrromethane cofactor (shown in black) is covalently linked to cysteine 242 which is part of a loop that projects from domain 3 into the catalytic cleft. This picture was produced using Rasta3D graphics.

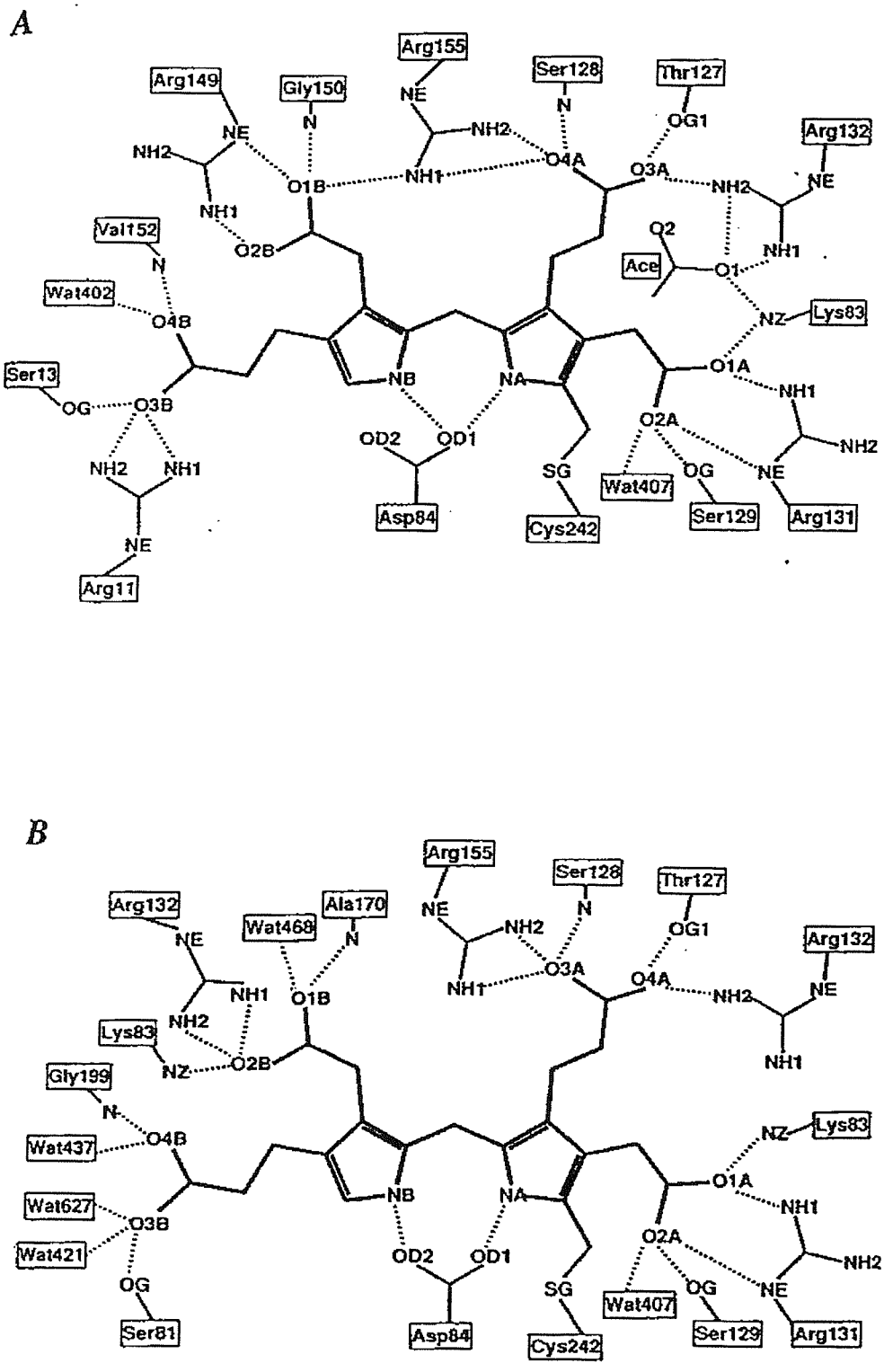


Figure 1.10 The interactions of the dipyrromethane cofactor: A: in the oxidised conformer, and B: in the reduced conformer.

1.3 Disorders of haem biosynthesis

Haem is an essential prosthetic group for haemoglobin, cytochromes, peroxidases and catalases. In humans haem is synthesised from glycine and succinyl-CoA by a pathway involving eight enzymes (figure 1.2). A complete deficiency of any one of these enzymes would be incompatible with aerobic life and hence lethal. However, partial defects in all eight of the enzymes occur and cause disease. Hereditary sideroblastic anaemias result from mutations of the first enzyme of the pathway 5-ALA synthase. Deficiencies in each of the seven remaining enzymes is associated with a particular type of porphyria (table 1.3).

Enzyme	Porphyria	Inheritance	Symptoms
ALA dehydratase	ALA-dehydratase deficiency porphyria	Autosomal recessive	Neurovisceral
Porphobilinogen deaminase	Acute intermittent porphyria	Autosomal dominant	Neurovisceral
Uroporphyrinogen III synthase	Congenital erythropoietic porphyria	Autosomal recessive	Photosensitivity
Uroporphyrinogen decarboxylase	Porphyria cutanea tarda	Variable	Photosensitivity
Uroporphyrinogen decarboxylase	Hepatoerythropoietic porphyria	Autosomal recessive	Photosensitivity
Coproporphyrinogen oxidase	Hereditary coproporphyria	Autosomal dominant	Neurovisceral ± photosensitivity
Protoporphyrinogen oxidase	Variegate porphyria	Autosomal dominant	Neurovisceral ± photosensitivity
Ferrochelatase	Erythropoietic protoporphyria	Autosomal dominant	Photosensitivity

Table 1.3 The porphyrias resulting from deficiencies of enzymes in the haem biosynthetic pathway.

1.3.1 Hereditary sideroblastic anaemias

This disease can result from mutations in the first enzyme of the haem biosynthetic pathway 5-ALA synthase. In humans there are two ALA synthase enzymes, a ubiquitous enzyme and an erythroid specific enzyme. There are no reports of a disorder associated with mutations in the ALAS-1 gene that encodes the ubiquitous enzyme. Such a mutation could be lethal even in the heterozygous state, or the defect may be undetected if induced expression of the normal gene allele can compensate for the mutant allele (May *et al.*, 1990).
The erythroid enzyme, encoded by the ALAS-2 gene, is located on the X-chromosome. Mutations in the ALAS-2 gene can cause X-linked hereditary sideroblastic anaemia where there is usually a decrease in bone marrow 5-ALA synthase activity with hypochromic/microcytic erythrocytes in the blood, leading to anaemia of variable severity. The decreased production of haem and accumulation of iron in erythrocytes differentiates hereditary sideroblastic anaemia from iron-deficiency anaemia. Genetic analysis has shown that a large percentage of the mutations in the ALAS-2 gene are clustered in and around exon 9. This is unsurprising as exon 9 contains the site that binds pyridoxal phosphate, which is an essential cofactor for the enzyme. Approximately thirty percent of patients with hereditary sideroblastic anaemias respond to treatment with the vitamin pyridoxine and 5-ALA synthase activity in the bone marrow is restored by subsequent administration of pyridoxal phosphate.
It is important to note that not all patients with X-linked sideroblastic anaemias have defects in the ALAS-2 gene. Linkage studies have shown that mutations at two other sites on the X-chromosome may cause a form of sideroblastic anaemia.

1.3.2 Porphyrias

Deficiencies in the seven remaining enzymes in the haem biosynthetic pathway can cause diseases collectively known as the porphyrias. The disease symptoms (table 1.3) are attributable not only to the decrease in haem biosynthesis but also as a result of the accumulation of toxic pathway intermediates. The biosynthesis of haem is primarily controlled through a feedback mechanism whereby an

increase in the cellular haem pool level triggers decreased synthesis and activity of 5-ALA synthase, reducing the flux through the pathway. ALA-synthase is the first and rate-controlling enzyme in the pathway. This feedback loop ensures tight regulation of the pathway, preventing the build-up of pathway intermediates. Insufficient amounts of any of these seven enzymes results in the depletion of the cellular haem pool, leading to induction of 5-ALA synthase. However, due to a defective enzyme at some step in the pathway, this induction is not followed by the usual increase in haem biosynthesis, but instead by the accumulation of toxic pathway intermediates. Essentially, there is a breakdown of normal regulatory mechanisms, leading to the over-induction of 5-ALA synthase. A deficiency of enzymes early in the pathway causes the overproduction of the intermediates 5-ALA and porphobilinogen. Lesions later in the pathway leads to the build-up of excess tetrapyrrole intermediates. Thus for each porphyria there is a specific pattern of accumulation of intermediates which are correlated with various pathological symptoms. Porphyrias can be classified, according to observed symptoms, as acute or non-acute, but some have both symptoms. In the acute porphyrias patients suffer from abdominal pain, constipation, vomiting, paralysis and psychiatric disorders. The neurological abnormalities are thought to result from the accumulation of ALA and porphobilinogen. It has been suggested that the elevated levels of ALA may be responsible for the neurological symptoms due to a structural resemblance to the inhibitory neurotransmitter 4-aminobutyric acid (GABA) (Bagust *et al.*, 1985). However, haem deficiency could also compromise the central nervous system. In non-acute porphyrias the predominant symptom is photosensitization which is due to the reaction of accumulated porphyrins in the tissues with molecular oxygen producing damaging free-radicals (Warren *et al.*, 1996). In the mixed porphyrias patients have both acute and non-acute symptoms.

All the enzymes that are defective in the porphyrias have been characterised and their genes located. This information has aided the study of the pathogenesis of the porphyrias at the molecular level, revealing a great deal of genetic heterogeneity (Sassa and Kappas, 2000). The following is a brief outline of the eight types of acute and non-acute porphyrias.

1.3.2.1 The acute porphyrias

There are four types of acute porphyria; ALA dehydratase deficiency porphyria, acute intermittent porphyria (AIP), hereditary coproporphyria and variegate porphyria. ALA dehydratase deficiency porphyria is inherited in an autosomal recessive fashion and results in approximately 1% of the normal enzyme activity. The other three acute porphyrias are autosomal dominant and are characterised by a 50% reduction in the level of active enzyme. More that 80% of those who inherit such a defect remain asymptomatic, making detection and diagnosis problematic. Clinical manifestations are often in the form of acute attacks of neurological dysfunction which may be life threatening. These attacks are usually precipitated by exogenous and endogenous factors including certain drugs, alcohol, fasting and hormones. Many of these factors activate flux through the pathway due to an increased demand for haem. Detection of carriers is therefore very important to provide appropriate counselling for the avoidance of precipitating factors.

ALA dehydratase deficiency porphyria

5-ALA dehydratase deficiency porphyria, or Doss-porphyria, is an autosomal recessive hepatic porphyria. It is a very rare porphyria, and only a few cases have been reported (Sassa and Kappas, 2000). The enzyme 5-ALA dehydratase is normally present in large amounts; consequently a partial deficiency does not produce any clinical symptoms. Biochemical analysis of patients with presenting neuropsychiatric symptoms indicates they have only 1% of the normal enzyme activity remaining. The finding that this reduced activity is not restored to normal by the *in vitro* addition of sulphydryl reagents is used to eliminate the possibility of lead poisoning, which displays overlapping symptoms with this porphyria. Treatment of this disease consists of intravenous haem-arginate and avoidance of exposure to harmful drugs (Kappas *et al.*, 1989).

Acute intermittent porphyria

Mutations in the human gene encoding porphobilinogen deaminase (PBGD) result in the disease acute intermittent porphyria (AIP). This is an autosomal dominant trait that results in activity depletion (usually by 50%) of PBGD enzyme. The disease is characterised by acute attacks, typically accompanied by colicky abdominal pain, vomiting, hypertension, muscular weakness and psychiatric disturbances. Abnormally high levels of the haem precursors ALA and porphobilinogen in the urine accompany clinical manifestations. During acute attacks intravenous haem administration and provision of large amounts of carbohydrates is often beneficial. AIP is discussed in more detail in section 1.3.5

Hereditary coproporphyria

This is an autosomal dominant porphyria caused by mutation of the gene encoding the enzyme copropophyrinogen oxidase, resulting in 50% of the normal enzyme activity. The symptoms of hereditary coproporphyria are similar to AIP, although milder, with only slightly raised amounts of ALA and porphobilinogen (Sassa and Kappas, 2000). Treatment of hereditary coproporphyria is similar to AIP with additional avoidance of sunlight to prevent skin photosensitivity (Kappas *et al.*, 1989).

Variegate porphyria

Variegate porphyria is an autosomal dominant hepatic porphyria due to mutations in protoporphyrinogen oxidase resulting in 50% of the normal enzyme activity. The clinical symptoms and disease management is identical to AIP. In addition, there is photosensitivity due to the accumulation of porphyrins in the skin, which can be minimised by preventing exposure to sunlight (Kappas *et al.*, 1989).

1.3.2.2 The non-acute porphyrias

Congenital erythropoietic porphyria, porphyria cutanea tarda, hepatoerythropoietic porphyria and erythropoietic porphyria comprise the four non-acute porphyrias. These diseases are characterised by burning pain, itching, pigmentation and skin fragility with blisters and scarring.

Congenital erythropoietic porphyria

This is an autosomal recessive erythropoietic porphyria, caused by deficiency of uroporphyrinogen III synthase. Symptoms include marked skin photosensitivity due to overproduction of porphyrins in tissues under the skin. Red blood cells show intense fluorescence due to the accumulation of uroporphyrin I. The age of onset of congenital erythropoietic porphyria is highly variable, but in most cases severe photosensitivity has developed soon after birth. Congenital erythropoietic porphyria can be treated by protecting the patients from sunlight.

Porphyria cutanea tarda and hepatoerythropoietic porphyria

Porphyria cutanea tarda and hepatoerythropoietic porphyria are due to heterozygous and homozygous deficiency, respectively, in the enzyme uroporphyrinogen decarboxylase. About 25% of porphyria cutanea tarda is inherited (familial, or type II) and 75% acquired (sporadic, or type I). Porphyria cutanea tarda, like some of the acute hepatic porphyrias, is latent until the patient has been exposed to precipitating factors including alcohol, hormones and liver disease. The clinical symptoms of porphyria cutanea tarda are characterised by marked skin photosensitivity, due to the oxidation of accumulated porphyrinogens to porphyrins. The disease can be treated by avoiding precipitating factors and, in some cases, by phlebotomy that reduces the over accumulation of iron (Kappas *et al.*, 1989).
Hepatoerythropoietic porphyria is a very rare disease with clinical symptoms similar to those of congenital erythropoietic porphyria. This porphyria can be

treated by avoidance of sunlight and the use of topical sunscreens.(Kappas et al 1989).

Erythropoietic protoporphyria

Erythropoietic protoporphyria is an autosomal dominant disorder caused by decreased activity of the enzyme ferrochelatase. The disease is characterised cutaneous photosensitivity due to a massive accumulation of protoporphyrin in erythrocytes, plasma and stools (Kappas *et al.*; 1989). It can be treated by avoiding sunlight and taking oral administration of β-carotene (Sassa and Kappas, 2000). Fatal liver disease may occur in a proportion of patients and liver transplantation has been used to prevent progression of the disease (Moore, 1993).

1.3.3 Acute intermittent porphyria

This is the most common acute porphyria that results from mutations in the gene encoding PBGD. This gene is transcribed from two promoters to produce ubiquitous and erythroid isoenzymes. The only difference between these is that the erythroid enzyme is 17 amino acids shorter at the amino terminus. Acute intermittent porphyria is classified into two subtypes. The predominant disease is classical AIP where both of the enzymes are deficient, but about 5% of families have variant AIP in which only the ubiquitous enzyme is affected. Approximately 90% of individuals who inherit a PBGD deficiency remain clinically unaffected. AIP is usually latent before puberty and symptoms are more frequently found in females than in males. Clinical manifestations of AIP include abdominal pain and neurological dysfunction. These symptoms occur during acute attacks, which are often precipitated by drugs, alcohol, low caloric intake or infections. In this respect, AIP has been defined as pharmacogenetic in nature, because clinical expression of the underlying defect is significantly influenced by non-genetic factors. The biochemical abnormalities in AIP result from defects in PBGD and consecutive overexpression of the first enzyme of the pathway 5-ALA synthase. These enzyme disturbances lead to the synthesis of

excess haem precursors with large amounts of 5-ALA and porphobilinogen being excreted in the urine. It has been proposed that 5-ALA mimics the inhibitory neurotransmitter GABA, causing a toxic effect in the nervous system. Another theory is that the neurological symptoms are a result of insufficient haem synthesis in the neural tissues. However, the pathogenesis of the neuropathy remains to be elucidated. The development of a transgenic porphobilinogen deaminase deficient mouse provides a system in which these hypotheses can be tested.

1.3.3.1 Diagnosis and treatment of acute intermittent porphyria

Diagnosis relies primarily on biochemical tests, including the estimation of urinary ALA and porphobilinogen and measurement of erythrocyte PBGD activity, which is usually decreased by 50%. However, these methods are limited, as there is an overlap between the normal range of enzyme activity and that of gene carriers. Additionally, in some patients the enzyme deficiency is restricted to the ubiquitous isoenzyme and therefore the measured erythroid PBGD activity is normal, leading to misdiagnosis. Additional biochemical tests consist of measurements of the different porphyrins in the urine and faeces. This helps to distinguish AIP from other acute porphyrias, but is not essential for short-term management of the acute attack.

Upon diagnosis of an acute attack, any precipitating factors, including specific drugs (Gorchein, 1997) must be sought and removed. Treatment consists of intravenous administration of exogenous haem. This deregulates the production of 5-ALA synthase and therefore the concentration of 5-ALA and porphobilinogen are lowered to normal levels. Supportive treatment includes analgesia, antiemetics, rehydration and parenteral nutrition. In women with repeated pre-menstrual attacks, analogues of lutenizing hormone have been used to inhibit ovulation, which can reduce or attenuate attacks.

1.3.3.2 Prevalence of acute intermittent porphyria

The estimated frequency of clinically overt AIP is 1-2 in 100,000, with disease being reported in many races (Elder, 1993). Lapland, Scandinavia, and the

United Kingdom are believed to have the highest incidence of AIP at 5 in 100,000 (Waldenstrom, 1956;Lee and Anvret, 1991). The prevalence is reported to be as high as 210 per 100,000 in psychiatric populations. Taking into account the large percentage of asymptomatic gene carriers, the total penetrance of AIP could be as high as 1-2 in 10,000. Environmental or acquired factors such as drugs, alcohol, or steroid hormones may precipitate attacks. Thus, the diagnosis of AIP heterozygotes is crucial as the primary form of medical management in the avoidance of specific precipitating factors.

1.3.3.3 Detection of AIP mutations

Until the identification of AIP mutations the diagnosis of gene carriers among relatives of AIP patients was based on erythrocyte PBGD activity. DNA methods are much more accurate than conventional biochemical techniques for detecting latent AIP and direct mutation testing has become the method of choice. To date more than 200 mutations have been identified in the PBGD gene, which consists of 15 exons that span 10 kilobases (Brownlie *et al.*, 1994). Approximately 40% of mutations are found in exons 10 and 12. Single-base substitutions accounts for roughly 75% of these mutations and the remaining 25% comprise small insertions or deletions in the PBGD gene. The majority of mutations are private or are only found in a few unrelated families. However, in Sweden almost 50% of AIP families share the same mutation, W198X, while another mutation, R116W, accounts for 30% of AIP cases in Holland. This geographically high frequency of specific mutations has been ascribed to founder effects (Elder, 1998).

1.3.3.4 Classification of AIP mutations

Historically AIP and subsequent AIP mutations were classified according to the CRIM (cross-reacting immunological material) status of the patients. Four major classes of AIP were identified by immunological and activity studies on porphobilinogen deaminase isolated from the erythrocytes of patients. Of the four classes, two are CRIM -ve and two CRIM +ve (Desnick *et al.*, 1985). A CRIM -ve response occurs when the decrease of cross-reacting material is proportional to the decrease in enzyme activity, that is the CRIM/activity ratio is

one. A CRIM +ve response occurs when the decrease of CRIM is not proportional to the decrease in enzyme activity. It has been proposed that those mutations involving substrate-binding groups are CRIM +ve, indicating that the gross protein structure is little affected. Mutations involving essential cofactor binding residues are CRIM -ve highlighting the importance of the dipyrromethane cofactor in maintaining protein stability. Residues important for protein folding or stability are also more likely to be CRIM –ve. The CRIM –ve phenotype represents over 80% of AIP patients. This method of classification is limited due to experimental variability, between laboratories, in the determination of the CRIM status. Additionally, more recently discovered mutations have not been correlated with their respective CRIM status. Despite this, the determination of the CRIM status can play an important part in the understanding of protein structure/function relationships.

1.3.3.5 Structure function relationship of AIP mutations

The three-dimensional structure of *E. coli* deaminase has been defined by X-ray analysis. Comparison of the *E. coli* and human PBGD sequences indicates strong conservation of structurally and important amino acids, and thus has allowed analysis of reported AIP mutations for likely effects on the structure and catalytic properties of the human enzyme. Additionally, a molecular model was constructed for the human enzyme (figure 1.10), based on the *E. coli* enzyme (Wood *et al.*, 1995). This model has been used to show where specific AIP mutations are located in the three-dimensional structure in an attempt to correlate the type of the mutation with clinical severity.

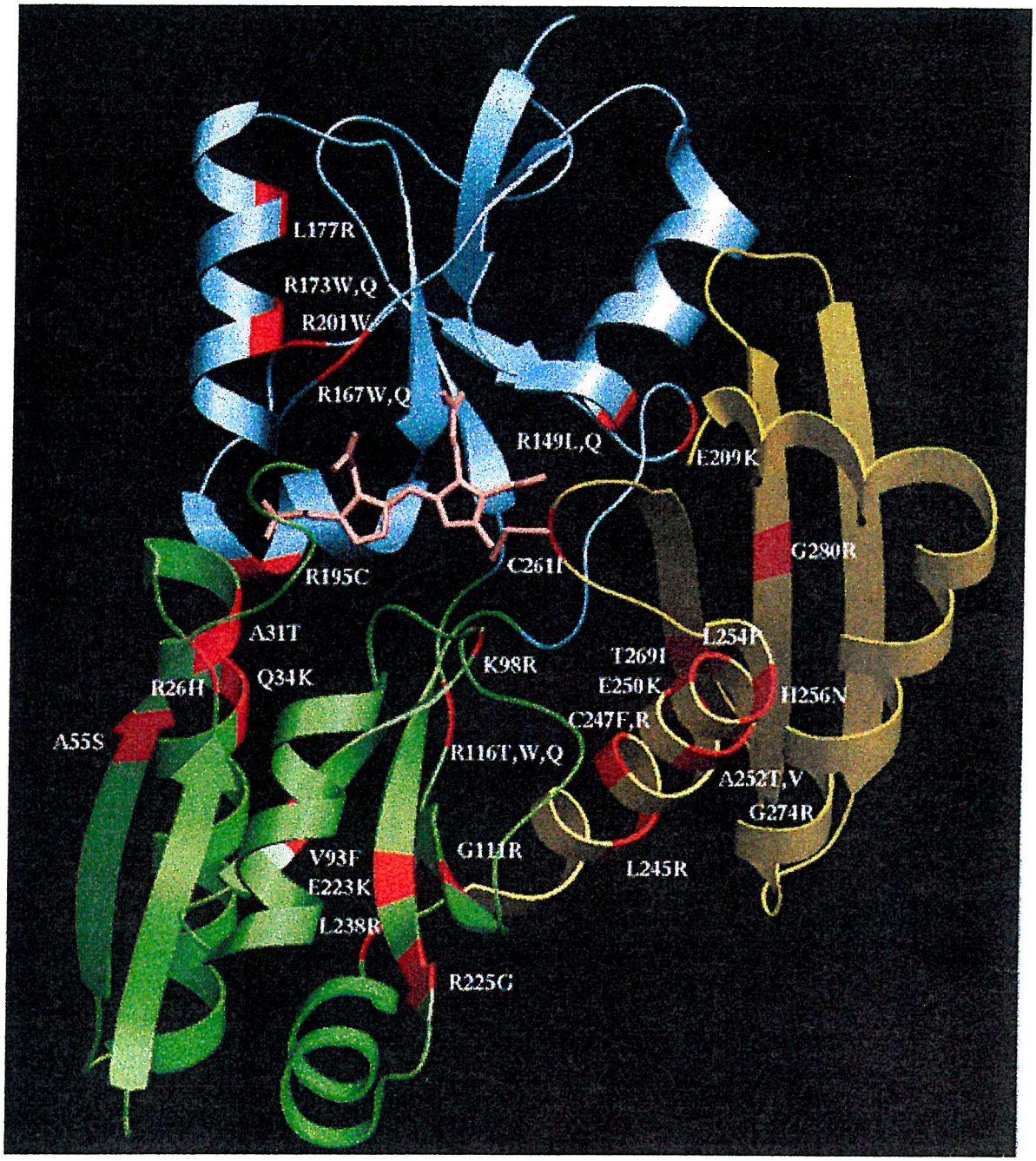


Figure 1.10 The three-dimensional structure of human ubiquitous PBGD modelled on *E. coli* PBGD. Domain 1 is depicted in green, domain 2 in blue, domain 3 in yellow and the cofactor in pink. The mutations shown are acute intermittent porphyria-associated mutations (reproduced from Wood *et al.*, 1995).

1.4 Aims of the Thesis

1. To investigate further the importance of the proposed flexible loop region in *E. coli* PBGD by site-directed mutagenesis of residue glycine 57 to alanine and subsequent characterisation of the mutant protein.
2. To study interdomain flexibility of *E. coli* PBGD by mutation of residue glycine 264 in domain three to cysteine. It is postulated that, under oxidising conditions, the residue cysteine 264 in the mutant would form a disulphide bridge with cysteine 134 in domain 2. Consequently the two domains 2 and 3 would be interlinked rigidly to one another and this could prevent the conformational change observed during catalysis, with a consequent loss of activity.
3. To overexpress, purify and crystallise ubiquitous and erythroid human porphobilinogen deaminase in order to obtain a detailed X-ray crystal structure of the human enzyme. This would allow the particular structural and functional characteristics to be confirmed, including the configuration and possible role of the additional N-terminal amino acids and the location of the additional insert in domain 3 of the ubiquitous human enzyme.

Chapter 2 Materials and Methods

2.1 Materials

All chemicals were obtained from Sigma Chemical Company, with the following exceptions: Tryptone and yeast extract - Difco Laboratories. Agar-LabM. Agarose–Pharmacia. Electrophoresis equipment and BIO-RAD reagent were purchased from BIO-RAD laboratories limited. All column chromatographic equipment was supplied by Pharmacia. Restriction enzymes and the Wizard DNA miniprep kit were obtained from Promega. Sculptor *in vitro* mutagenesis kit, Sequenase DNA sequencing kit, version 2.0, and PD10 columns were purchased from Amersham International plc (now Amersham Bioscience limited). Geneclean *II* kit was from BIO 101 Inc.

2.1.1 Bacterial strains and vectors

The strains of *Escherichia coli* and the vectors utilized are described in tables 2.1 and 2.2 respectively.

Strain	Genotype	Source
TG1	*sup*E, *thi* Δ(*lac*-proAB), *hsd*D5, F'[traD36, proA$^+$; *lac*I^q, *lac*Z M15]	Gibson, 1984
TB1	ara, (lac-proAB), rpsL, φ80, traΔ 36, proAB, lacI^{q}Z M15, hsdR, (r^-_k, m^+_k)	Focus, 1984
DH5α	F-, φ80dlacZΔM15, recA1, endA1, gyrA96, thi-1, hsdR17 (rk^-, mk^+), supE44, deoR, relA1, Δ(lacZYA-argF)U169, λ-	Woodcock, *et al.*, 1989
BL21 (DE3)	F-, ompT, (Ion), $hsdS_B$ (r_B-, mB, an E.coli B strain) with DE3, a λ prophage carrying the T7 RNA polymerase gene.	Studier, *et al.*, 1990
HMS174-DE3	F$^-$ recA1 hsdR($r_{K12}{}^-$ $m_{K12}{}^+$) Rif^R(DE3)	As above

Table 2.1 Bacterial strains utilised for cloning and expression of porphobilinogen deaminase

Plasmid	Properties	Plasmid size (kb)	Source
M13mp19	Cloning vector	7.23	Messsing, 1983
pUC18/19	Cloning vector	2.69	Messing, 1983
pT7-7	Cloning vector	2.47	Yanisch-Perron, 1985
pBM3	pUC18 with a 1.68 kb *Bam*HI/*Sal*I fragment encoding *hemC* (*E. coli* PBGD gene)	4.38	Dr B. Mgbeje, 1989
pG57A	pUC18 with a 1.68 kb *Bam*HI/*Sal*I fragment encoding *hemC* with mutation G57A	4.38	J. Mosley, this thesis
pG264C	pUC18 with a 1.68 kb *Bam*HI/*Sal*I fragment encoding *hemC* with mutation G264C	4.38	J. Mosley, this thesis
pGEM (Bluescript)	Cloning vector	2.87	Promega
pUPBGD	pT7-7 with 1142 base *Bam*HI/*Nde*I fragment encoding human ubiquitous PBGD	3.57	Dr M. Sarwar
pEPBGD	pT7-7 with 1090 base B*am*HI/*Nde*I fragment encoding human erythroid PBGD	3.57	Dr M. Sarwar

Table 2.2 Vectors utilised in cloning of porphobilinogen deaminase

2.1.2 Media and agar plates

Media

All media were autoclaved at a pressure of 15 p.s.i. for 15 minutes and stored at room temperature.

LB (Luria broth) media

For 1 litre: -

Bacto tryptone	10g
Bacto yeast extract	5g
NaCl	5g

2 x TY media

For 1 litre: -

Bacto tryptone	16g
Bacto yeast extract	10g
NaCl	5g

Agar plates

Once poured, all agar plates were stored at 4°C.

LB plates

1L of LB media was mixed with 15g of bacto agar prior to autoclaving (1.5% w/v). This was used to pour approximately 40 plates.

LB + Ampicillin plates

Ampicillin was added to molten LB agar (once it was less then 50°C) to a final concentration of 100μg/mL.

Glucose minimal medium plates

The following reagents were autoclaved separately and cooled before mixing aseptically: -

M9 Salts with 15g bacto agar	1Litre
1M $MgSO_4$	1mL
1M Thiamine HCl	1mL (filter sterilised)
0.1M $CaCl_2$	1mL
20% Glucose	10mL

M9 salts – 10 x

For 1 litre: -

Na_2HPO_4	60g
KH_2PO_4	30g
NH_4Cl	10g
NaCl	5g

H Top Agar

For 1 litre: -

Bacto tryptone	10g
Bacto agar	8g
NaCl	8g

L plates

For 1 litre: -

Bacto tryptone	10g
Bacto yeast extract	5g
Sodium chloride	10g
Bacto agar	15g

2.1.3 Solutions

DNA purification solutions

All the following solutions were autoclaved and stored at 4°C unless otherwise indicated.

TE buffer

10mM Tris/HCl, pH 8.0

1mM EDTA, pH 8.0

STE

0.1M NaCl

10mM Tris/HCl, pH 8.0

1mM EDTA, pH 8.0

Solution I

50mM Glucose

25mM Tris/HCl, pH 8.0

10mM EDTA, pH 8.0

Solution II

This was freshly prepared for each use

0.2M NaOH

1% SDS

Solution III

5M Potassium acetate	60mL
Glacial acetic acid	11.5mL
Distilled water	28.5mL

This was stored at room temperature.

20%(w/v) PEG 6000/2.5M NaCl

For 100mL: -

PEG 6000	20g
NaCl	14.6g

13%(w/v) PEG 8000/1.6 M NaCl

For 100mL: -

PEG 8000	13g
NaCl	9.35g

DNase free pancreatic RNase

Pancreatic RNase was dissolved in 10mM Tris/HCl, pH 8.0, to a final concentration of 10mg/mL and boiled for 15 minutes. The solution was left to cool to room temperature and stored in 50µL aliquots at -20°C.

Wizard Miniprep Solutions

The following were all components of the Wizard miniprep kit supplied by Promega: -

Cell Resuspension Solution

50mM Tris/HCl, pH 7.5

10mM EDTA

100µg/mL RNase A

Cell Lysis Solution

0.2M NaOH

1% SDS

Cell Neutralisation Solution

1.32M Potassium acetate, pH 4.8

Wash Solution (stock)

20mL of 200mM NaCl

20mM Tris/HCl, pH 7.5

5mM EDTA

The above stock solution was diluted with 95% ethanol to give a final ethanol concentration of 55%.

DNA electrophoresis solutions

50 x TAE

242g Trizma base

37.2g Na_2EDTA

57.1mL Glacial acetic acid

Loading buffer

0.05g Bromophenol blue

0.05g Xylene cyanol FF

6mL Glycerol

The above were made up to 20mL with distilled H_2O and stored at 4°C.

Ethidium bromide solution

A stock of 10mg/mL was made in sterile distilled water. It was used at a final concentration of 0.5µg/mL

Acrylamide stock

38g Acrylamide

2g *N,N*-Methylenebisacrylamide

Distilled water was added to a final volume of 100mL and the solution was stirred gently for 30 minutes. The solution was then filtered through a Whatman 0.2µm nylon membrane filter and stored at 4°C in the dark.

10 x TBE

For 1 litre: -

108g Trizma base

27.5g Boric acid

20mL 0.5M EDTA, pH8.0

Sequencing gel mix

210g Urea

75mL Acrylamide stock (see above)

50mL 10x TBE buffer

The above solutions were made up to 500mL with distilled water and filtered through a 0.45µm Millipore filter. The solution was stored at 4°C in the dark. Prior to use 70µL of TEMED and 300µL of 10% ammonium persulphate were added to 50mL of the sequencing gel mix for the main gel or to 20mL of the sequencing gel mix for the plug.

DNA sequencing solutions

5 x reaction buffer

200mM Tris/HCl pH 7.5

100mM $MgCl_2$

250mM NaCl

Labelling mix

dGTP, dCTP and dITP all at 7.5µM

Dideoxynucleotide termination mix

80µM of dGTP, dCTP, dATP and dTTP

8µM of one ddNTP

50mM NaCl

Stop solution

95% Formamide

20mM EDTA pH 8.0

0.05% Bromophenol blue

0.5% Xylene blue

Protein analysis and electrophoresis solutions

Modified Ehrlich's reagent

1g 4-Dimethylaminobenzaldehyde

42mL Glacial acetic acid

8mL 70% Perchloric acid

This solution was stored in the dark at 4°C. The shelf life was a maximum of one week.

Acrylamide stock

Same as acrylamide stock used for DNA sequencing gels (see DNA electrophoresis solutions)

10 x SDS-PAGE running buffer

For 1 litre: -

144g Glycine

30.3g Trizma base

10g Sodium dodecyl sulphate

The components were dissolved in distilled water and stored at room temperature. For use, 100mL was mixed with 900mL of distilled water.

The same recipe was used to make native-PAGE running buffer, except SDS was omitted.

5 x Loading buffer (for SDS-PAGE)

For 200mL: -

6.1g Trizma base

0.2g SDS

0.1g Bromophenol blue

100mL 50% Glycerol

Before use 982.5µL was mixed with 17.5µL of mercaptoethanol.

For native-PAGE the SDS and mercaptoethanol were left out.

Destain for SDS-PAGE and Native-PAGE

100mL Glacial acetic acid

500mL Methanol

400mL Distilled H_2O

1.5g of Coomassie brilliant blue was added to make stain solution.

2.2 Molecular Biology Methods

This section includes the various techniques for DNA purification and subsequent manipulation to produce the wild type and mutant clones specifying the *E.coli* and human PBGD proteins.

2.2.1 Large scale dsDNA purification by alkaline lysis and polyethylene glycol

1. The bacterial cell pellet from a 500mL culture was washed with 100mL ice-cold STE.
2. The cell suspension was centrifuged (3000g) for 15 minutes at 4°C.
3. The washed pellet was resuspended in 10mL of solution I by vortexing and transferred to a 50mL centrifuge tube. To this resuspension 1mL of a freshly prepared solution of lysozyme (10mg/mL in 10mM Tris/HCl, pH 8.0) was added to the pellet suspension and the mixture incubated at room temperature for 10 minutes
4. Solution II (20mL) was added and gently mixed by inverting. The mixture was

stored at room temperature for 10 minutes.

5. Addition of 15mL of ice-cold solution III was followed by gentle inversion of the tube several times.

6. This lysate was stored on ice for 10 minutes allowing a white flocculent precipitate to form before centrifuging (5500g for 25 minutes at 4°C) and transferring the supernatant to a clean tube.

7. The supernatant was filtered through four layers of Myracloth, 0.6 volume of isopropanol was added and the mixture stored at room temperature for 10 minutes.

8. The nucleic acid was recovered by centrifugation (3000g for 15minutes at 4°C). The supernatant was discarded and the tube was inverted to allow any remaining supernatant to drain away.

9. The pellet was washed with 70% ethanol at room temperature and the tube was then inverted on tissue to dry the pellet.

10. The pellet was dissolved in 3mL of TE (pH 8.0).

11. To the nucleic acid solution from step 10, 3mL of ice-cold 5M LiCl was added. The solution was mixed well and centrifuged at 11,500g for 10 minutes at 4°C.

12. The supernatant was transferred to a clean 50mL centrifuge tube and an equal volume of *iso*-propanol was added and, after mixing, the nucleic acid was recovered by centrifugation as in 1, but at room temperature.

13. The supernatant was drained away, the pellet rinsed with 70% ethanol and any remaining ethanol removed by aspiration at room temperature.

14. The pellet was dissolved in 500µl of TE, pH8.0, containing DNase-free pancreatic RNase (20µg/mL), transferred to a microfuge tube and stored at room temperature for 30 minutes. (The RNase was made up by diluting 1.5µl of 10mg/mL stock into 500µl of TE).

15. To the resulting solution, 500µl of 1.6M NaCl containing 13% (w/v) PEG 8000 was added and the DNA was recovered by centrifugation at 12,000g for 5 minutes at 4 °C in a microfuge.

16. The supernatant was removed by aspiration and the pellet was dissolved in 400 µl of TE, pH 8.0. The solution was extracted once with phenol, once with

phenol:chloroform (1:1 v/v) and once with chloroform.

17. The aqueous solution was transferred to a fresh microfuge tube, mixed with 100µl of 10M ammonium acetate and 2 volumes (~1mL) of ethanol and the solution stored at room temperature for 10 minutes. The precipitated DNA was recovered by centrifugation at 12,000g for 5 minutes at 4°C in a microfuge.

18. The supernatant was removed by aspiration and 200µl of 70% ethanol at 4°C was added. The tube was vortexed and centrifuged at 12,000g for 2 minutes at 4° C in a microfuge.

19. The supernatant was removed and the tube was left open until all the ethanol had evaporated.

20. The pellet was dissolved in 500µl of TE (pH 8.0) and the concentration of the plasmid DNA was measured at 260nm (dissolved in 0.1-1.0mL of buffer depending on the pellet size).

1 OD unit at 280nm ≡ 50µg dsDNA

2.2.2 Small scale dsDNA purification by alkaline lysis

1. An overnight culture of bacteria, harboring the recombinant plasmid of interest, was grown by incubating a single colony in 10mL of LB, supplemented with the necessary antibiotic solution.

2. The cells were harvested by centrifuging at 10,000g for 15 minutes.

3. The supernatant was discarded and the pellet was resuspended in 1mL of ice-cold STE buffer and transferred to a microfuge tube.

4. The resuspended material was centrifuged at 12000g for 3 minutes and the supernatant was removed by aspiration.

5. Ice cold solution I (200µl) was added and the tube contents vortexed.

6. Next, 400µl of solution II (freshly prepared) was added and the tube was rapidly inverted 5 times, ensuring that the entire surface of the tube came into contact with solution II. The tube was then stored on ice.

7. Ice cold solution III (300µl) was added, before vortexing in an inverted position for 10 seconds (to disperse solution III through the bacterial lysate). Again, the tube was stored on ice for 3-5 minutes, before centrifuging at 12,000g for 5

minutes at 4°C in a microfuge. The supernatant was transferred to a fresh tube.

8. An equal volume of phenol:chloroform was added (900µl) and mixed by vortexing. After microfuging for 2 mins at 4°C the supernatant was transferred to a fresh tube.

9. The dsDNA was precipitated by adding 900µl (100% v/v) isopropanol and vortexing, before leaving the tube to stand at room temperature for 2 mins.

10. The tube was centrifuged again for 5 minutes at 4°C, the supernatant was discarded, centrifuged and any remaining supernatant was removed by aspiration.

11. The pellet was washed with 1mL of 70% ethanol (ice cold). The mixture was centrifuged at 4°C, the supernatant removed by aspiration and the pellet left to dry in air for 10 minutes.

12. Finally, the dried pellet was resuspended in 50µl of TE pH 8.0 containing DNase free RNase, vortexed briefly and left at room temperature for at least 30 minutes to allow digestion to occur. The final mixture was stored at –20°C.

2.2.3 Small scale purification of plasmid DNA using Wizard Plus SV Minipreps System

Purification of dsDNA

1. 5mL of a bacterial culture was pelleted by centrifuging at 10,000g for 10 minutes.

2. After discarding the supernatant the pellet was resuspended in 250µL of Cell Resuspension Solution by vortexing.

3. Next 250µL of Cell Lysis Solution was added and mixed with the cell resuspension by inverting it until it became clear.

4. 10µL of Alkaline Protease Solution was added and mixed by inverting the tube four times prior to incubating the tube at room temperature for 5 minutes

5. The lysed bacterial culture was neutralized by the addition of 350µL of Neutralization Solution, again mixing by inverting the tube several times.

6. The bacterial lysate was centrifuged at 14,000g for 10 minutes in a microfuge at room temperature.

7. The supernatant was loaded onto a Wizard Plus Minipreps spin column using a

5mL syringe. The whole syringe was then detached, the plunger removed, before re-attaching the barrel to the spin column.

8. 750μL of New Wash solution was pipetted into the syringe barrel and pushed through the spin column by re-attachment and use of the plunger. This was repeated with 250μL of Wash Solution.

9. The Wizard Miniprep Spin Column was transferred to a 2mL collection tube and centrifuged at 14,000g for 1 minute to remove any remaining Wash Solution.

10. The Wizard Miniprep Spin Column was transferred to a sterile microfuge tube and the plasmid DNA eluted by the addition of 50-100μL sterile water to the spin column.

11. The DNA solution was collected in the microfuge tube by centrifuging at 14,000g for 1 minute.

Purification of ssDNA

The same protocol was used as for dsDNA except a separate resin, consisting of 1g diatomaceous earth suspended in 50mL of 7M guanadinium-HCl, was used. 1mL of this ssDNA-binding resin was mixed with the supernatant before loading onto the miniprep column in step 7 above.

2.2.4 Oligonucleotide synthesis and purification

Once the particular oligonucleotide had been designed, it was synthesized on a 40nM scale using phosphoroamidite chemistry on an Applied Biosystems 381A DNA synthesizer. After 'trityl-on' synthesis, the oligonucleotide was cleaved from the support by overnight incubation at 55°C in 2mL of concentrated ammonia. The purification was carried out using oligonucleotide purification cartridges (from Applied Biosystems or Cruachem). Firstly, the oligonucleotide-containing solution was separated from the glass particles by spinning the mixture briefly in a microcentrifuge and transferring the supernatant to a fresh tube. The OPC cartridge was prepared by flushing with 5mL of HPLC grade acetonitrile using a 5mL syringe. This was followed by 5mL of 2M triethylamine acetate, keeping the flow rate at 1-2 drops per second. Analytical grade water (1mL) was

added to the deprotected oligonucleotide solution and this was loaded onto the OPC at the same rate. The eluted fraction was collected and passed through the OPC again, allowing further oligonucleotide binding to the column. The OPC was flushed three times with 5mL of dilute ammonium hydroxide solution (1:10), followed by two washes with 5mL of analytical grade (AnalaR) water. The bound oligonucleotide was detritylated by washing it twice with 5mL of 2% TFA solution. After washing the oligonucleotide with 10mL of AnalaR water, it was eluted using 20% acetonitrile (3mL). The eluted fractions were dispensed equally into 6 microcentrifuge tubes and dried using a Speedvac. The resultant pellet was dissolved in sterile water and the absorbance measured at 260nm to calculate the oligonucleotide concentration.

2.2.5 Oligonucleotide annealing

Annealing of two complementary oligonucleotides was done by mixing the following in a microfuge tube to a total volume of 20μL: -

100 pmoles Oligonucleotide 1

100 pmoles Oligonucleotide 2

2μL Ligation buffer - (200mM Tris/HCl, pH 7.5, containing 100mM $MgCl_2$, but no ATP)

The total volume was made up to 20μL with sterile H_2O The mixture was heated at 90°C for 2 minutes and then allowed to cool slowly to room temperature, before storing at -20°C.

2.2.6 Oligonucleotide phosphorylation

100 pmoles of oligonucleotide solution

5μL 10 × T4 kinase buffer (700mM Tris/HCl, pH7.6, containing 100mM $MgCl_2$, and 50mM DTT)

5μL 10mM ATP

2μL T4 Kinase (10 units)

The above mixture was made up to a total volume of 50µL with sterile H_2O and incubated at 37°C for 1 hour. Finally it was heated to 70°C for 10 minutes before either using the phosphorylated oligonucleotide immediately or storing it at -20°C.

2.2.7 Agarose gel electrophoresis

Agarose gels were prepared by the addition of the required amount of powdered agarose to 50mL of 1x TAE (depending on the percentage of the gel required). The slurry was heated in a microwave until the agarose dissolved, then left to cool to below 60°C. Next, 5µl of a 10mg/mL ethidium bromide solution was added and mixed before pouring the molten agarose into a 100 x 80 mm gel-casting tray. After the gel had set, the comb was removed and the gel mounted in the electrophoresis tank containing 1x TAE buffer. The DNA samples (including DNA standards) were mixed with agarose gel loading buffer and loaded into the submerged wells. Electrophoresis was carried out at 100 volts for approx. 1-1.5 hours. The DNA was visualised by use of a UV-transluminator and the size of DNA fragments estimated by comparison to standards (table 2.3).

Lambda DNA restricted with *Eco*RI/*Hin*dIII kb	**PhiX174 DNA restricted with *Hae*III** kb
21.226	1.353
5.1482	1.078
4.973	0.872
4.268	0.603
3.530	0.310
2.027	0.281
1.904	0.271
1.584	0.234
1.375	0.194
0.947	0.118
0.831	0.072
0.564	
0.125	

Table 2.3 DNA standards used in agarose gel electrophoresis

Photographs were taken using a Polaroid camera fitted with a Kodak Wratten filter 23A filter and a Polaroid type 667 film at an aperture of f 4.5 with an exposure time between 30-60 seconds.

2.2.8 Extraction of DNA from an agarose gel by Geneclean™

DNA fragments required for cloning were excised from the agarose gel, three volumes of 6M NaI solution was added and the mixture was heated at 45-55°C until the agarose melted. Next, the appropriate volume of Glassmilk (a component of the Geneclean*II* kit containing a silica matrix suspended in water) to bind the DNA was added and the suspension was vortexed every 1-2 minutes, whilst being incubated on ice for 10 minutes. The mixture was centrifuged (12,000g for 5 seconds) and the supernatant discarded leaving the DNA bound to the silica matrix. The pellet was centrifuged again and any remaining NaI removed by aspiration. The pellet was washed with 700µl of ice-cold New Wash™ solution (a mixture of NaCl, ethanol, water, EDTA, Tris/HCl pH 7.0-8.5) resuspended and then centrifuged for 5 seconds (12,000g). This wash was repeated 3 times, each time discarding the supernatant. An equal volume of sterile water (as Glassmilk) was added and the mixture incubated at 45-55°C for 3 minutes in order to elute the bound DNA. Finally the mixture was centrifuged (30 seconds at 12,000g) and the DNA supernatant pipetted into a sterile Eppendorf tube.

2.2.9 Digestion of dsDNA by restriction enzymes

All restriction enzymes were purchased from Promega and used with one of the four supplied enzyme buffers A, B, C or D depending on which one provided conditions for optimal activity.

A general 10µL restriction is as follows: -

1µl DNA (≡ 1µg)

1µl Restriction enzyme

1µl 10x Enzyme buffer (A, B, C or D)

7µl Sterile water

This mixture was incubated at 37°C for 2 hours. Digested DNA was analyzed on an agarose gel and the required fragment(s) identified by comparison with DNA standards (table 2.3).

2.2.10 Ethanol precipitation of dsDNA

A quick method was used to concentrate a DNA sample after a restriction digest. The DNA solution was made up to100µl with sterile water, 250 µl of ethanol and 10µl of 3M sodium acetate were added before precipitating the DNA using a dry ice-ethanol bath for 30 minutes. The DNA was pelleted (12,000g for 10 minutes at 4°C) and the supernatant removed. After drying in air, the pellet was resuspended in sterile water or an appropriate buffer for a further restriction digest.

2.2.11 Ligation and cloning of gene fragments

The optimum ligation conditions were achieved by varying the molar ratios of vector DNA: insert DNA by 1:1, 1:2, and 1:3. In a typical reaction 1µg of vector DNA was mixed with an equimolar amount of insert DNA in an Eppendorf tube containing 1µL of 10x ligase buffer (300mM Tris/HCl, pH 7.8 containing 100mM $MgCl_2$, 100mM DTT and 10mM ATP) and 3 units of T4 DNA ligase. The reaction mixture was made up to 10µL before incubating at 16°C overnight for sticky-end ligations or at 4°C overnight for blunt end ligations.

2.2.12 Transformation of *Escherichia coli*

LB media (10 mL) was inoculated with 100µl of overnight culture and grown for 2.5 hours at 37°C to an O.D. of 0.3. The culture was centrifuged at 3000g for 10 minutes, the supernatant discarded and the pellet resuspended in 10mL of ice cold $CaCl_2$. This was stored on ice for 40 minutes before centrifuging (10 minutes at 4000g) and resuspending the pellet in 2mL of ice cold $CaCl_2$.

In a prechilled Eppendorf tube, 200µl of competent cells were gently mixed with 0.1µg of plasmid DNA using a pipette. After leaving the cells on ice for 5 minutes, they were heat shocked for 2 minutes at 42°C, then 0.8mL of LB media was added and the cell suspension was incubated for 1 hour at 37°C. Finally, 200 µl of this transformation mixture was spread onto LB+A plates, which were incubated at 37°C overnight.

2.2.13 Alpha complementation technique

(Namba *et al.*, 1991).

This is used for selection of recombinant vectors and is based on the insertional inactivation of the *lacZ* gene (Ullmann *et al.*, 1967). To an LB plate containing the appropriate antibiotics, 40µL of a solution of X-gal (20mg/mL in dimethlyformamide) and 4 µL of a solution of IPTG (200mg/mL) were spread onto the surface. The plate was incubated for 5 minutes at room temperature, allowing the solutions to be absorbed into the agar. Transformed cells (200µL), containing the recombinant plasmid under study, were spread onto the surface of the plate before incubating at 37°C overnight in an inverted position. Colonies harboring the recombinant plasmid are selected by their white colour, whereas colonies containing non-recombinant plasmids are blue.

2.2.14 Sequencing of ssDNA and dsDNA

The Sequenase^{R}, version 2.0 kit, supplied by United States Biochemical Corporation, was used for sequencing purified DNA

The sequencing reaction

A single annealing step was carried out for each set of four reactions (C, T, A, G). The DNA (diluted if necessary to a final volume of 7µl), 5x reaction buffer (2µl) and primer (1µl) were mixed and incubated at 65°C for 2 minutes. The mixture was allowed to cool to room temperature before placing on ice.

The labelling reaction was carried out by addition of the following to the annealing mix:

DTT 0.1M	1µl
Diluted labelling mix	2µl
α-^{35}S dATP	0.5µl (0.5µCi)
Diluted Sequenase™	2µl (1.5U)

The reaction tubes were mixed thoroughly and incubated at room temperature for 5 minutes. The termination reactions were carried out by transferring 3.5 µl of the labelling mix to four tubes containing the appropriate dideoxynucleotide (ddNTP) termination mix. Each tube was mixed and incubated at 37°C for 5 minutes before the addition of 'stop' solution. The final solutions could then be stored at -20°C for up to a week.

Preparation and electrophoresis of the DNA sequencing gel

The BIO-RAD Sequi-gen^{R} system was used for DNA sequencing. This consists of two glass plates, one of which is bonded to the upper buffer chamber, clamped together and separated by 4mm spacers. The plate bonded to the buffer chamber was siliconized using 'Sigmacote' to prevent the gel sticking on separation of the two plates. All components were washed thoroughly and cleaned with ethanol prior to use. The bottom of the plate assembly was sealed using 20mL of gel solution in a casting tray, then the main gel was cast using 50mL of gel mix. After the main gel had set, the gel plate assembly was placed in the lower buffer tank before filling the lower and upper buffer tanks with 350mL and 500mL of 1x TBE respectively. The sequencing gel was pre-run at 1800V until the temperature of the apparatus reached 50°C. The comb was then inverted and the wells were cleaned immediately before loading. The samples were loaded in sets of four in the order C, T, A, G. The gel was electrophoresed at 1800V for the required length of time. After electrophoresis, the buffer was carefully poured from the tanks and the plates were carefully separated, the gel adhering to the non-siliconised plate.

The gel was immediately soaked in 12% methanol/10% acetic acid solution for 15 minutes to remove the urea. The gel was then transferred to 3MM paper and covered with cling-film before drying under vacuum at 80°C for 2 hours on a

BIO-RAD model 583 gel dryer. After drying, the cling-film was removed and the gel was placed in an autoradiography cassette in direct contact with an X-ray film (Hyperfilm β-max, Amersham). After suitable exposure (usually 24 hours), the gel was developed according to the manufacturer's instructions.

2.2.15 Site-directed mutagenesis using the Eckstein method

The Sculptor™ *in vitro* mutagenesis kit from Amersham was used to produce the site-specific mutants of the *hemC* gene. It is based on the phosphorothioate technique of Eckstein's group (Nakamaye and Eckstein, 1986), which has been improved to make it quicker and more efficient.

The principle of the Amersham™ technique is illustrated in figure 2.1. The *hemC* gene cloned into M13 DNA is the template for the mutagenesis reaction. An oligonucleotide with the desired mutation is annealed to the template DNA. The annealed primer is extended in the presence of dCTPαS using T7 DNA polymerase, with T4 DNA ligase, to produce a closed circular heteroduplex. The extension reaction thus incorporates dCMPαS into the mutant strand. The base analogue dCTPαS has a sulphur atom substituting the oxygen at the α-phosphate. This allows the selective destruction of the non-mutant strand leading to a high yield of the mutant strand. The restriction enzyme *Nci*I is used to digest the heteroduplex DNA. This enzyme cannot cleave the phosphorothioate bonds in the mutant strand, so that DNA is preferentially nicked in the non-mutant strand. Exonuclease III is then used to digest most of the nicked non-mutant strand, leaving a fragment of approximately 800 bases which acts as a primer for the repolymerisation. A double-stranded homoduplex DNA is generated using DNA polymerase I and T4 DNA ligase. The resultant DNA should possess the mutation in both strands. The homoduplex is then transformed into a suitable *E. coli* host strain and the plaques are sequenced in order to obtain a mutant clone.

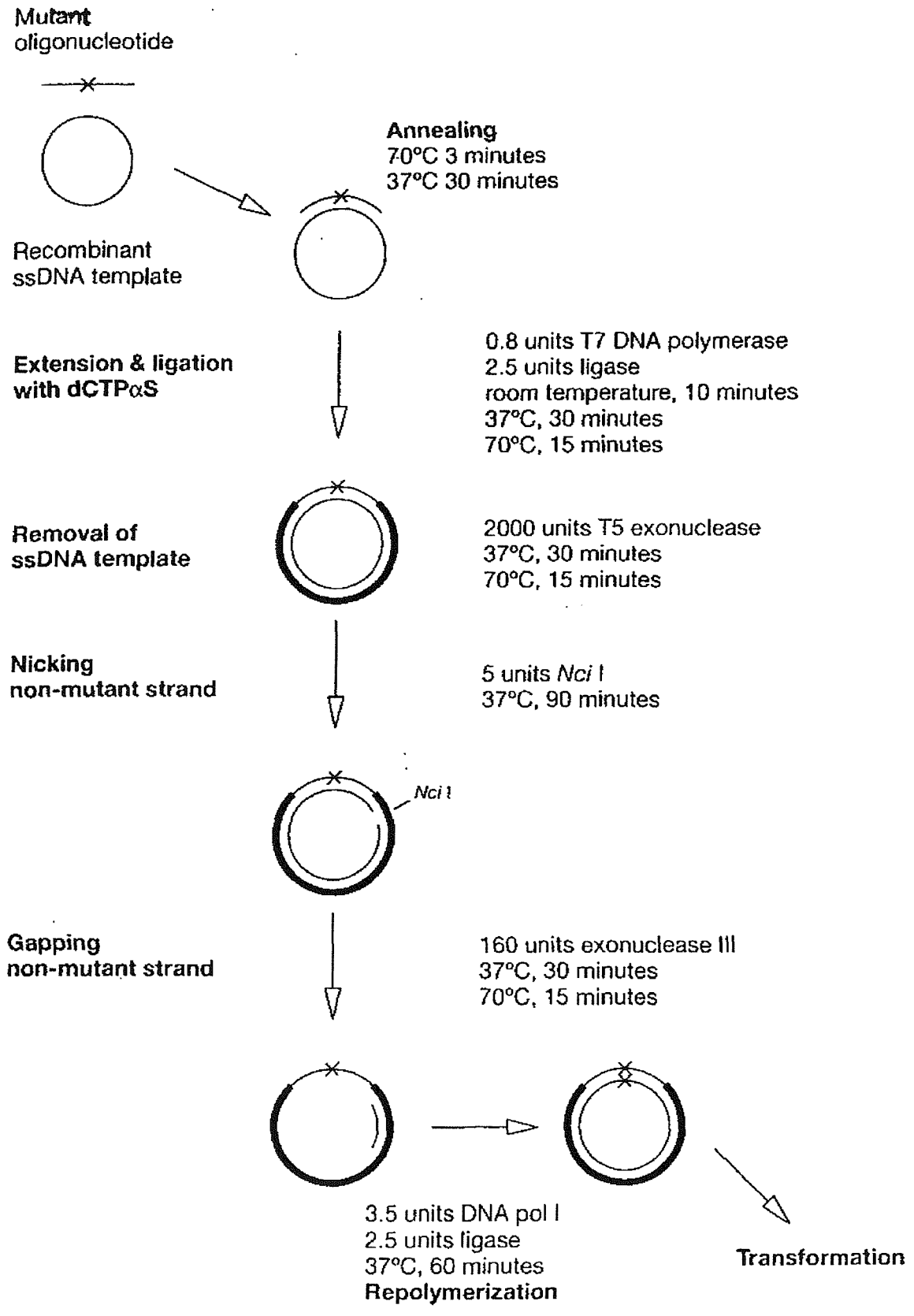


Figure 2.1 Site-directed mutagenesis protocol of the Sculptor™ *in vitro* mutagenesis kit.

2.2.15.1 The mutagenesis protocol

Annealing the mutant oligonucleotide

In a microfuge tube the following were mixed: -

ssDNA template (1μg/μl)	2μl
Phosphorylated oligonucleotide (1.6pmol/μl)	1μl
Buffer A	1μl
Water	5μl

The mixture was heated at 70°C for 3 minutes, then 37°C for 30 minutes before placing on ice.

Extension and ligation

To the annealing reaction, the following were added: -

dNTP mix A	10μl
T4 DNA ligase	1μL (2.5 units)
T7 DNA polymerase	1μL (0.8 units)

After thorough mixing, the tube was left at room temperature for 10 minutes, followed by incubation at 37°C for 30 minutes, with a final inactivation step at 70°C for 15 minutes.

Removal of single-stranded DNA

To the extension reaction the following were added

Buffer B	50μl
T5 Exonuclease	2μl (2000 units)

The mix was incubated at 37°C for 30 minutes, before heat inactivation as in the previous step.

Nicking the non-mutant strand

To the above the following were added: -

Buffer C	5μl
*Nci*I	1μl (5 units)

After mixing, the solution was incubated at 37°C for 90 minutes

Digestion of the non-mutant strand

To the nicking reaction the following were added: -

Buffer D	20µl
Exonuclease III	1µl (160 units)

The mixture was incubated at 37°C for 30 minutes, followed by heat inactivation (as before).

Repolymerisation of the gapped DNA

To the digestion the following were added: -

dNTP mix B	20µl
T4 DNA ligase	1µl (2.5 units)
DNA polymerase I	1µl (3.5 units)

The above were mixed and incubated for 1 hour at 37°C

Analysis of the mutagenic reaction products

Samples were taken at each stage of the mutagenic reaction and subsequently electrophoresed on a 1 % agarose gel. This allows the different DNA species (e.g. ssDNA, circular dsDNA etc.) to be visualised, in order to confirm the success of the mutagenic process.

2.2.16 Expression Cassette PCR

In standard PCR two short oligonucleotides are annealed to the DNA template, one to each strand of the double helix. These oligonucleotides, known as the forward and reverse primers, delimit the region that will be amplified. With this method, the amplified fragments are an exact copy of the original target sequence. The expression-cassette polymerase chain reaction method has distinct advantages in comparison to standard PCR. Expression-cassette polymerase chain reaction facilitates the production of overexpression plasmids by 'effecting site-specific replacement of the 5' and 3' ends of the gene with expression sequences derived from synthetic oligonucleotides'. Figure 2.2 shows a

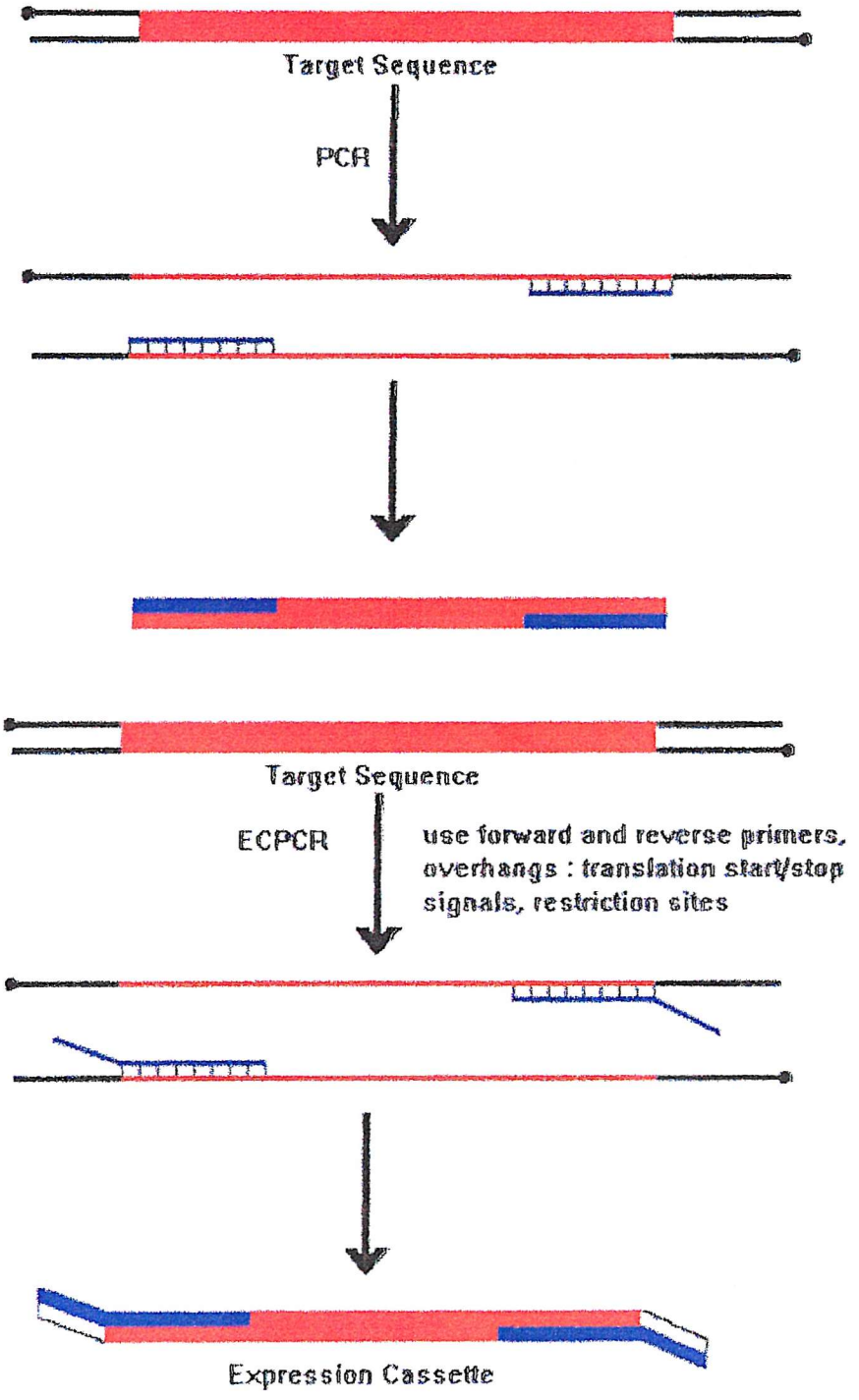


Figure 2.2 Comparison of standard PCR with ECPCR. In standard PCR the amplified fragments are an exact copy of the original target sequence. ECPCR allows the incorporation of additional sequences, required for transcription and translation, into the amplified fragments.

comparison of ECPCR with standard PCR. In ECPCR, sequences required for protein translation and restriction endonuclease digestion are incorporated into the primers. Consequently, the ECPCR of the target DNA results in the synthesis of an expression cassette bearing all the necessary information for cloning into a suitable vector and translation in *E. coli* (MacFerrin *et al.*, 1990).

2.2.16.1 Expression cassette PCR protocol

A pair of forward (sense, or upstream) and reverse (antisense, or downstream) primers were designed, synthesised and purified.
Then the following components were mixed in a small Eppendorf: -

Template DNA	100 to 1000ng, depending on the size of the template
Forward primer	30 to 40ng, depending on the size of the primer
Reverse primer	30 to 40ng, depending on the size of the primer
10 × amplification buffer	4.0µL
100mM $MgSO_4$	1.0µL
dNTPs (2.5mM each)	1.5µL
Total volume	38.0 µL

Two control reactions were carried out in parallel, one with no DNA template, and the other without any primers. The products were stored at 4°C prior to analysis.

PCR cycling parameters

Denaturation 95°C for 5 minutes
Addition of 2µL diluted Vent™ polymerase (0.5µL diluted to 2µL with enzyme buffer)

94°C / 1 minute
60°C / 1 minute
72°C / 1 minute
} 30 cycles

2.3 Protein Chemistry methods

2.3.1 Determination of protein concentration

This was carried out using BIO-RAD reagent, which is based on the Bradford assay (Bradford, 1976). The protein binds to the dye Coomassie brilliant blue G-250 in a concentration dependant manner. This assay was routinely used for protein samples between 200 and 1,400 μg/mL, a typical reaction is as follows: -

10μl Protein sample

790μl Distilled water

200μl Biorad reagent

After incubation at room temperature for 5 minutes, the absorbance was measured at 595 nm and the protein concentration was calculated using the conversion: - 0.1 OD = 1.95 μg protein.

Alternatively, for a pure sample of *E. coli* or human PBGD the absorption was measured at 280nm and the concentration calculated using the molar extinction co-efficient value of 7898 $M^{-1}L^{-1}$.

2.3.2. Determination of PBGD activity

Enzyme solution (50μl) of known concentration was added to 350μl of 20mM Tris/HCl assay buffer, pH 8.2, and prewarmed for 5 minutes at 37°C. 50μl of 2mM PBG solution (similarly prewarmed) was added to the enzyme mixture and incubation was continued at 37°C for 5 minutes. The reaction was stopped by the addition of 125μl of 5M HCl to cyclise any preuroporphyrinogen, followed by 50 μl of benzoquinone solution (0.1% w/v in methanol) to oxidise the uroporphyrinogens to uroporphyrins. The assay mixture was stored in the dark for 20 minutes, before centrifuging for 1 minute at 10,000g. A 100μl sample of the supernatant was added to 900μl of 0.1M HCl and the absorbance was measured at 405.5 nm using a Hitachi U-200 spectrophotometer. The enzyme activity was calculated in μmoles PBG consumed per mg enzyme per hour.

The kinetic parameters, K_m and V_{max}, were determined by assaying the enzyme at a range of substrate concentrations.

2.3.3 Identification of the dipyrromethane cofactor

Porphobilinogen deaminase (500µl of a 1mg/ml solution) was mixed with an equal volume of modified Ehrlich's reagent at room temperature. The absorption was measured by scanning between 380 and 600 nm at time zero and 15 minutes. Observation of a peak with

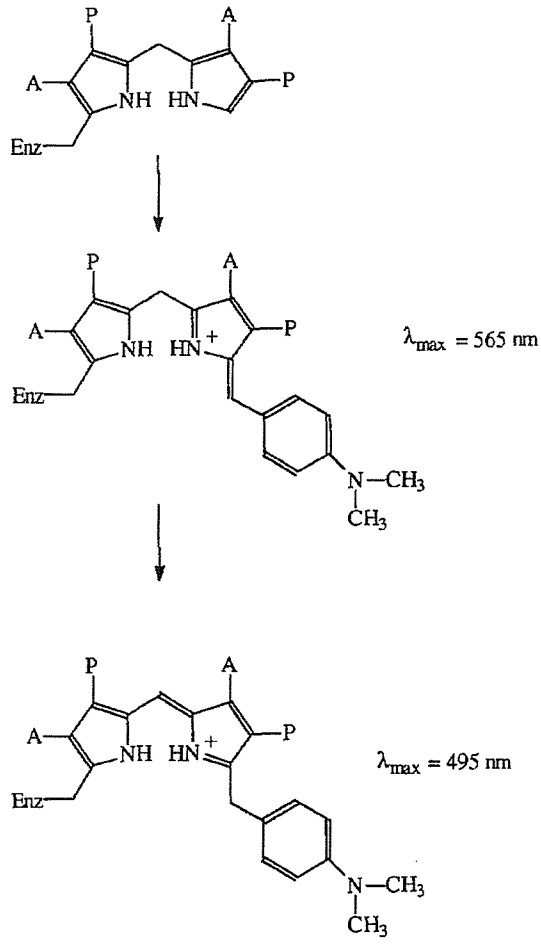


Figure 2.3 The reaction of Ehrlich's reagent with the dipyrromethane cofactor.

λ_{max} of 565 nm and subsequent shift to a λ_{max} of 495 nm confirmed the presence of the cofactor. The reaction was blanked using buffer and Ehrlich's reagent. Figure 2.3 shows the reaction between the enzyme containing the dipyrromethane cofactor and Ehrlich's reagent.

2.3.4 Enzymic preparation of PBG

The following procedure was carried out in the absence of light. ALA (5mM) was dissolved in 50mM Tris/HCl, pH 7.5, containing 50μM $ZnCl_2$, 1mM $MgCl_2$, and 20mM β-mercaptoethanol. ALA dehydratase (10 – 20 mg depending on activity) was added and the mixture was gently stirred at room temperature until the PBG concentration, monitored by reaction with modified Ehrlich's reagent, did not increase further. The PBG was then bound to the anionic exchange resin Dowex acetate (1X8-400) using the batch method. The Dowex acetate resin was prepared by washing with glacial acetic acid followed by rinsing with distilled water that had been adjusted to pH 8.0. Once PBG had bound to the resin, washing was continued with distilled water (pH 8.0) until the pH of the elutant was 8.0. This ensured that excess ALAD and ALA were removed. Finally the washed resin was loaded into a Pharmacia column (20cm x 2.5cm) and the PBG was eluted using 0.3M acetic acid. Again the process was monitored using Ehrlich's reagent and the fractions containing PBG were rotary evaporated and freeze-dried. The resultant solid PBG was stored in the dark at -20°C for up to 6 months.

2.3.5 Quantification of porphobilinogen

10μL of PBG solution was added to 490μL of H_2O and mixed with 500μL of modified Ehrlich's reagent. The sample was left in the dark for 20 minutes. The absorption was measured at 555nm and the concentration of PBG calculated using a molar extinction coefficient of 60200 $M^{-1}cm^{-1}$.

2.3.6 Polyacrylamide gel electrophoresis

All denaturing polyacrylamide gel electrophoresis was carried out according to the Laemmlli method (Laemmli, 1970). The composition of the gels are shown in tables 2.4 and 2.5. Protein samples were prepared by the addition of the appropriate amount of disruption buffer followed by boiling for 2 minutes. After

the samples were loaded, electrophoresis was carried out at 120 volts until the bromophenol blue dye front had reached the bottom of the gel. Coomassie brilliant blue stain was used to visualise the protein samples, which were compared with standard protein markers (table 2.6).

	Small 12% SDS -PAGE	**Large 12% SDS-PAGE**	**10% Native-PAGE**
Acrylamide	3.2mL	18.20mL	3mL
1.5M Tris/HCl pH 8.8	2mL	11.25mL	2.8mL
dH_2O	2.8mL	15.1mL	3.7mL
10% SDS	80µL	450µL	—
10% APS	100µL	40µL	100µL
TEMED	20µL	40µL	40µL

Table 2.4 Compositions of the resolving polyacrylamide gels gels used for protein analysis.

	Small 12% SDS-PAGE	**Large 12% SDS-PAGE**	**10% Native-PAGE**
Acrylamide	1mL	3mL	1mL
0.5M Tris/HCl pH 6.8	1mL	7.5mL	2mL
dH_2O	3mL	19.5mL	2mL
10% SDS	80µL	300µL	—
10% APS	100µL	300µL	100µL
TEMED	20µL	45µL	40µL

Table 2.5 Compositions of stacking polyacrylamide gels used for protein analysis.

Protein	Molecular weight - Daltons
Bovine serum albumin	66,000
Egg albumin	45,000
Glyceraldehyde-3-phosphate dehydrogenase	36,000
Carbonic anhydrase	29,000
Trypsinogen	24,000
Soybean trypsin inhibitor	20,100
α-Lactalbumin	14,200

Table 2.6 Dalton VII marker protein molecular weight standards

2.3.7 Mass spectrometry of proteins

Samples of protein were prepared by diluting the enzyme to 0.1 mg/mL in 50:50:1 acetonitrile:water:formic acid. Samples were analysed on a Micromass VG Quattro II mass spectrometer. All data was processed using MassLynx and Maximum Entropy programs.

2.3.8 Western blotting and N-terminal amino acid sequencing

A large 12% SDS-PAGE resolving gel was cast and overlaid with water saturated *n*-butanol, before leaving to polymerise overnight. The following day, the stacking gel was cast, using resolving buffer in place of stacking gel buffer. To remove any free radicals, the gel was pre-run at 60 volts with 0.125M Tris/HCl, pH 6.5, containing 2% (w/v) SDS (stacking gel buffer) (Dunbar and Wilson, 1994). This buffer exchange procedure allows the resolving gel buffer in the stacking gel to be replaced with stacking gel buffer. Completion of this exchange is judged by migration of loading buffer to the stacking/resolving gel boundry. Next the anode buffer was replaced with standard SDS-PAGE running buffer and the protein samples prepared and loaded as usual. After electrophoresis the blotting was performed using a Biorad Transblot apparatus and PVDF membrane.

The gel was soaked in 10mM CAPS pH 11.0 to remove any traces of glycine. The gel was laid onto the membrane and both were sandwiched between sheets of blotting paper. The membrane had been washed in methanol and pre-soaked in 10mM CAPS buffer. The sandwich assembly was loaded into the blotting apparatus, and the Transblot cell was filled to the top with 10mM CAPS pH 11.0. Blotting was carried out for one hour at a current of 0.5Amps. The membrane was then stained using amidoblack for approximately 20 minutes. Finally, destaining allowed transferred protein to be visualised. For *N*-terminal amino acid sequencing a portion of the blot was excised and analysed by Edman degradation.

2.3.9 Protein crystallisation

The hanging drop method was used to crystallise purified protein. This relies on the vapour diffusion technique (McPherson, 1976) where the droplet containing the protein to crystallise with buffer, crystallising agent and additives, is equilibrated against a reservoir containing a solution of crystallising agent at a higher concentration. Equilibration proceeds by diffusion of the volatile species (water and organic solvents) until the vapour pressure in the droplet equals that in the reservoir. Consequently, the concentration of all the constituents in the protein drop will increase resulting in protein crystallisation or precipitation.
To set up the hanging drop crystallisation screens, 1mL of crystallising agent was pipetted into each well of a 24 well Linbro tissue culture plate. On a siliconised glass coverslip, 1µL of protein solution was mixed with 1µL of crystallising agent before inverting over the appropriate well which had been previously coated with vacuum grease.

2.3.10 Expression studies

For recombinant plasmids containing the T7-promoter a single colony was inoculated in 10ml of LB media supplemented with ampicillin (100µg/ml) and grown to an O.D of 1.0 at 600nm. Next IPTG was added to a final concentration of 1mM and the cultures grown for a further 3 hours at 37°C before harvesting by

centrifugation (3000g /10 minutes/4°C). For pUC18/19 recombinant plasmids a single colony was inoculated in 10mL of LB media supplemented with ampicillin and grown for 16 hours before harvesting the bacteria as above. The bacterial cell pellet was resuspended in 0.5ml of 20mM Tris/HCl pH8.2 + 1mM DTT in an eppendorf. The extract was sonicated at an amplitude of 10 microns in five bursts (30 seconds on / 60 seconds off) using the smallest probe. The crude sonicates were centrifuged (10 minutes at 10,000g and 4°C) before transferring the supernatant to a separate tube .The pellets were centrifuged again and any remaining supernatant was removed by aspiration.

2.3.11 Large-scale culture growth

After selecting the most active clone, starter cultures were grown overnight (16 hours) at 37°C with shaking (at 100 rpm). For expression of genes under the control of the T7-promoter six baffled flasks each containing 800ml of LB media, were inoculated with 10ml of overnight culture and grown at 37°C with shaking (200rpm) to an OD_{600nm} of 1.0 before inducing with 1mM IPTG (final concentration). Growth was continued for an additional 3 hours before centrifuging at 4000g for 25 minutes at 4°C to harvest the bacteria. Similarly, for expression of genes in the pUC vectors, six flasks were inoculated with 1mL of starter culture and grown at 37°C with shaking for 16-18 hours. The cells were then harvested by centrifuging at 3000g for 25minutes at 4°C

2.3.12 Isolation of bacteriophage-resistant bacterial strains

Chloroform, 50µL, was mixed with 0.5mL of the bacteriophage infected culture by inversion, before centrifuging at 5000g to clarify the solution. The supernatant was serially diluted using 10mM Tris/HCL pH 8.0 containing 10mM $MgCl_2$. Competent cells of the same bacterial culture (section 2.2.12) were mixed with 3mL of molten H-top agar by rolling / inversion and poured onto a LB plate. Once set, the agar was spotted with the serially diluted solutions containing the bacteriophage and grown overnight at 37°C. The bacteriophage resistant cultures

were selected from the agar plate and transformed with the recombinant plasmid of interest before checking for protein expression (section 2.3.10). This process was repeated until a strain was isolated that expressed sufficient levels of protein in comparison to the original wild-type strain.

2.3.13 Selecting and optimising an ionic-exchange purification step

Minicolumns were filled with 5mL of the selected ion-exchange material and equlibrated with an array of different buffer solutions at varying pH values. The enzyme solution (1mL) was loaded onto the minicolumn, before washing through with five column volumes (CV's) of buffer and collecting the elutate in 1mL fractions. Washing the columns was repeated with buffer plus 0.5M KCl and again the elutate was collected. All collected fractions were dialysed into the original buffer to remove any salt. Activity and protein concentration measurements confirmed if the enzyme was firmly bound or eluted from the column under the different conditions. From these results a buffer was selected in which the enzyme bound strongly to the ion-exchange material, but was eluted by the addition of 0.5M KCl to the buffer. Optimal separation of desired protein from contaminating proteins was determined using a range of salt gradients (0.1 – 0.5mM NaCl) in the elution buffer. Analysis the eluted fractions by SDS-PAGE and activity assays indicated the salt concentration and elution volume that gave maximum resolution.

2.3.14 Protocol for the purification of *E. coli* porphobilinogen deaminase and heat stable mutants

Bacteria containing the recombinant plasmid of interest were grown and harvested as described in section 2.3.11.

Sonication and heat treatment

The bacterial cell pellet was resuspended in 0.1M potassium phosphate buffer, pH 8.0, 14mM β-mercaptoethanol, 0.2mM PMSF and lysed by sonication using an

MSE Soniprep ultrasonic disintegrator. The lysate was heated at 60°C for 10 minutes and then rapidly cooled by placing on ice. The precipitated heat labile proteins were removed by centrifugation at 8000g at 4°C for 20 minutes and the supernatant was collected.

Ammonium sulphate fractionation and dialysis

Solid ammonium sulphate was added slowly to the supernatant to give a final saturation of 30% (16.4g / 100 mL). After leaving the extract to equilibrate, the mixture was centrifuged as before and the precipitate was discarded. On further addition of ammonium sulphate to give 60% saturation (36.1g / 100mL), the mixture was centrifuged (20 minutes at 8000g) and the pellet was collected. The pellet was dissolved in a minimal amount of 0.1M potassium phosphate buffer, pH 8.0, 14mM β-mercaptoethanol and pipetted into dialysis tubing. Dialysis with slow stirring was carried out overnight against 4 litres of cold, 0.1M potassium phosphate buffer, pH 8.0, 14mM mercaptoethanol.

DEAE Sephacel ion exchange chromatography

The dialysed protein sample was loaded onto a DEAE-Sephacel column (20cm x 2.5cm) that had been equilibrated with 0.1M potassium phosphate buffer, pH 8.0 containing 14mM mercaptoethanol. The fractions collected were examined by assay for uroporphyrinogen I formation and by SDS-PAGE. The fractions containing enzyme were pooled and concentrated using an Amicon 8200 ultrafiltration cell fitted with a PM10 membrane. After use the column was cleaned by washing it with 2M NaCl, followed by distilled water, before storing in 20% ethanol at 4°C.

Desalting protein solutions using PD10 columns

The concentrated enzyme solution was desalted into AnalaR water prior to freeze-drying or buffer for high-resolution ion-exchange chromatography using a PD10 column. The PD10 columns were pre-packed with 9mL of Sephadex G-25 medium and were used as follows: -

The column was equilibrated with 25mL of AnalaR water or the required buffer.

The sample was loaded to a total volume of 2.5mL. If the sample volume was less than 2.5mL, the sample was run into the column followed by water or buffer to a total volume of 2.5mL. The protein was eluted with 3.5mL of water or buffer and the fractions containing protein identified using the BIO-RAD assay (see 2.3.1) before pooling.

High resolution ion-exchange chromatography

The concentrated enzyme solution was applied to a high-resolution anion exchange Mono-Q HR5/5 column using a Pharmacia FPLC system. The column was pre-equilibrated in 20mM Tris/HCl buffer, pH 7.5 and a linear gradient of NaCl was applied to the column to elute the enzyme (100-350mM total volume 20mL). The elution was monitored at 280nm and the peaks were collected and assayed for deaminase activity. The enzyme was desalted through a PD10 column and freeze-dried before storing at -20°C.

2.3.13 Separation of enzyme-substrate intermediates by high resolution ion exchange chromatography

This technique was also used for the isolation of enzyme-substrate intermediates ES, ES_2, ES_3 (Warren and Jordan, 1988). Enzyme was mixed with varying molar ratios of porphobilinogen and immediately applied to the Mono-Q column, before salt elution. Variations of this method are included in the relevant chapters.

Chapter 3 Studies on recombinant human porphobilinogen deaminase: cloning, expression and purification of the erythroid and ubiquitous isoenzymes.

3.1 Introduction

Porphobilinogen deaminase from humans is encoded by a single gene of 10kb encompassing 15 exons (Raich *et al.*, 1986;Grandchamp *et al.*, 1987) located on chromosome 11 in the region 11q24.1Ø11q24.2 (Namba *et al.*, 1991). The gene contains two promoters and is differentially transcribed and spliced in a tissue-specific manner to give ubiquitous (housekeeping) and erythroid mRNAs (Chretien *et al.*, 1988). The ubiquitous isoenzyme, specified by exon 1 and exons 3-15, has 361 amino acids whereas the erythroid isoenzyme, specified by exons 2-15 is smaller with 344 amino acids (figure 3.1). The ubiquitous isoenzyme has a M_r of 44,000 and includes an N-terminal extension that is 17 amino acids longer than the erythroid enzyme, M_r of 42,000.

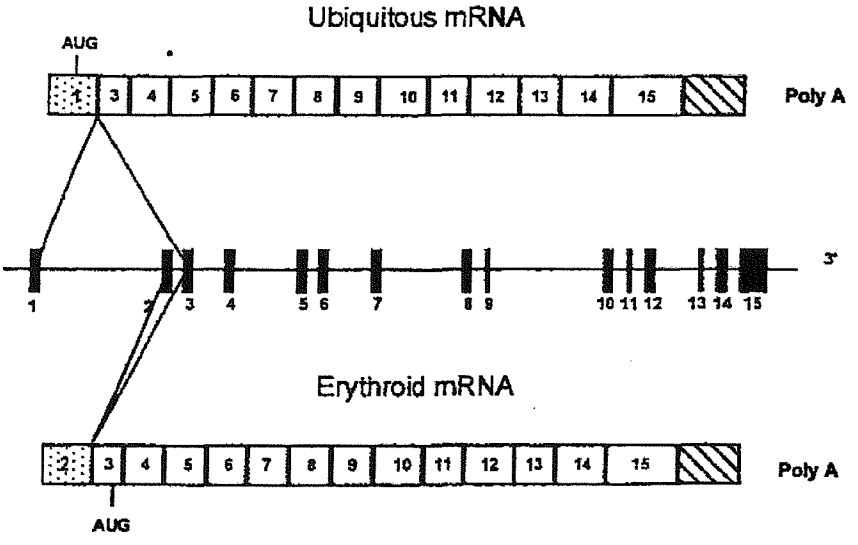


Figure 3.1 Genomic organisation of the human porphobilinogen deaminase gene and alternative splicing of the ubiquitous (housekeeping) and erythroid specific transcripts (Deybach and Puy, 1995).

Human erythroid porphobilinogen deaminase (also called uroporphyrinogen synthase) has been purified over 42,000 fold from human eythrocytes with a specific activity of 2,300 nmoles uroporphyrin formed/hr/mg (Anderson and Desnick, 1980). The enzyme has a pH optimum of 8.2 and a K_m of 6μM and is monomeric and exists in several forms attributable to enzyme-substrate intermediates. Experiments with tritiated porphobilinogen suggested that these forms represented enzyme bound to one, two, three and four substrate molecules. Purification of the human erythrocyte enzyme (also called hydroxymethylbilane synthase) has also been achieved by the use of affinity dye chromatography (Smythe and Williams, 1988).

The ubiquitous form of porphobilinogen deaminase has been purified over 1,000 fold from human liver hepatocytes with a specific activity of 32 nmole/hr/mg, substantially lower than that of the erythroid enzyme. Enzyme substrate covalent intermediates were also detected (Mazzetti and Tomio, 1993) and studies with dicarbonyl reagents and Woodward's reagent K (*N*-ethyl-5-phenylisoxazolium-3-sulphonate) highlighted the importance of arginine and carboxyl residues, respectively, for the functioning of the enzyme (Mazzetti and Tomio, 1997).

Studies on deaminase from rat liver (Mazzetti and Tomio, 1988) and the enzyme from Rhodobacter sphaeroides (Jordan and Shemin, 1973;Davies and Neuberger, 1973) and spinach also indicate that deaminases are monomeric enzymes with similar properties. This has been confirmed by comparing the nucleotide-derived amino acid sequence of over twenty deaminases that show a remarkable conservation whether from eukaryotes or prokaryotes.

Defects in the human PBGD gene can lead to acute intermittent porphyria, an autosomally inherited disease with low penetrance. The overt disease is characterised by acute attacks of abdominal pain, muscular weakness and a complicated array of neurovisceral and psychiatric symptoms. The condition can deteriorate and become life threatening if incorrectly diagnosed and appropriately treated. Acute attacks may be induced by certain drugs, chemicals, alcohol and dietary or hormonal factors (Elder *et al.*, 1997). During an acute attack, haem biosynthesis is induced and, as a result of a rise in ALAS levels the early

precursors, 5-ALA and porphobilinogen, are overproduced and accumulate in the blood and urine. Whether it is the direct action of the accumulated metabolites or the intracellular haem deficiency, or both, that leads to the clinical features of the acute attacks is of some debate (Thunell, 2000).

Over 200 mutations affecting the human PBGD gene are known (Brownlie *et al.*, 1994). These include missense and nonsense mutations and a number of splicing mutations as well as deletions and additions. Examples of mutations are found in all 15 exons. Using the X-structure of the *E. coli* PBGD, many of the mutations in the human enzyme have been mapped and their functional consequences evaluated (Wood *et al.*, 1995). This has been possible because of the high amino acid sequence similarity between the *E. coli* and human proteins - 60% similarity and 40% identity.

Despite the broad similarities between the *E. coli* and human deaminases, there are a few specific differences between the enzymes that make it essential to determine the X-ray structure of both human deaminases independently. One of the major differences between the *E.coli* and the human enzymes relates to domain 3 where there is an insertion of 29 amino acids. In addition, the ubiquitous human enzyme has an *N*-terminal sequence of 15 extra amino acids compared with the *E. coli* deaminase. There are also differences in amino acid residues in a number of residues where human mutations have been identified, making it difficult to use the *E. coli* PBGD structure to predict effects on human enzyme function.

This chapter describes the construction of plasmids harbouring the cDNA specifying ubiquitous and erythroid porphobilinogen deaminases. Using PCR, two segments of DNA have been generated each with a conveniently located restriction site to allow cloning into the plasmid pT7-7 for expression in BL-21. The DNA thus expressed in strains of *E. coli* yield milligramme amounts of ubiquitous and erythroid human recombinant porphobilinogen deaminases that have been purified to homogeneity prior to crystallisation and X-ray crystallography.

3.2 Cloning of human PBGD cDNA

A full-length human PBGD cDNA construct was provided as a gift from Professor Bernard Grandchamp (Department of Molecular Genetics, Faculty of Medicine, Hôpital X. Bichat, 75018, Paris, France). This construct consisted of the plasmid pGEM3 (Bluescript) with a 1.492 kb human PBGD cDNA insert between the restriction sites *Pst*I and *Acc*I, encoding both human PBGD isoenzymes (figure 3.2). This recombinant plasmid was used only as a source of the human PBGD cDNA and was not the final construct used for protein expression. Initially, the plasmid encoding the cDNA was purified and analysed to confirm its sequence and restriction sites (section 3.2.2). Once isolated, two different methods were attempted to clone the human PBGD cDNA, the second method enabling the overexpression and purification of both the human erythroid and ubiquitous PBGDs.

3.2.1 Amplification and purification of the human PBGD cDNA construct

The pGEM construct (figure 3.2) was transformed into the *E. coli* transcription strain DH5α and transformants were selected on LB agar plates containing ampicillin (100 μg/mL) using the α-complementation screening technique (section 2.2.13). To obtain larger quantities of the recombinant plasmid, a large-scale dsDNA preparation was carried out using the ethylene glycol method of purification (section 2.2.1).

3.2.2 Analysis of the human cDNA construct by restriction and DNA sequencing

The purified cDNA construct was digested with three restriction endonucleases (*Kpn*I, *Sma*I and *Hin*dIII) and the digest analysed by agarose gel electrophoresis (figure 3.3) for comparison with theoretical fragment sizes (table 3.1). Identity between the actual and theoretical fragment sizes, combined with DNA

sequencing (figure 3.4) confirmed the correct construct has been isolated. Alignment of the sequencing results for the human PBGD cDNA region with the PDB sequence revealed two single point mutations: leucine 188 CTA → leucine CTG, a silent mutation and lysine 210 AAA → glutamate GAA, a natural variant in the human PBGD gene sequence (Raich *et al.*, 1986). Spectrophotometry of the purified DNA indicated a pure sample free from RNA contamination.

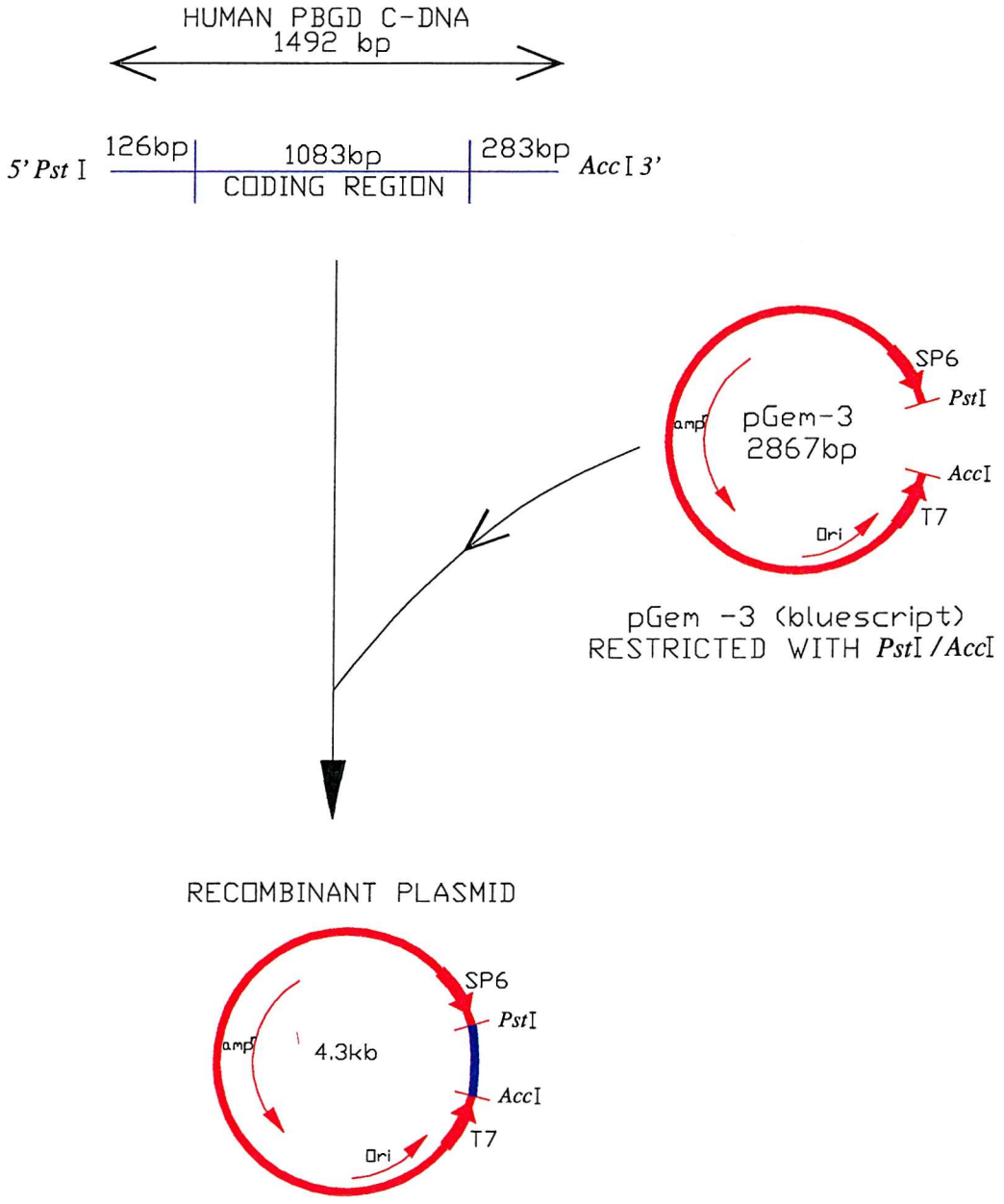


Figure 3.2 A map of the recombinant pGEM-3 plasmid harbouring human PBGD C-DNA. The 1.492 kb cDNA is inserted between the *Pst*I and *Acc*I restriction sites, in the 5' to 3' orientation. The 1.492 kb cDNA insert is divided into a coding region of 1083 bases (≡ 361 amino acids), plus 5' and 3' untranslated regions of 126 and 283 bases respectively.

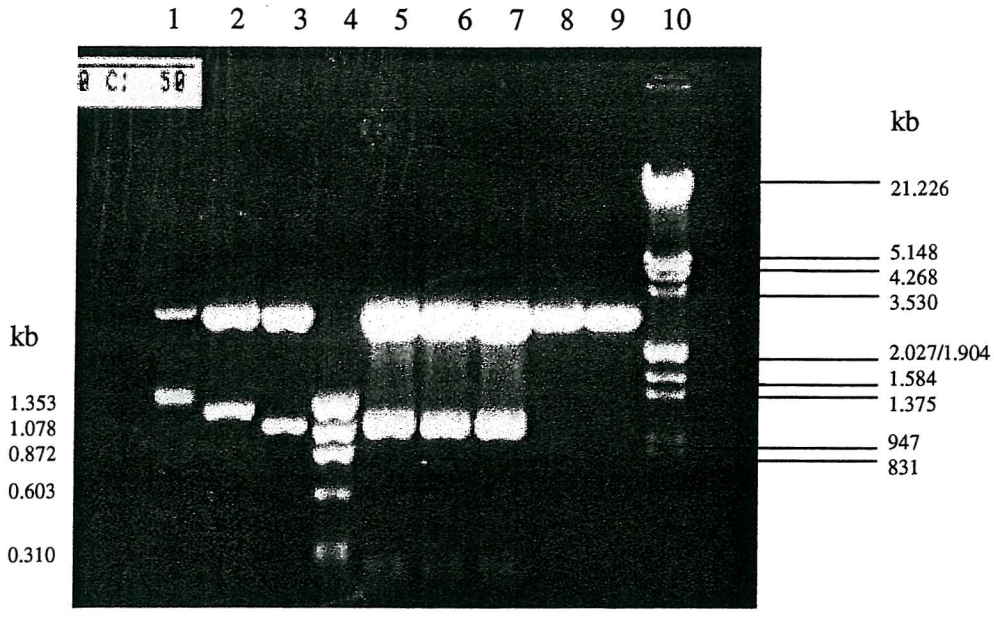


Figure 3.3 Agarose gel eletrophoresis showing the restriction analysis of the recombinant human porphobilinogen deaminase cDNA construct. Lane 1 = restriction with *Kpn*I; Lane 2 = restriction with *Sma*I; Lane 3 = restriction with *Kpn*I + *Sma*I; Lane 4 = ΦX174 DNA standard; Lanes 5 – 7 = *Kpn*I + *Sma*I; Lane 8 = pUC19 restricted with *Sma*I; Lane 9 = pUC19 restricted with *Hin*dIII; Lane 10 = λ *Eco*RI / *Hin*dIII DNA standard. The lower band seen in lanes 5 – 7 is the *KpnI/ SmaI* fragment utilised in cloning method 1.

***Kpn*I restriction** **Fragment sizes (bases)**	***Sma*I restriction** **Fragment sizes (bases)**	***Kpn* + *Sma*I restriction** **Fragment sizes (bases)**
3040	2918	2912
1320	1193	1065

Table 3.1 Theoretical fragment sizes produced by restriction of the recombinant human porphobilinogen deaminase cDNA clone with the restriction enzymes *Kpn*I and *Sma*I. Comparison of the theoretical fragment sizes with the actual restriction fragments produced (figure 3.3) combined with DNA sequencing (figure 3.4) confirmed the correct construct has been isolated.

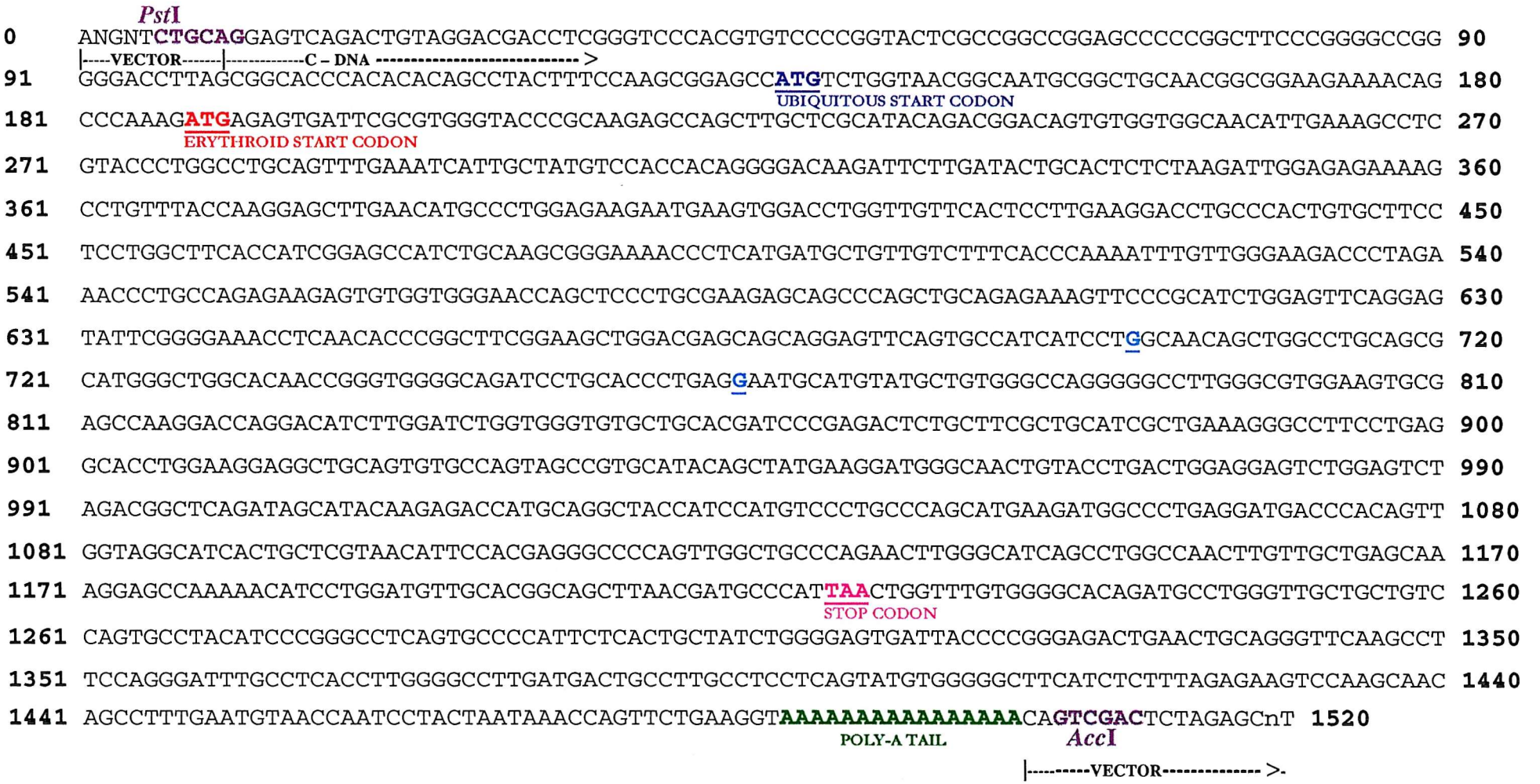


Figure 3.4 Nucleotide sequence of human PBGD cDNA. The ubiquitous start codon is highlighted in blue, preceding the erythroid start codon (red) 49 bases downstream. There is a single stop codon shown in pink followed by the poly-A tail (green). The restriction sites utilised for cloning the human PBGD cDNA into the transcription vector pGEM are coloured in purple. Comparison of the above sequence with the published sequence for human porphobilinogen deaminase reveals two single point mutations (cyan) leucine 188 CTA → CTG, a silent mutation, and lysine 210 AAA→ glutamate GAA, a natural variant in the human PBGD gene sequence. Additionally flanking vector sequences are shown at the 5' and 3' ends of the cDNA.

3.2.3 Cloning of human porphobilinogen deaminase cDNA– method 1

The basic concept was to modify the human PBGD cDNA for efficient protein expression in a bacterial host as illustrated in figure 3.5. This method was divided into two main stages. The first involved ligation of a double-stranded oligonucleotide, encoding an *E. coli* promoter, ribosome-binding site and unique restriction sites to the purified human PBGD cDNA fragment. The second part of the method comprised insertion of this ligation product, the human PBGD expression cassette, into the vector pUC19, prior to transformation into *E. coli* strain DH5α. Initially, cloning of the shorter erythroid cDNA was carried out.

3.2.3.1 Stage 1 Production of the expression cassette encoding human erythroid porphobilinogen deaminase

Design and synthesis of the double-stranded oligonucleotide

Essential features of the double-stranded oligonucleotide (promoter, ribosome-binding site and translational spacer element) were selected by copying known *E.coli* consensus sequences (figure 3.6). Unique restriction sites were included at the 5' and 3' ends for sub-cloning into the vector pUC19 and ligating to the purified erythroid PBGD cDNA fragment. The twenty-three bases at the 3'-end of the double-stranded oligonucleotide were incorporated to complete the erythroid PBGD gene sequence as these are absent in the isolated erythroid PBGD cDNA fragment. This substitution allowed the first six codons of the erythroid PBGD gene to be changed to equivalent codons that express well in *E. coli* (Sharp *et al.*, 1988).

The synthetic DNA fragment was synthesised as two single-stranded oligonucleotides, which were purified and annealed to give the double-stranded synthetic DNA fragment (section 2.2.5). Finally the oligonucleotide mixture was phosphorylated as in section 2.2.6.

Purification of erythroid PBGD cDNA fragment

The erythroid PBGD cDNA fragment was generated by restriction of the human PBGD cDNA construct (figure 3.2) with the enzymes *Kpn*I / *Sma*I. The 1065

base fragment was identified by agarose gel electrophoresis (figure 3.3) and gel purified using the 'Geneclean™' kit.

Ligation of the double-stranded oligonucleotide to the gel purified cDNA fragment

The double-stranded oligonucleotide was ligated to the purified cDNA fragment as described in section 2.2.11. Analysis of the ligation mixture (5μL) on a 1.2 % agarose gel and comparison with the original 1.1 kb *KpnI/SmaI* cDNA fragment indicated the correct ligation products (figure 3.7). The remaining ligation mixture was restricted with *Sma*I and *Hin*dIII, ethanol precipitated and resuspended into 10 μL of sterile H_2O.

3.2.3.2 Stage two – ligation of the expression cassette with the vector pUC19

The product from Stage 1, the human erythroid PBGD expression cassette, was ligated to the vector pUC19 that had been restricted with *Hin*dIII and *Sma*I. The ligation mixture was analysed by agarose gel electrophoresis (section 2.2.7). The gel indicates that incorrect ligation products were formed. The reason for such low ligation efficiency between the expression cassette and the vector could be due to the incorporation of the 'blunt-ended' *Sma*I restriction site at the 3' end of the erythroid cDNA fragment. Blunt-end ligation is often problematic, as many of the ligated molecules, are self-ligated, due to the low specificity of the technique. Additionally, it is difficult to optimise a ligation reaction that incorporates both sticky and blunt-ends, because they have different optimal reaction conditions.

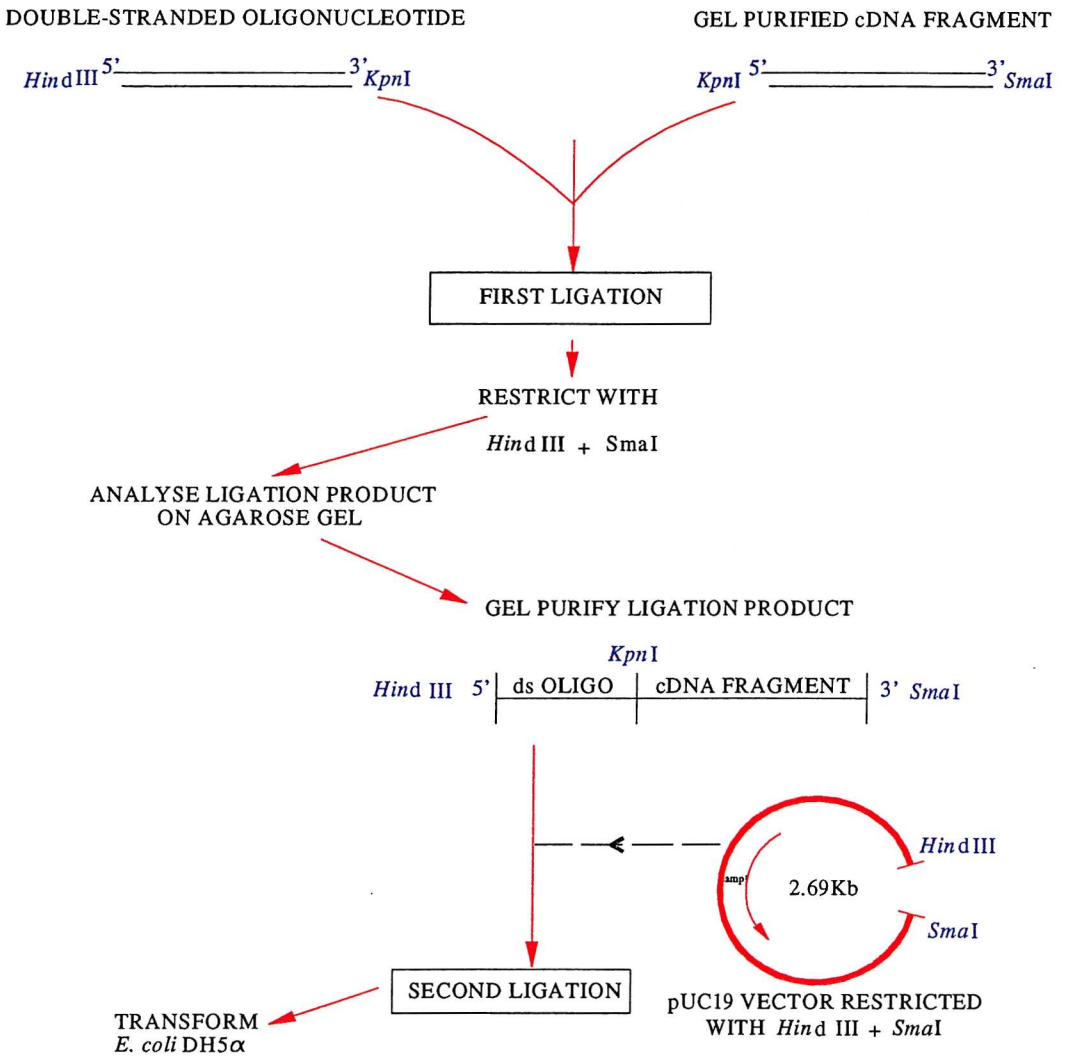


Figure 3.5 Cloning of human porphobilinogen deaminase. The first step involved ligation of the gel purified cDNA fragment to a double-stranded oligonucleotide encoding essential transcription and translation initiation sequences. The second stage comprised insertion of the first ligation product, the human PBGD expression cassette, into the vector pUC19. This recombinant DNA construct was transformed into *E .coli*

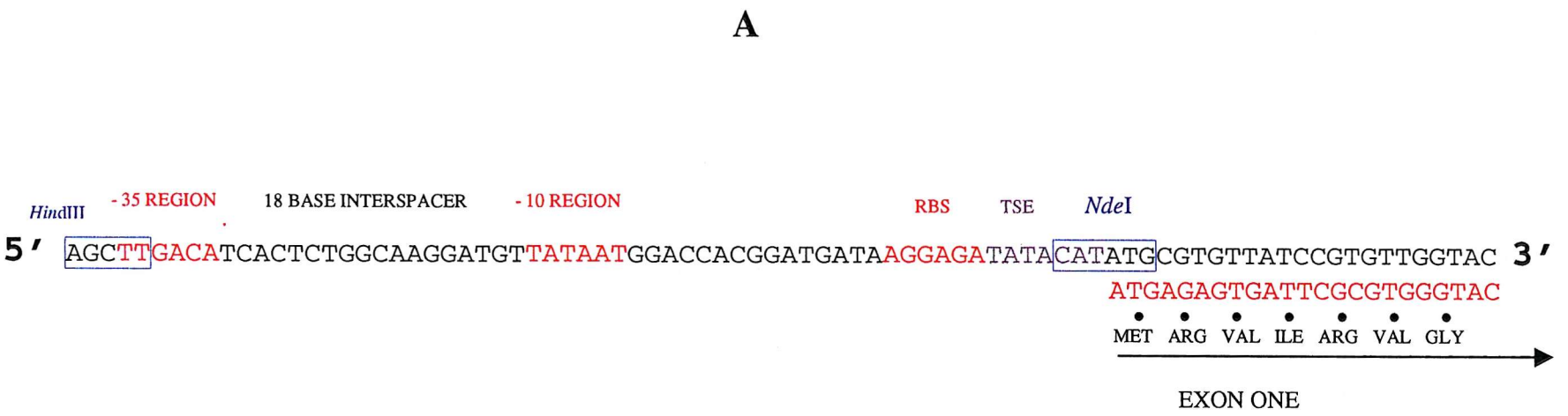


Figure 3.6 The double-stranded oligonucleotide was constructed by the annealing of two complementary single-stranded oligonucleotides A and B.

A) Consensus sequences taken from *E. coli* were copied for the promoter, ribosome-binding site (RBS) and translational spacer element (TSE). The promoter consists of two regions known as the –35 and –10 regions. The 18 base interspacer was duplicated from the *E. coli hemC* gene, as this gene is efficiently expressed. Restriction sites (boxed in blue) are necessary for sub-cloning into the vector. The first six codons after methionine have been substituted for codons that are favoured in *E. coli*. For comparison the original gene sequence is shown in red. B) Sequence of primer B. This is the complementary strand to primer A, required to make a double-stranded fragment upon annealing.

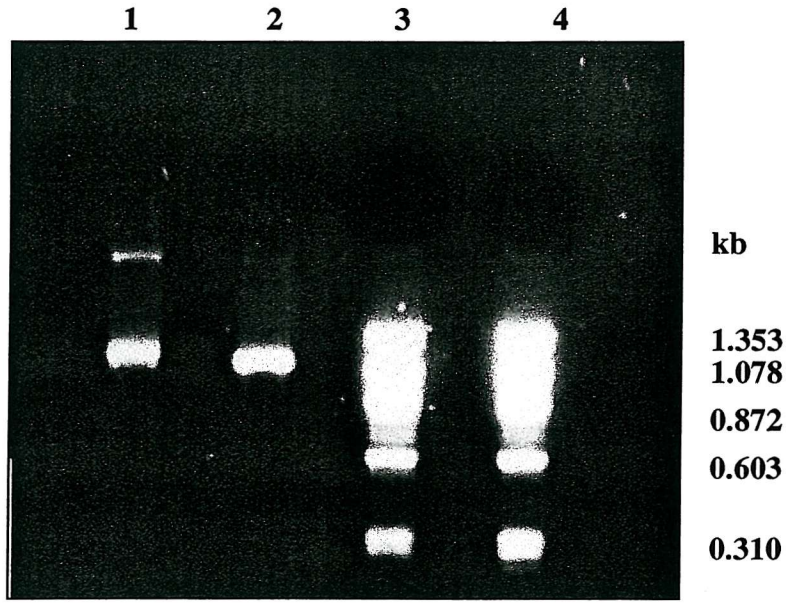


Figure 3.7 Analysis of ligation products from stage 1 of the cloning of human erythroid PBGD cDNA by agarose gel electrophoresis. Lane 1 = 5μL of the ligation of the double-stranded oligonucleotide to the human erythroid PBGD cDNA fragment; Lane 2 = 1μL of the original*Kpn*I / *Sma*I gel purified cDNA fragment; Lanes 3 – 4 = ΦX174 DNA standard. The reduced mobility of the main band in lane 1 in comparison to the band in lane 3 and the additional upper band in lane 1 indicates the correct ligation product.

3.2.4 Cloning of human PBGD cDNA using expression cassette PCR – Method 2

The method is outlined in figure 3.8. It involved the use of expression cassette PCR (ECPCR) to produce an expression cassette, which was ligated to the expression vector pT7-7 yielding a recombinant plasmid. This was done in collaboration with Dr. Sarwar.

3.2.4.1 Primer design

Two sets of primers were designed (Dr. Sarwar), one for the amplification of the non-erythroid coding sequence and one for the erythroid. The two forward (5') primers were designed with a *Nde*I restriction site. One single reverse (3') primer was designed with a *Bam*HI restriction site (table 3.2). The incorporation of *Nde*I and *Bam*HI restriction sites into the amplified cDNA facilitated subcloning into

the expression vector pT7-7. The restriction sites selected were absent in the PBGD coding region so that the cDNA coding region was not disrupted upon subcloning. Once the primers had been designed, they were synthesised and purified as in section 2.2.4.

Primer	Oligonucleotide sequence
Primer one non-erythroid forward primer (62 bases)	**5'**GGGAATTCCATATGTCTGGTAACGGTAACGCTGCTGCAA *Nde*I CGGCGGAAGAAAACAGCCCAAAG **3'**
Primer two Erythroid forward primer (56 bases)	**5'**GGGAATTCCATATGCGTGTTATCCGTGTTGGTACC *Nde*I CGCAAGAGCCAGCTTGCTCGC **3'**
Primer three Reverse primer (34 bases)	**5'** GCGCGGATCCGGGATGTAGGCACTGGACAGCAGC **3'** *BamH*I

Table 3.2 The oligonucleotide primers employed for PCR of human PBGD cDNA specifying the non-erythroid and erythroid forms. The engineered restriction sites are underlined. The mismatched bases are shown in red and the complementary bases in black. The additional bases 5' of the restriction sites are essential for optimum binding of the restriction enzymes. Bases 3' of the restriction sites have been designed to maximise translational efficiency for the first five/six codons proceeding the initial methionine residue (Dr. Sarwar).

3.2.4.2 ECPCR of the human PBGD cDNA

The human ubiquitous and erythroid PBGD cDNA were amplified by ECPCR (section 2.2.16). This produced amplified fragments of sizes 1142 and1090 bases, for the ubiquitous and erythroid coding sequences respectively (figure 3.10). The amplified fragments were extracted from the agarose gel and restricted with *Nde*I and *Bam*HI before ligating into the expression vector pT7-7 (section 2.2.11).

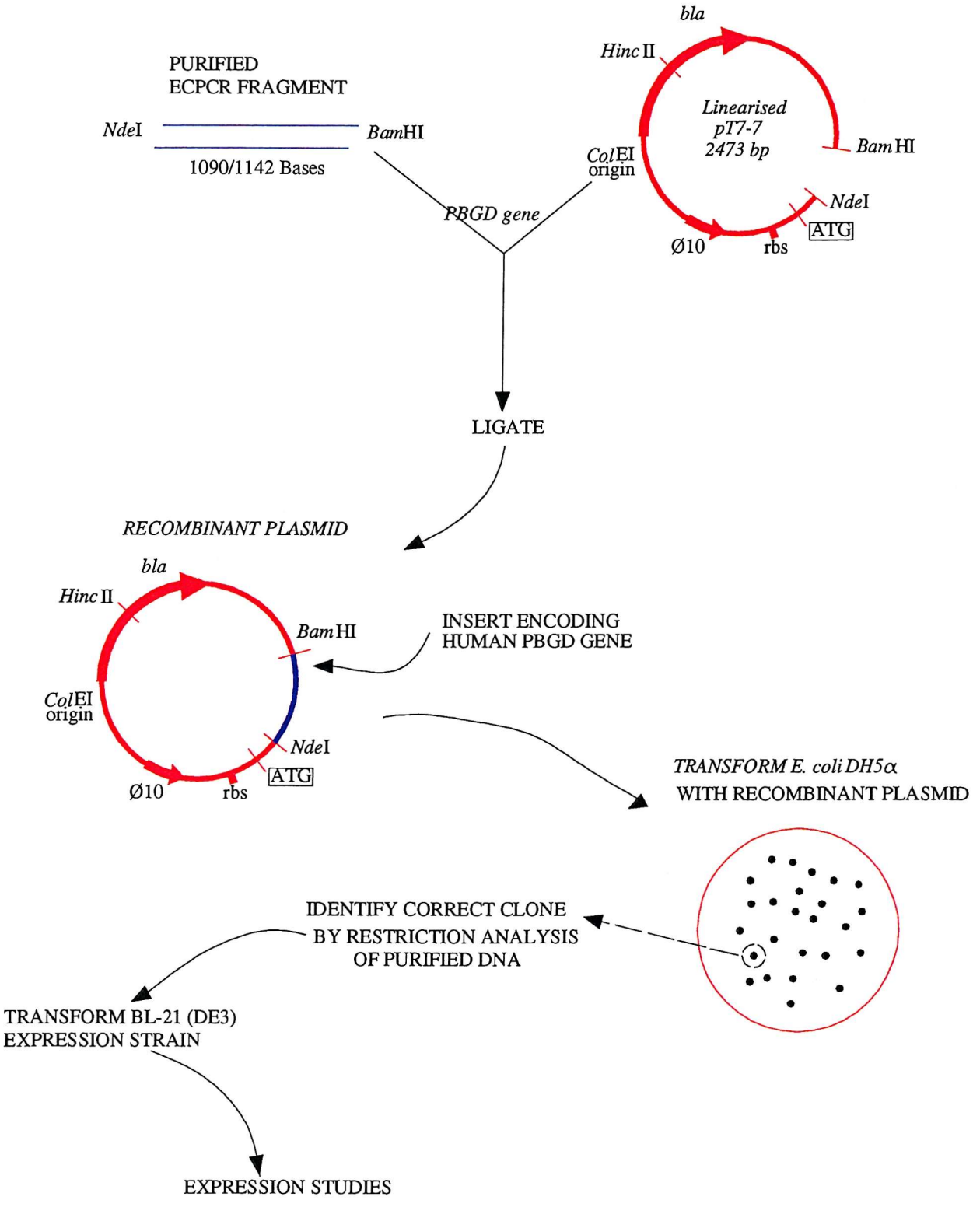
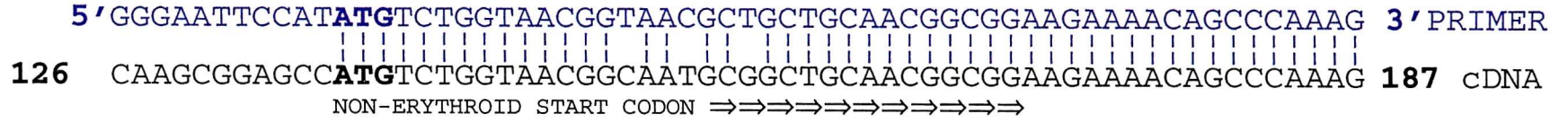


Figure 3.8 Cloning of human PBGD by expression cassette PCR. Firstly the expression cassette is produced by ECPCR before purifying and ligating into the expression vector pT7-7. This recombinant plasmid is transformed into *E. coli* strain DH5α for DNA replication in order to confirm the correct recombinant DNA has been isolated. Finally this recombinant DNA is transformed into *E. coli* strain BL21-DE3 (This was done in collaboration with Dr. Sarwar).

Primer one ubiquitous PBGD forward primer



Primer two erythroid PBGD foward primer

Primer three –Reverse primer

Figure 3.9 Annealing sites of the three PCR primers to the human porphobilinogen deaminase gene. Primers sequences are in blue and template DNA in black. Translation initiation and termination sites are indicated in bold.

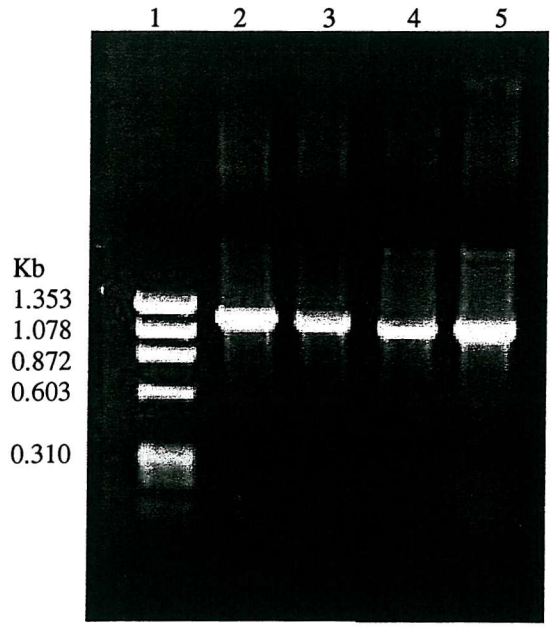


Figure 3.10 Analysis of the human PBGD cDNA ECPCR products by agarose gel electrophoresis.

Lane 1 = φX174 DNA standard; Lanes 2/3 = amplified 1142 bp fragment encoding human ubiquitous PBGD; Lanes 4/5 = amplified 1090 bp fragment encoding human erythroid PBGD (Dr. Sarwar).

3.2.4.3 Cloning of the gene fragments into vector pT7-7

ECPCR produced amplified fragments of sizes 1142 and1090 bases, for the ubiquitous and erythroid coding sequences respectively. The DNA fragments were electrophoresed on a 1% agarose gel and the required bands were extracted from the gel using Geneclean^{TM} (figure 3.10). The PCR products were resuspended and restricted with *Nde*I and *Bam*HI before ligating into the expression vector pT7-7 (under the control of the T7 bacteriophage promoter) (figure 3.11). The recombinant plasmids were transformed into DH5α and, again selected on LB agar plates containing ampicillin (100μg/ml). After identification of positive clones, checked by restriction analysis (figure 3.12), the plasmids harbouring the non-erythroid and erythroid sequences were transformed into expression strain BL-21(DE3) (Dr. Sarwar). This strain contains a single copy of the gene for T7 RNA polymerase under the control of the inducible lacUV5

promoter. This is essential for high-level expression of the genes under the control of the T7 promoter.

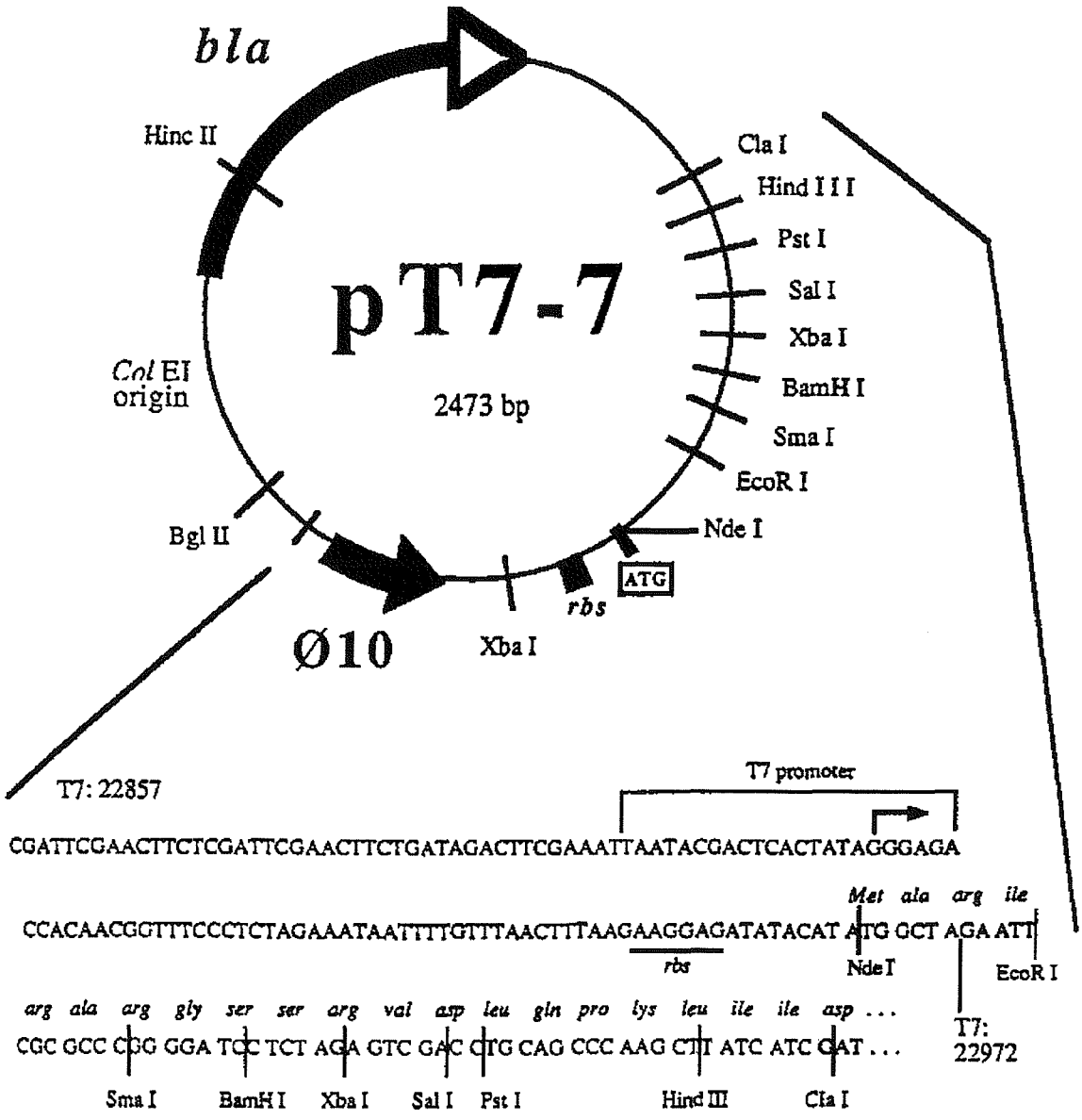


Figure 3.11 Map of the plasmid pT7-7. The ECPCR fragment was ligated between the sites *Nde*I / *Bam*HI in the 5' to 3' orientation. The inserted cDNA is in frame with the ATG start codon of the vector DNA and the optimal distance from the ribosome binding site (rbs), both being critical for efficient translation of the transcribed mRNA. The *bla* gene confers ampicillin resistance, allowing selection of recombinant plasmids.

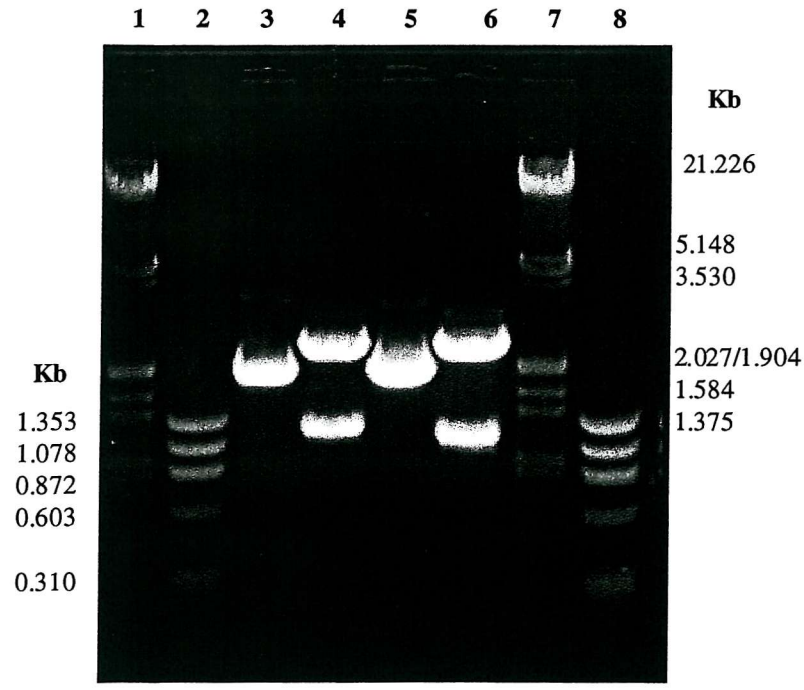


Figure 3.12 Restriction analysis of the recombinant plasmids encoding the cDNA specifying the erythroid and ubiquitous PBGD. Lane 1 = Lambda *Hin*dIII / *Eco*RI DNA standard; Lane 2 = φX174 DNA standard; Lane 3 = undigested 3.6 kb recombinant plasmid encoding ubiquitous PBGD; Lane 4 = recombinant plasmid encoding ubiquitous PBGD digested with *Nde*I and *Bam*HI producing 2.5kb and 1.142 kb fragments; Lane 5 = undigested 3.6 kb recombinant plasmid encoding erythroid PBGD; Lane 6 = as lane 5, but digested with *Nde*I and *Bam*HI producing 2.5kb and 1.09kb fragments; Lane 7 = Lambda *Hin*dIII / *Eco*RI DNA standard; Lane 8 = φX174 DNA standard.

3.3 Purification of recombinant human porphobilinogen deaminase

3.3.1 Initial expression studies

Eleven colonies harbouring the recombinant plasmids were selected (6 erythroid, 5 non-erythroid) for characterisation by SDS-PAGE and activity assays. Cultures were grown for each clone, sonicated and centrifuged to give supernatant and pellet fractions (section 2.3.10). SDS-PAGE analysis of the supernatant and pellet fractions (figure 3.13 a and b) indicated that >50% of the required protein

was insoluble (due to the greater intensity of the bands in the pellet fractions). Despite this relative insolubility, activity measurements showed there was enough soluble enzyme for protein purification and possible crystallisation. The clone selected for purification of the non-erythroid enzyme is shown in lane eleven (both gels).

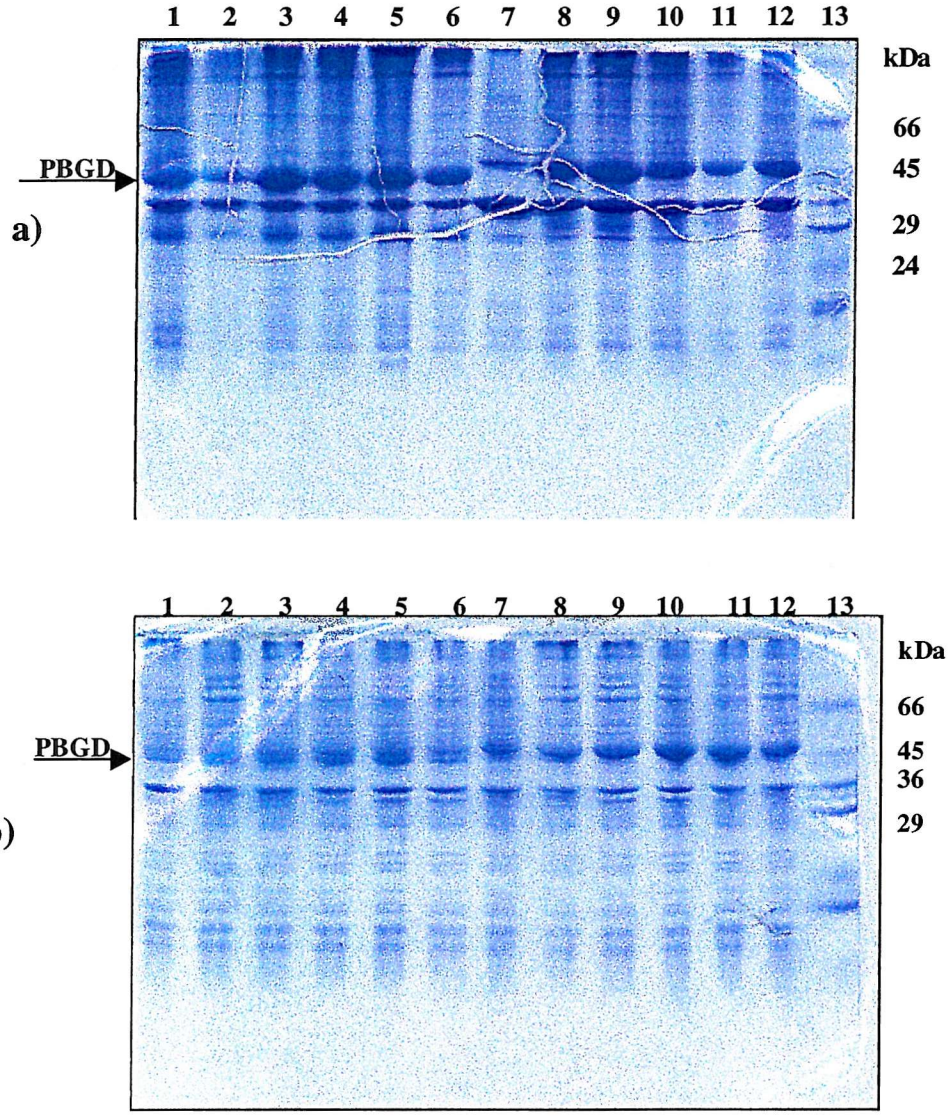


Figure 3.13 SDS-PAGE of selected strains of *E. coli* expressing recombinant erythroid and ubiquitous human PBGD. a) Supernatant fractions. Lanes 1 – 6 = erythroid clones; Lane 7 = plasmid pT7-7 (no insert); Lanes 8 – 12 = ubiquitous clones; Lane 13 = Dalton VII protein standards.

b) Pellet fractions. Lanes as for a).

3.3.2 Large-scale culture growth

Large-scale cultures were grown following confirmation of soluble protein expression. The scale-up proved to be reproducible and the bacterial cell pellets could be stored for up to one year at -20°C, without any significant loss of enzyme activity. A bacteriophage resistant strain of BL21-DE3 was isolated and used for expression to overcome loss of protein due to bacteriophage infection (section 2.3.12).

3.3.3 The purification of recombinant human ubiquitous PBGD

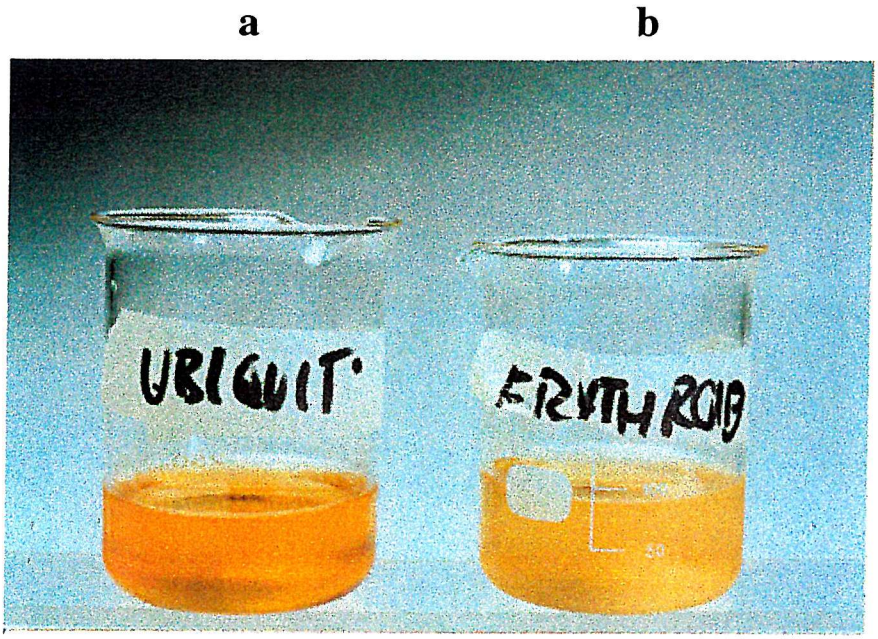
In order to characterise and crystallise both recombinant PBGD isoenzymes an effective and reproducible purification protocol was developed as follows.

3.3.3.1 Sonication

The bacterial pellet from a 5L culture was resuspensed in 100ml of 20mM Tris/HCl, pH8.2 containing 5mM DTT and 0.02% sodium azide. Sonication was optimal at 10 microns for 6 cycles (30 seconds on, 60 seconds off), as deduced by time course experiments. The addition of 0.2mM PMSF sufficiently inhibited any cellular proteases released and immersing the sonication vessel in an ice/salt mixture prevented excessive overheating of the cell extract and hence protein denaturation.

3.3.3.2 Heat step and ultracentrifugation

Subsequent heat treatment of the crude extract was tested as a concept was taken from the purification of the in-vivo enzyme (Mazzetti and Tomio, 1993). Heating the sonicated extract at 60°C for 10 minutes, followed by rapid cooling on ice to 4°C precipitated a significant amount of the contaminating proteins. Ultracentrifugation at 40,000 rpm/4°C for one hour pelleted the precipitate yielding a clear supernatant containing soluble human porphobilinogen deaminase (figure 3.14). It was observed that the supernatant from the strain



UBIQUITOUS ERYTHROID

Figure 3.14 Supernatant fractions after ultracentrifugation of *E. coli* strains expressing the recombinant human ubiquitous and erythroid PBGD enzymes. a) Ubiquitous extract and b) erythroid extract.

Figure 3.15 Pellet fractions after ultracentifugation of a) Ubiquitous extract b) Erythroid extract.

expressing the ubiquitous enzyme was more intensely coloured in comparison to the supernatant from the strain expressing the erythroid enzyme (figure 3.14). Also, the erythroid pellet contained a large creamy coloured precipitate, which commonly indicates the presence of inclusion bodies (figure 3.15).

3.3.3.3 Column chromatography

Various chromatographic materials were tested for binding the recombinant PBGD isoenzymes and separation from contaminating proteins.

DEAE-Sephacel anionic exchange column chromatography

DEAE-Sephacel, an anionic exchanger, was selected as it has desirable properties being a robust material that is readily available and inexpensive. Arrays of different buffers, including Tris/HCl, phosphate and MOPS at different pH values, were evaluated as described in section 2.3.13. These results were further refined producing the following effective capture step. Firstly a Pharmacia XK50/20 column (50 x 18mm) was packed with DEAE Sephacel to a bed volume of 150mL. The column was equilibrated with 10 column volumes (1.5 litres) of 20mM Tris/HCl pH8.2, containing 5mM DTT and 0.02% azide at a flow rate of 2mL/minute. The supernatant was carefully loaded onto the column at a flow rate of 1mL/minute and washed with approximately 10 column volumes of buffer, or until the OD_{280nm} of the elute was less than 0.1 (compared to a blank of buffer). This ensured that any unbound proteins had been washed off the column and maximised protein separation. A gradient of KCl (0 to 70mM) was then applied to the column, before eluting the desired protein isocratically with 20mM Tris/HCl, containing 1mM DTT and 70mM KCl. The protein was eluted over a volume of 300-400 ml. The fractions containing enzyme were identified by SDS-PAGE (figure 3.16) and pooled before concentrating to 10ml in an Amicon ultrafiltration cell fitted with a PM10 membrane. The protein solution was further concentrated using a Centricon 10,000 to a volume of 1mL (centrifuged at 5000 rpm).

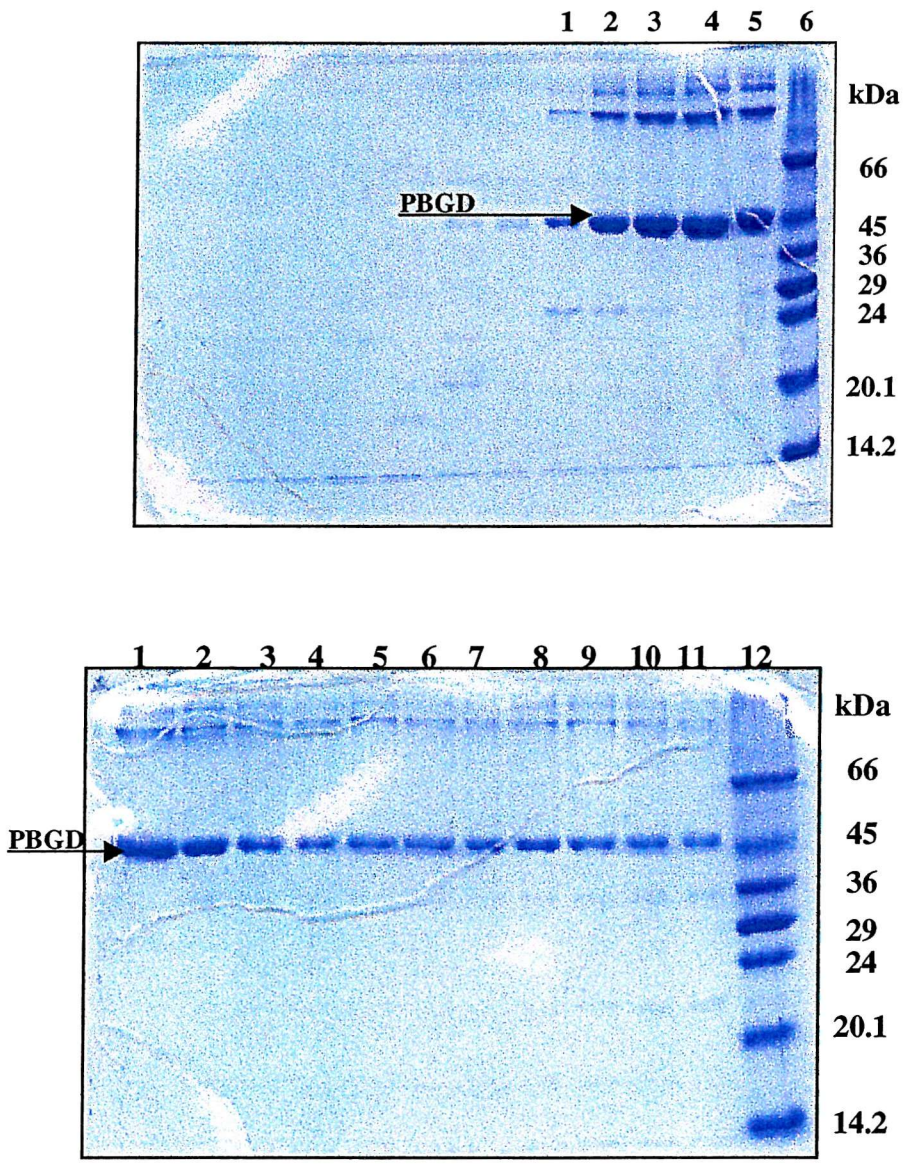


Figure 3.16 SDS-PAGE analysis of fractions of recombinant human PBGD eluted from a DEAE Sephacyl column. Gel a) Lane 1 identifies the first fraction containing ubiquitous PBGD; Lanes 2 – 5 show additional fractions containing enzyme; Lane 6 – Dalton VII protein standard. Gel b) Lanes 1 – 11 show the later fractions, containing ubiquitous PBGD, eluted from the DEAE-Sephacel column.

Hi-load 16/60 Superdex-75 gel filtration

After the first column, the concentrated protein sample was still impure (figure 3.22), so an additional column step was required. Various materials were tested for protein separation including phenyl-sepharose, hydroxyapatite, G-100 Sephadex, and G-200 Superdex. The gel-filtration column Superdex-75, used in conjunction with a Pharmacia FPLC system, gave the best resolution. Gel filtration chromatography works on the principle of size exclusion, with larger molecules leaving the column first followed by the smaller molecules in the order of their sizes. Consequently, for a given column, the elution volume of a protein can be determined from its size and *vice versa*, provided there are no ionic interactions between the protein molecules and the gel matrix. The elution volume of human uPBGD from the Superdex-75 gel filtration column was significantly larger that expected when eluted with 20mM Tris/HCl pH 8.2. Increasing the ionic strength of the buffer to 100mM Tris/HCl, pH8.2 gave the expected results and allowed efficient separation of the human ubiquitous PBGD from the contaminants. These results suggest that at the lower ionic strength there are interactions between the enzyme and gel matrix, resulting in retention of the enzyme and a subsequent increase in the elution volume. After equilibration of the column in 100mM Tris/HCl, pH8.2, containing 5mM DTT, the 1mL of concentrated enzyme solution was loaded at a flow rate of 0.5mL/minute and eluted isocratically. A UVMII monitor detected protein and fractions were selected from the elution profile (figure 3.17) and analysed by SDS-PAGE (figure 3.18).

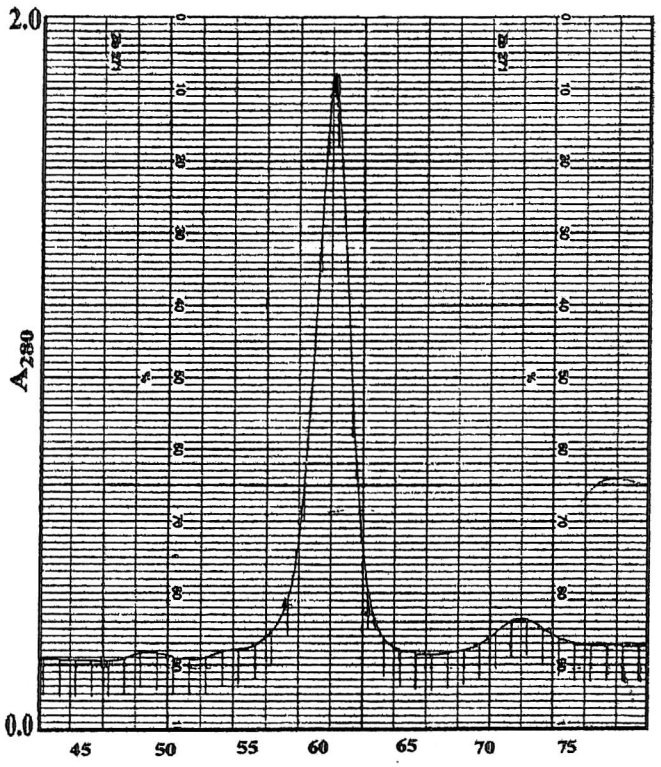
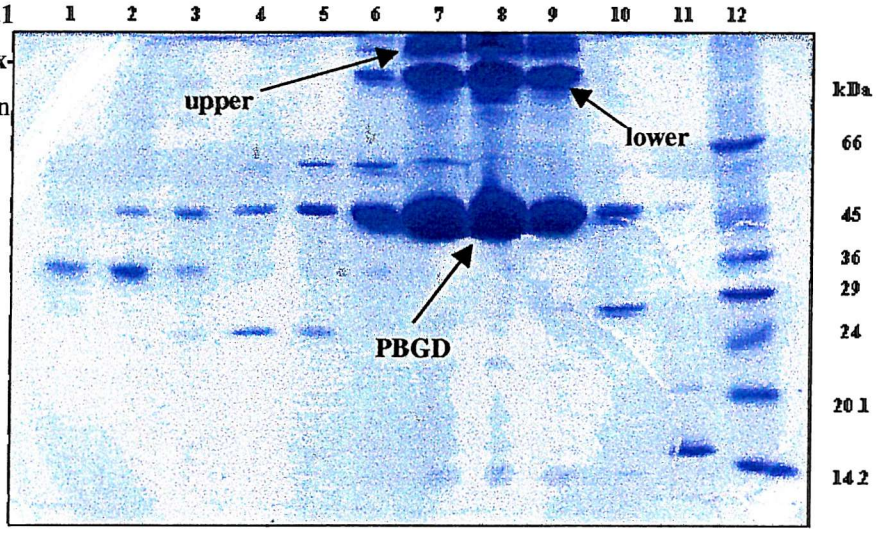


Figure 3.17 Elution profile of recombinant ubiquitous PBGD from a Superdex-75 gel filtration column. The main peak corresponds to uPBGD.

Figure 3.1 n
Superdex- ı in
the elution



3.3.3.4 Identification of the co-purified contaminating proteins

The SDS-PAGE analysis of the gel filtration fractions (figure 3.18) shows the major impurities are the two upper bands (at the top of the gel). It was suspected that these were not contaminating proteins, but aggregation /degradation products of the PBG-deaminase (as they appeared to be proportional to the intensity of the PBGD bands and could not be separated). These additional bands were unlikely to be apoenzyme (enzyme without the cofactor), as this would not bind efficiently to the DEAE-Sephacel column and hence be eluted before the native enzyme. To answer the question an additional analysis was undertaken. A standard 12% SDS-PAGE gel was run stained with 0.3M $CuCl_2$ and the two upper bands excised. After extraction of the gel fragments with acetonitrile they were reloaded onto a pre-prepared 12% SDS gel and electrophoresed. Both bands (termed unknown upper and lower, relative to each other) regenerated a band in an equivalent position to ubiquitous PBGD (figure 3.19). This suggests, as suspected that these bands were not contaminants but aggregated or denatured PBGD.

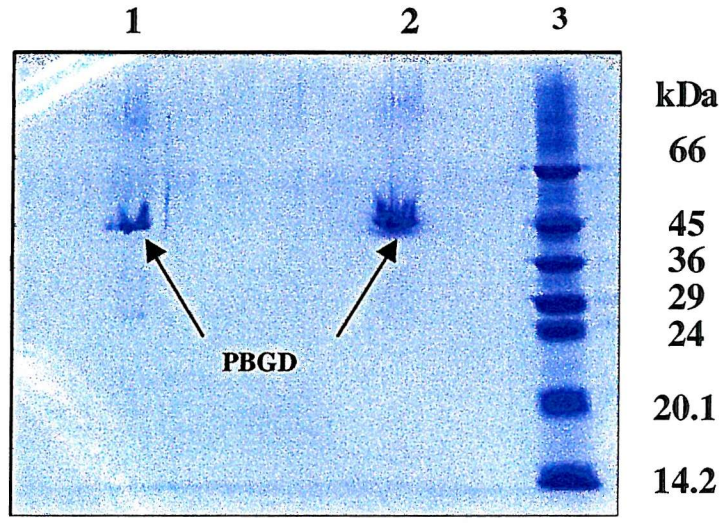


Figure 3.19 SDS-PAGE showing the regeneration of recombinant human ubiquitous PBGD from the two upper bands shown on SDS-PAGE after gel filtration. Lane 1 – unknown upper band reloaded. Lane 2 – unknown lower band reloaded. Land 3 – DaltonVII protein standard.

Through the purification procedure the fractions containing enzyme were identified by SDS-PAGE and activity assays. The entire procedure allowed the recombinant ubiquitous PBGD to be purified >90% as summarised in figure 3.20 and table 3.3. Analysis of the different steps of purification process shows that the heat treatment removed a considerable amount of proteins, confirming the thermostability of human ubiquitous PBGD. DEAE-ion exchange and subsequent concentration resulted in 3.6 fold increase in specific activity.

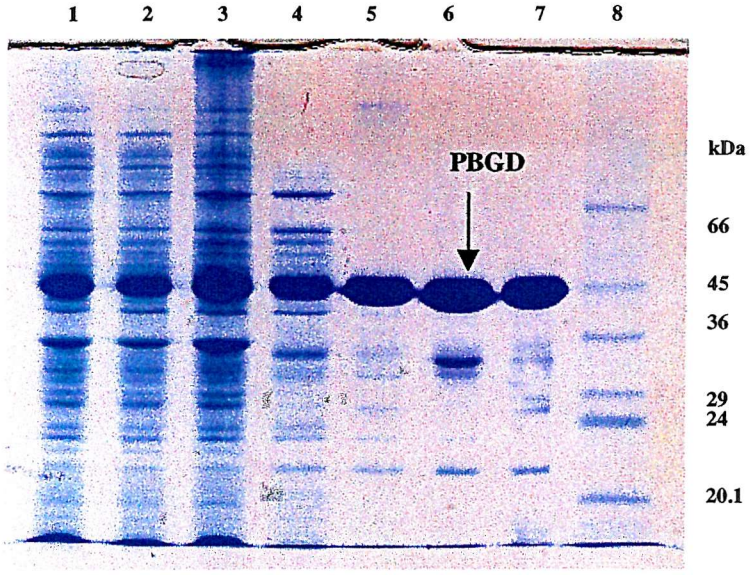


Figure 3.20 SDS-PAGE of samples taken from each stage of the recombinant human ubiquitous PBGD enzyme purification. Lane 1= resuspended cells, Lane 2= sonicated extract, Lane 3= heat-treated extract, Lane 4= supernatant after heat treatment and ultracentrifuging, Lane 5= after DEAE-Sephacyl column, Lane 6/7= after gel filtration, Lane 8= Dalton VII protein standards.

	[Protein] mg/ml	Total volume (ml)	Total protein (mg)	Activity (μmoles/ml/hour)	Specific activity (μmoles/mg/hour)	Total activity (μmoles/hour)	Yield (%)
Pellet resuspended	10.24	105	1075.2	2.442	0.238	256.41	94.1
Sonicate	17.199	101	1737.1	2.698	0.156	272.5	100
Heat step	13.34	93	1240.62	2.556	0.192	237.7	87.2
Supernatant	4.6	82	377.4	2.65	0.576	217.3	79.7
DEAE-Sephacel	0.156	415	64.7	0.327	2.1	135.71	50
Gel filtration	2.9	10	29	3.97	1.4	39.72	14.6

Table 3.3 Purification of recombinant human ubiquitous PBGD. Activity is measured in μmoles of product formed.

3.3.3.5 Studies on the heterogeneity of purified recombinant human ubiquitous PBGD

After gel filtration SDS-PAGE shows the PBGD protein is at least 90% pure. Mass spectrometry indicates one species and N-terminal sequencing gave one sequence, in concordance with a single protein (see chapter 4). However, analysis by native-PAGE indicates a heterogeneous solution containing three different species (figure 3.21). The different species were separated by high-resolution anionic exchange chromatography using a Pharmacia MonoQ/HR column (figure 3.22). Analysis of the separated peaks by native-PAGE showed all three species had a different electrophoretic mobility increasing in the order 6>7>8. Initially it was postulated that this heterogeneity could be attributed to a series of enzyme-substrate (ES) intermediates in concordance with the altered electrophoretic mobility. Subsequent investigations taking the two main peaks after MonoQ separation and incubating each with substrate to examine the formation of the other enzyme: substrate intermediates suggested this postulation was incorrect. MonoQ analysis showed three additional peaks to the right of the original peak representing ES, ES_2 and ES_3 suggesting that the heterogeneity of purified recombinant human ubiquitous PBGD is not due to different ES intermediates. Additional evidence for this was provided by experiments performed by Danica Butler (personal communication), where the two samples were reacted with hydroxylamine. This compound would cause the release of any bound pyrrole molecules from the cofactor, so an additional band would have been observed on a native polyacrylamide gel. The absence of an additional band confirms the absence of any enzyme: substrate intermediate; however the exact difference remains to be elucidated. A similar phenomenon was observed with the *E. coli* enzyme (Jordan *et al.*, 1988a).

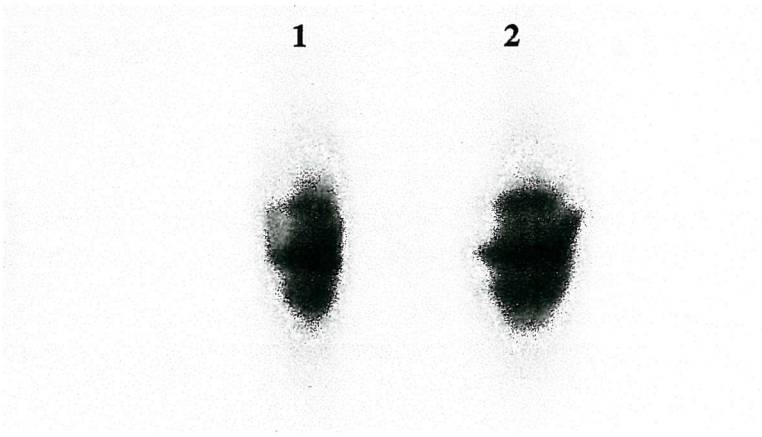


Figure 3.21 Native PAGE analysis of purified recombinant human ubiquitous PBGD. Lane1 and lane 2 are identical samples of purified human ubiquitous PBGD. The presence of three species indicates that the purified protein is heterogeneous.

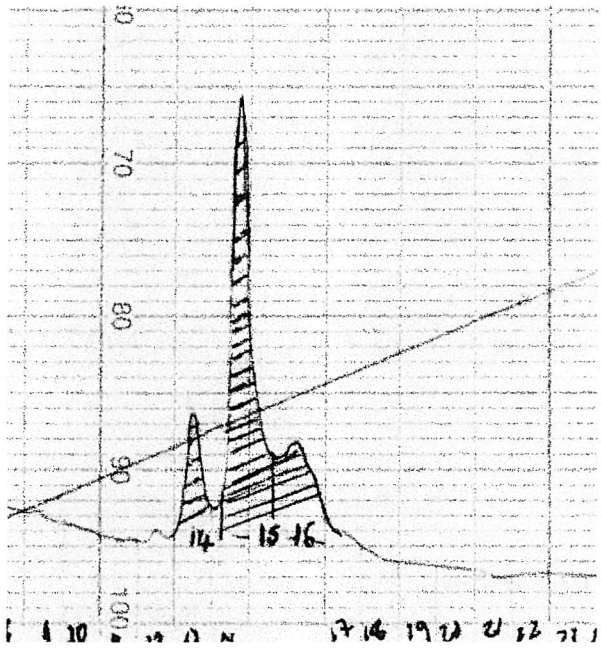


Figure 3.22 Elution profile of purified recombinant human ubiquitous PBGD from a MonoQ/HR5 column. Three main species corresponding to fractions 14, 15 and 16 can be seen. However, the fractions 15 and 16 are not completely resolved.

3.3.4 Purification of recombinant human erythroid porphobilinogen deaminase

The erythroid enzyme was purified using the same method as for the ubiquitous enzyme. Figure 3.23 shows the fractions eluted from the DEAE-Sephacel anionic exchange step, which were pooled and concentrated to a volume of 1mL before loading onto a Superdex-75 gel filtration column. The erythroid enzyme was eluted from the gel filtration column isocratically in 100mM Tris/HCl, pH8.2 containing 5mM DTT (figure 3.24) and the fractions concentrated for crystallisation. The purification of recombinant human erythroid PBGD is summarised in figure 3.25 and table 3.4. As for the ubiquitous enzyme, the protocol allowed the erythroid isoform of PBG-deaminase to be purified >90%. Heat treatment precipitated approximately 80% of contaminating proteins resulting in a 3.5 fold increase in specific activity. DEAE-ion exchange and subsequent concentration resulted in an additional 3.6 fold increase in specific activity. After gel filtration SDS-PAGE shows the PBGD protein is at least 90% pure. Mass spectrometry indicates one species and N-terminal sequencing gave one sequence, in concordance with a single protein (see chapter 4).

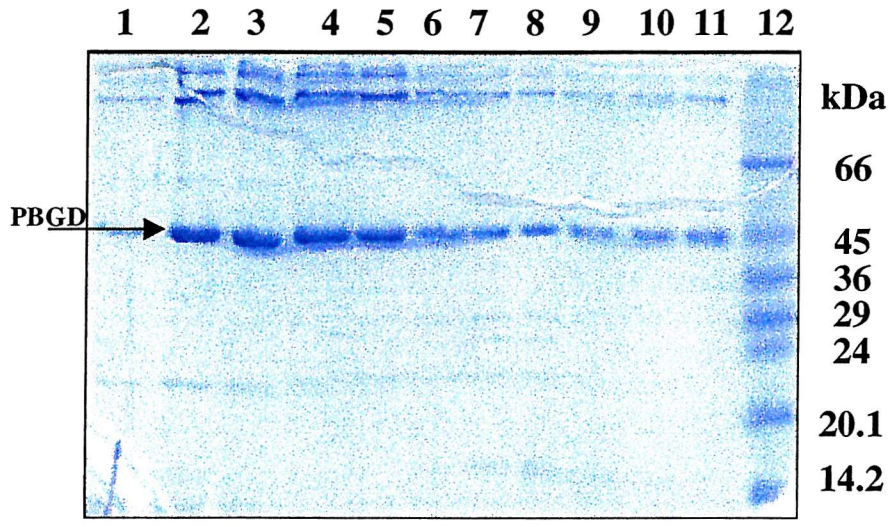


Figure 3.23 Recombinant human erythroid PBGD fractions eluted from DEAE – Sephacel shown by SDS–PAGE. Lane 1 identifies the first fraction containing erythroid PBGD, Lanes 2 – 11 show additional fractions containing enzyme, Lane 12 – Dalton VII protein standard.

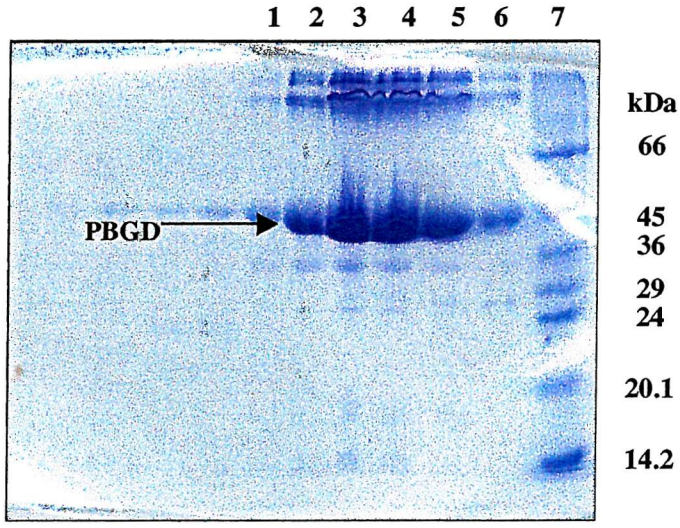


Figure 3.24 SDS-PAGE of recombinant human erythroid PBGD fractions eluted from Superdex 75 gel filtration column. Lanes 1-6 correspond to the main peak seen on the elution profile. These fractions were pooled and concentrated before crystallising. Lane 7= Dalton VII protein standard.

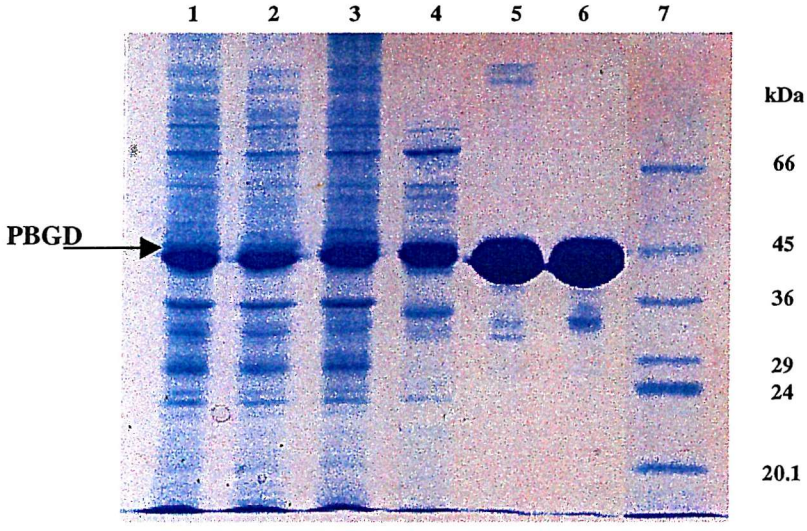


Figure 3.25 SDS-PAGE of samples taken from each stage of the purification of recombinant human erythroid PBGD. Lane 1= resuspended cells, Lane 2= sonicated extract, Lane 3= heat-treated extract, Lane 4= supernatant after heat treatment and ultracentrifuging, Lane 5= after DEAE-Sephacel column, Lane 6= after gel filtration, Lane 7= Dalton VII protein standards.

	[Protein] mg/ml	Total volume (ml)	Total protein (mg)	Activity (μmoles/ml/hour)	Specific activity (μmoles/mg/hour)	Total activity (μmoles/hour)	Yield (%)
Pellet resuspended	17.2	80	1216	8.5	0.49	680	88.5
Sonicate	28.2	80	2256	9.6	0.34	768	100
Heat step	21.3	74	1579.2	8.9	0.42	659	85.8
Supernatant	7.37	63	464.31	8.7	1.18	548	71.4
DEAE-Sephacel	0.2	425	85	0.85	4.25	361.25	47
Gel filtration	0.8	28.6	22.9	4.0	5.0	114.4	15

Table 3.4 Purification of recombinant human erythroid PBGD. Activity is measured in μmoles of uroporphyrin product formed

Conclusions

This chapter has described the construction of two different plasmids harbouring the cDNAs specifying ubiquitous and erythroid porphobilinogen deaminases. Using expression cassette PCR, two segments of DNA were generated, each with conveniently located restriction sites to allow direct ligation into the plasmid pT7-7 for expression in BL-21. The PCR primers also contained nucleotides specifying codons for the initial amino acids of the deaminases that are used in highly expressed proteins. The DNA in the plasmids thus obtained was expressed in strains of *E. coli* yielding milligramme amounts of recombinant ubiquitous and erythroid human deaminases (Dr. Sarwar).
Purification protocols were designed for both ubiquitous and erythroid deaminases yielding homogeneous enzymes that were suitable for crystallographic trials. The specific activity of purified ubiquitous deaminase was 1.4μmoles/hr/mg, significantly higher than the previously published figure. The specific activity of the erythroid enzyme was 5μmoles/hr/mg, twice that previously reported. The characterisation and crystallisation of these two recombinant human porphobilinogen deaminases are discussed in chapter 4.

Chapter 4 Characterisation and Crystallographic Studies on the recombinant human PBGD Isoenzymes

4.1 Introduction

The X-ray structure of porphobilinogen deaminase was first determined in 1992 using the *E. coli* enzyme (Louie *et al.*, 1992) and subsequently refined to a resolution of 1.76A (Louie *et al.*, 1996). The structure shows the deaminase to be a protein with three domains, each of about 100 aminoacids. Domains 1 and 2 provide a flexible structure between which the active site is located, with domain 3 providing the attachment point for the dipyrromethane cofactor. This three-dimensional structure has greatly increased our understanding of how the enzyme is folded and exactly how it accommodates the dipyrromethane cofactor. In addition the structure has provided insight as to how the enzyme recognises the substrate porphobilinogen and catalyses the polymerisation reaction. The structural and functional properties of several mutations made in the *E. coli* deaminases (Jordan and Woodcock, 1991; Lander *et al.*, 1991; Woodcock and Jordan, 1994) have also been made possible using this structure.

A comparison of the primary structures of all porphobilinogen deaminase enzymes reveals that there is a remarkable similarity between all deaminases suggesting that it is an evolutionarily ancient enzyme. For instance there is as much as 60% similarity between the deaminases from *E. coli* and human sources, indicating that the structural similarity extends to secondary and tertiary structure. With this similarity in mind, the X-ray structure of the *E. coli* enzyme has been used to map the position of several human mutations (Wood *et al.*, 1995). Coupled with *in vitro* experiments involving the expression and characterisation of human mutant deaminases, a clearer picture is being built up of structure functional relationships in the human enzyme, particularly with respect to the CRIM status of AIP associated mutations.

Despite the similarity between the *E. coli* and human deaminases, there are still substantial differences in the sequence of amino acids, particularly in domain 3 where there is an insertion of 29 amino acids. A model of the secondary structure of human PBGD, highlighting this difference, has been generated by overlaying the structure of *E. coli* enzyme with the exon boundaries of the human gene (figure 4.1). This predicts that the insertion loop in domain 3 follows strand $\beta 2_3$. However, this is not conclusive, as there could be a number of plausible models for this region. It was therefore important to determine the structure of the human deaminase independently so that the particular structural and functional characteristics could be determined. Additionally, this would allow the precise effects of the human mutations that lead to acute intermittent porphyria (AIP) to be evaluated. The previous chapter in this thesis described the cloning, expression and purification of recombinant human ubiquitous and erythroid porphobilinogen deaminase. This chapter describes the characterisation of the ubiquitous and erythroid enzymes and attempts to obtain crystals for determination of their X-ray structures.

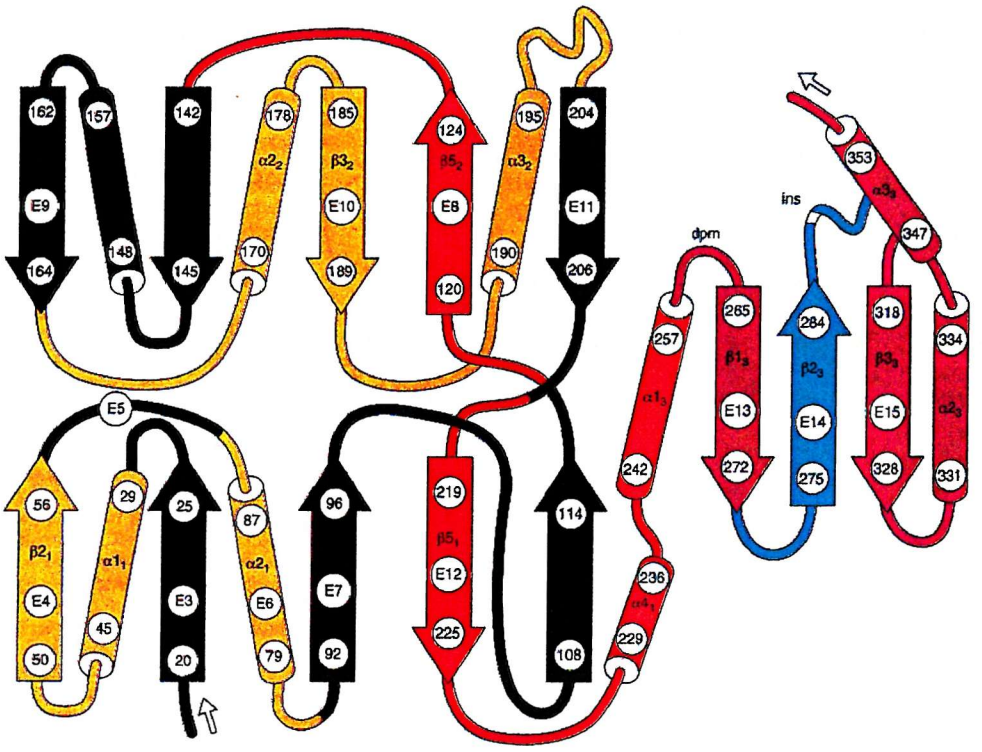


Figure 4.1 The topology of the secondary structure of human porphobilinogen deaminase, modelled from the *E. coli* enzyme, is overlaid with the exon boundaries of the human gene (depicted by a change of colour). Each regular element is labelled by the human sequence number, exon number and descriptor; for example, $\alpha1_1$ is the first helix subscript domain 1. Dpm is the cofactor site and ins denotes the domain 3 insertion of the human sequence relative to *E. coli* (reproduced from Wood *et al.*, 1995).

4.2 Characterisation of recombinant human ubiquitous and erythroid PBGDs

In chapter 3 the purification of recombinant ubiquitous and erythroid PBGDs were described. Figure 4.2 below shows the purified ubiquitous and erythroid enzymes respectively. From the gel the Mr of the ubiquitous enzyme was 44,000 and the erythroid enzyme was 42,000. This is somewhat different from the theoretical values of 39,751 and 38120 predicted from the cDNA sequences (table 4.1). It was also evident that high molecular weight species were present that were suspected of being aggregates of the monomeric enzymes. It was therefore necessary to carry out further experiments to characterise the two enzymes by N-terminal sequencing and accurate mass determination by electrospray mass spectrometry.

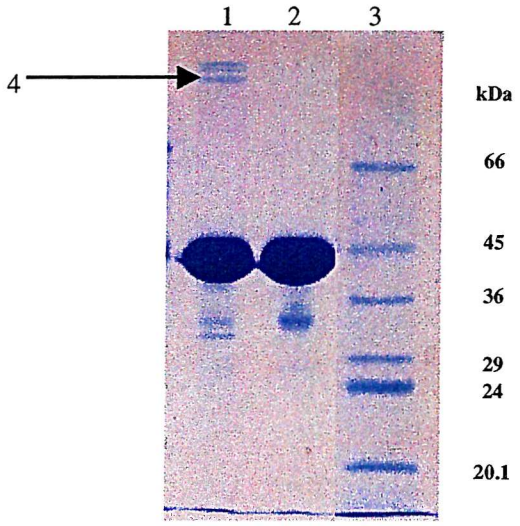


Figure 4.2 SDS-PAGE of the purified recombinant human ubiquitous and erythroid PBGDs. Lane 1 = purified ubiquitous enzyme, Lane 2 = purified erythroid enzyme, lane 3 = Dalton VII protein standards and 4 = higher molecular weight aggregates.

	Mr from SDS-PAGE	Mr from cDNA sequence	+ cofactor = 420.4 mass units
Erythroid PBGD	44,000	37,699	**38,120**
Ubiquitous PBGD	42,000	39,330.6	**39,751**

Table 4.1 Comparison of the molecular weights of recombinant human PBGDs determined from the cDNA sequence and by SDS-PAGE. The values differ significantly indicating that additional analysis is required to confirm the particular characteristics of each isoenzyme.

4.2.1 N-Terminal amino acid sequencing of recombinant ubiquitous and erythroid PBGDs

N-terminal sequences were determined using the Edman degradation method. After separation by SDS-PAGE the protein were transferred onto a PVDF membrane by Western blotting (figure 4.3) prior to sequencing as described in section 2.3.8. Samples 1 and 2, corresponding to erythroid PBGD and its aggregates, both gave the N-terminal sequence **M R V I R V** which is identical to the published sequence of erythroid PBGD (Raich *et al.*, 1986). Samples three and four, corresponding to ubiquitous PBGD and its aggregates had the same N-terminal sequence **S G N D G N A** which correlates with the Protein Database sequence (PDB) for human ubiquitous PBGD (Grandchamp *et al.*, 1987), except the N-terminal methionine is absent. This clipping of the initial methionine residue is a common phenomenon for eukaryotic proteins expressed in *E. coli.* Primarily these results provide evidence that the human ubiquitous and erythroid PBGDs have been isolated. Additionally, because the upper bands, suspected of being aggregates, have the same N-terminal sequence, the isolated proteins are homogeneous, free from contaminating proteins.

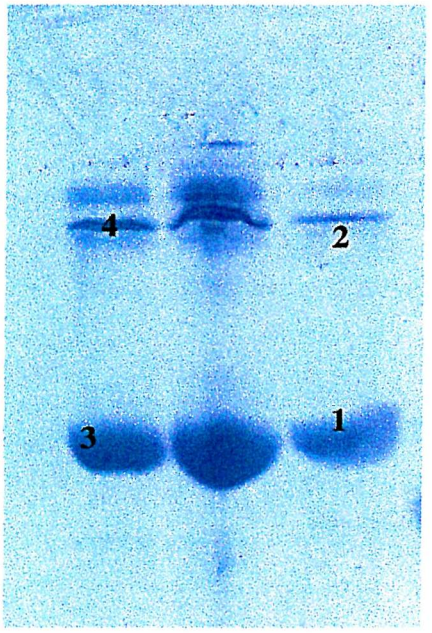


Figure 4.3 Western blot of recombinant ubiquitous and erythroid PBGD. 1 = erythroid PBGD, 2 = erythroid PBGD aggregates, 3 = ubiquitous PBGD and 4 = ubiquitous PBGD aggregates. All four bands were excised from the nitrocellulose membrane and used to determine the N-terminal sequence by Edman degradation (Paul Skipp, Southampton University).

4.2.2 Mass spectrometry studies on recombinant human porphobilinogen deaminase

Mass spectrometry was used to determine accurate molecular masses for the human erythroid and ubiquitous PBGD isoenzymes. Also, the reaction of each isoenzyme with the natural substrate porphobilinogen was analysed by ESMS. Optimal spectra were produced for native protein by mixing 50µL of enzyme at 0.2mg/mL with 50µL 50% acetonitrile/ 50% H_2O/ 1% formic acid. For production of enzyme: substrate intermediates 100µL of enzyme at 0.2mg/mL was incubated with 2.6µL PBG (1mM), reacted by vortexing, then immediately added to 100µL 50% acetonitrile/ 50% H_2O/ 1% formic acid. It was essential that all samples were freshly prepared and analysed immediately to produce accurate and reproducible results.

The theoretical molecular mass of the recombinant ubiquitous isoenzyme, 39,751, differs from the mass obtained by ESMS by 138 mass units (mu) (figure 4.3). This is easily explained in the light of N-terminal sequencing results, which show that the

N-terminal methionine is clipped, accounting for theoretical reduction of 131mu. The spectra obtained for recombinant erythroid porphobilinogen deaminase (figure 4.5) agrees with the theoretical molecular mass (table 4.1). Correlation between the theoretical and measured masses confirms that the correct gene expression has occurred and that the chemical integrity of the endogenous and recombinant proteins are identical, except the first methionine residue is missing from recombinant ubiquitous PBGD.

Comparison of the spectra obtained after reaction of the recombinant ubiquitous PBGD with substrate (figure 4.6) with the expected results (table 4.2) confirm the formation of the enzyme intermediates ES, ES_2 and ES_3. Reaction of the recombinant erythroid enzyme gave similar results (figure 4.7). From these observations it can be concluded that both recombinant isoenzymes have been purified in an active form. The results also support previous findings that the mechanism of the human enzyme, as for other deaminases, proceeds via the formation of the intermediates ES, ES_2 and ES_3 (Desnick *et al.*, 1985).

Recombinant protein	**Erythroid PBGD**	**Ubiquitous PBGD**
Native mass	38,120	39,620
Mass ES (+ 1 PBG)	38,330	39,830
Mass ES_2 (+2 PBG)	38,540	40,040
Mass ES_3 (+3PBG)	38,750	40,250

Table 4.2 The theoretical molecular masses of the recombinant ubiquitous and erythroid PBGDs and their corresponding enzyme: substrate intermediates. Comparison with the determined masses (figures 4.4 – 4.7) confirmed the isolation of the recombinant enzymes and the formation of enzyme: substrate intermediates.

A

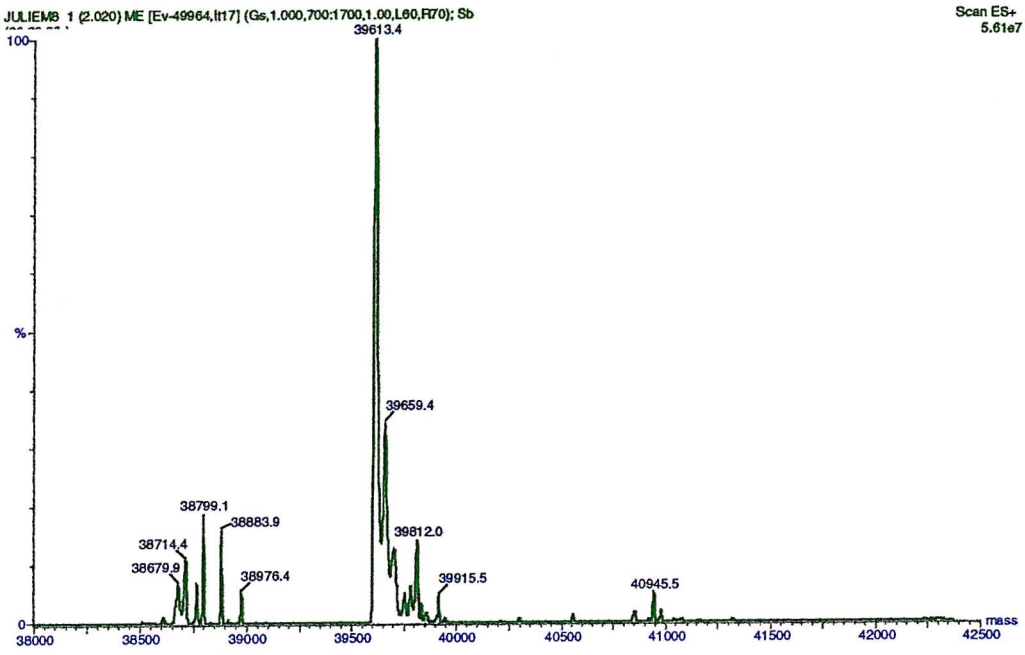
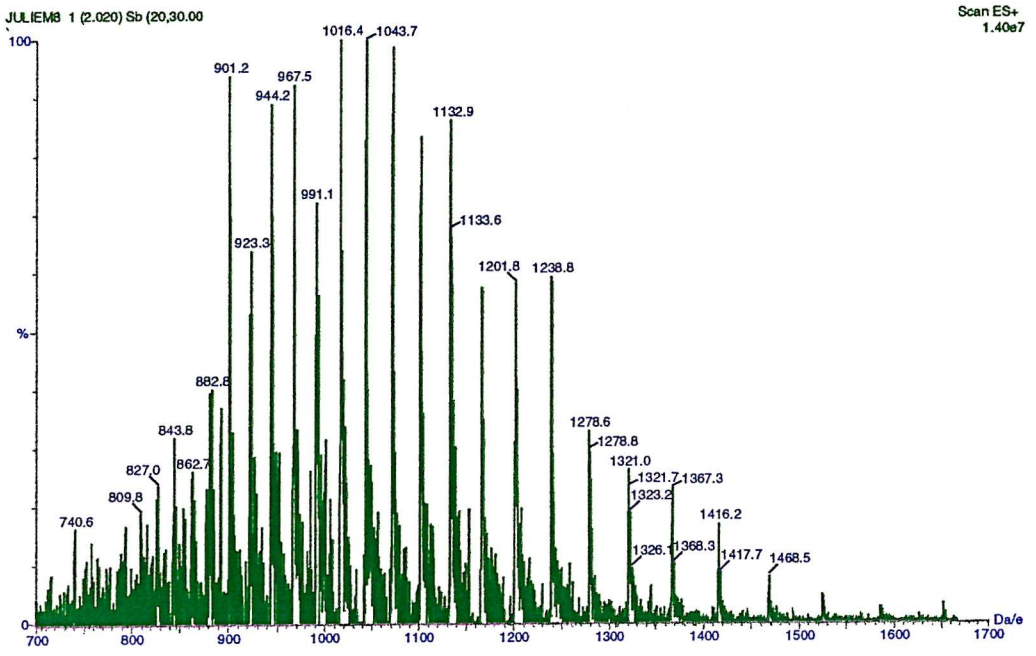


Figure 4.4 A) Maximum entropy deconvoluted mass spectrum of recombinant ubiquitous PBGD. The main peak at 39,613.4, has an error of 0.017% in comparison to the predicted molecular mass of 39,620 Da. **B) Corresponding raw data.**

B



A

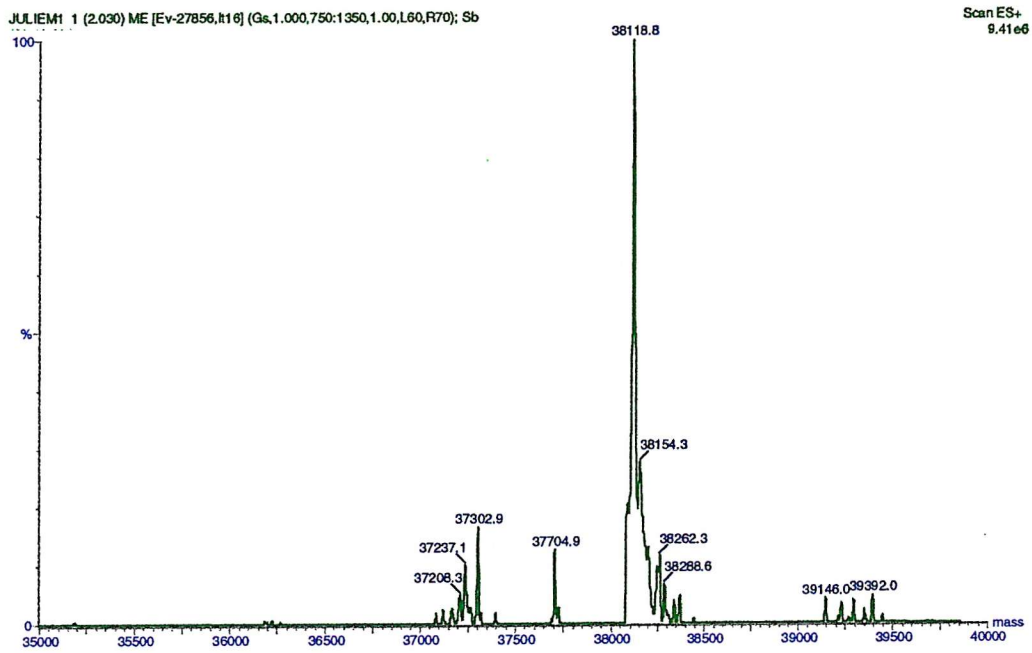
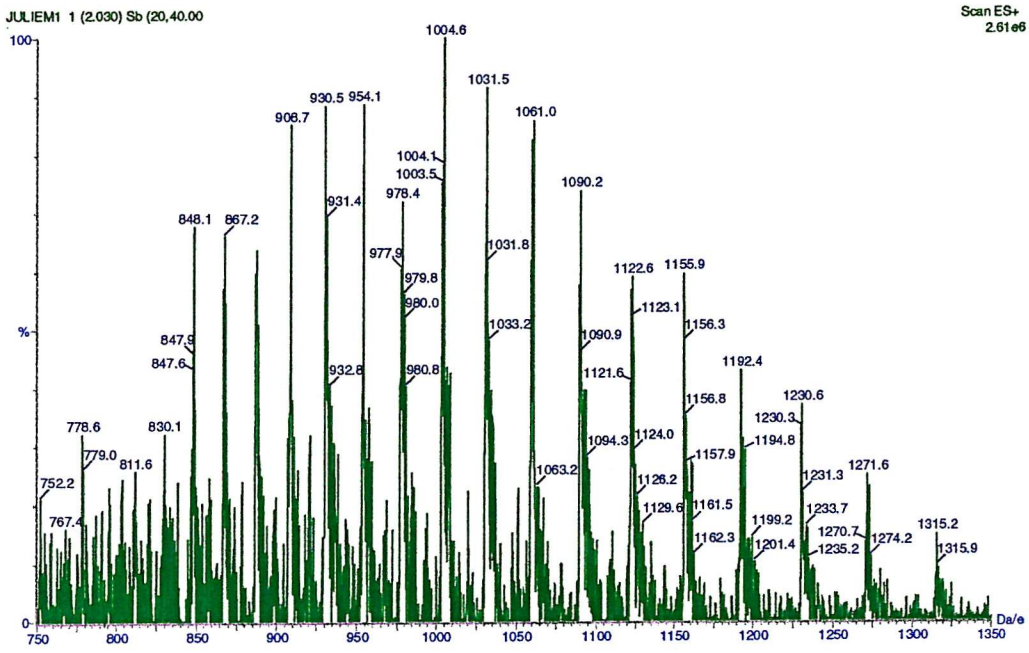


Figure 4.5 A) Maximum entropy deconvoluted mass spectrum of recombinant erythroid PBGD. The main peak at 38,118.8 has an error of 0.003% well within the accuracy value of 0.01%. **B) Corresponding raw data.**

B



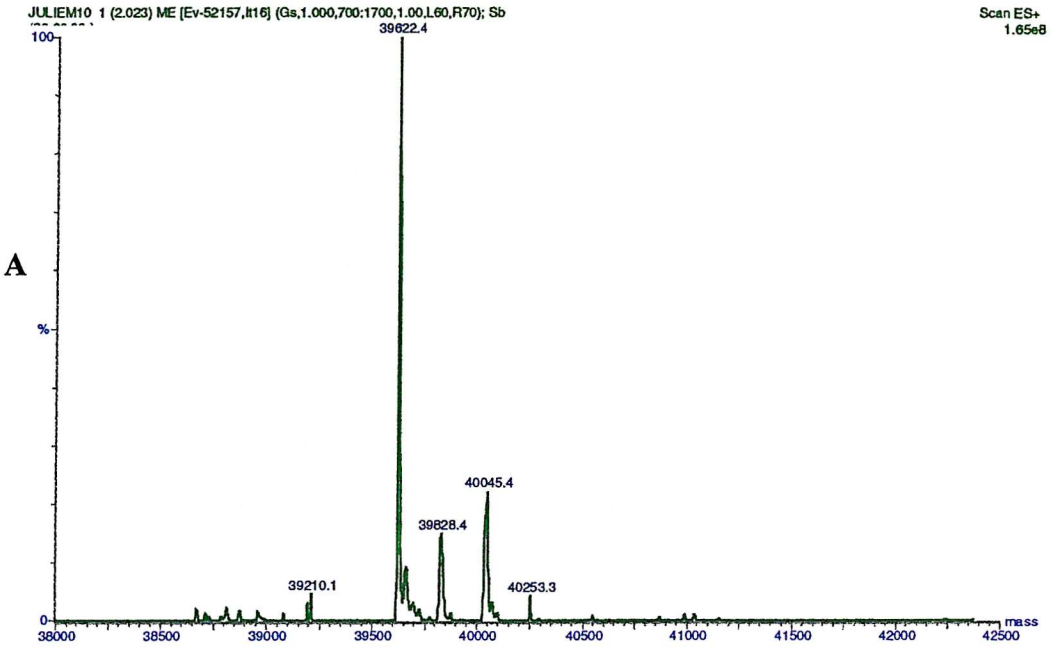
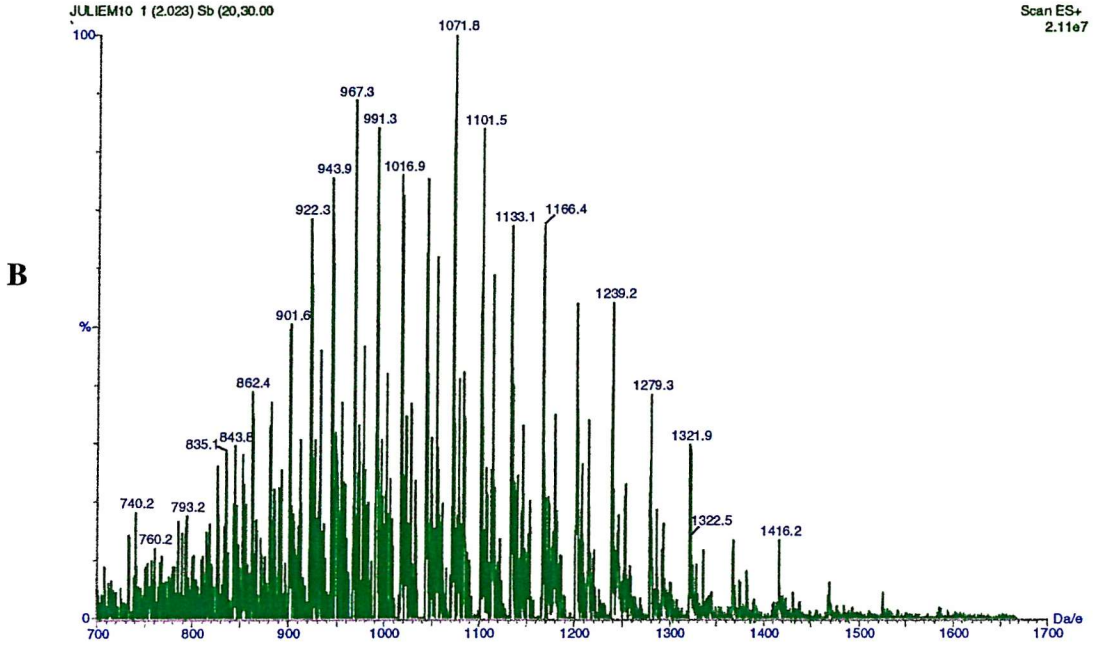


Figure 4.6 A) Maximum entropy deconvoluted mass spectrum of recombinant ubiquitous PBGD mixed with a five molar excess of PBG, its natural substrate. Peaks observed at 39,622.4, 39828.4, 40045.4, 40253.3 correspond to enzyme E, ES, ES_2, and ES_3 respectively. **B) Corresponding raw data.**



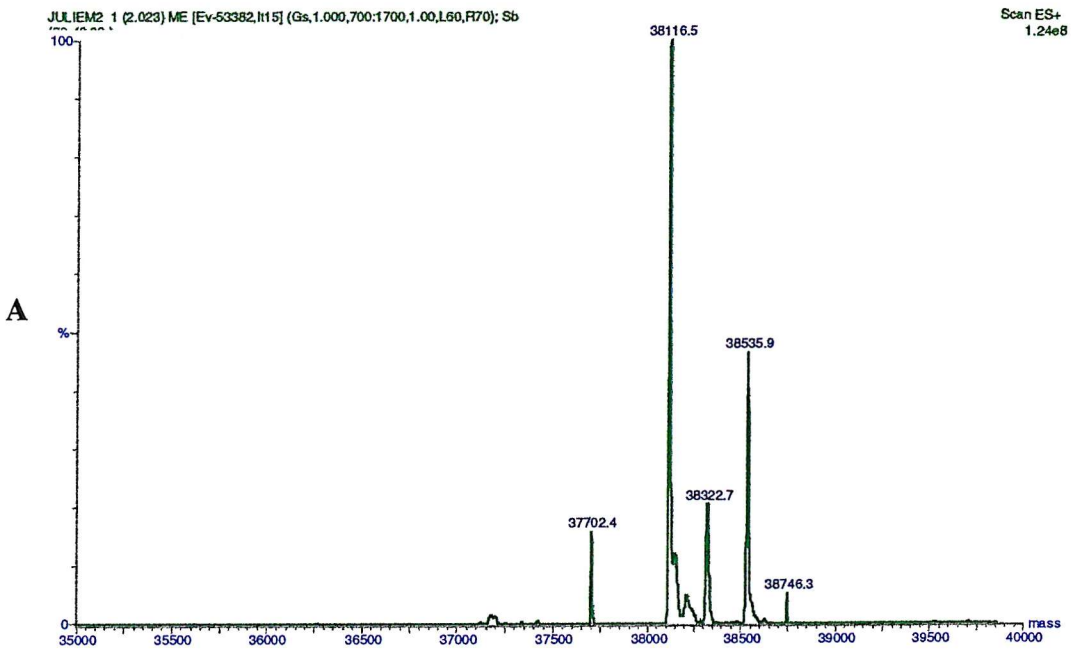
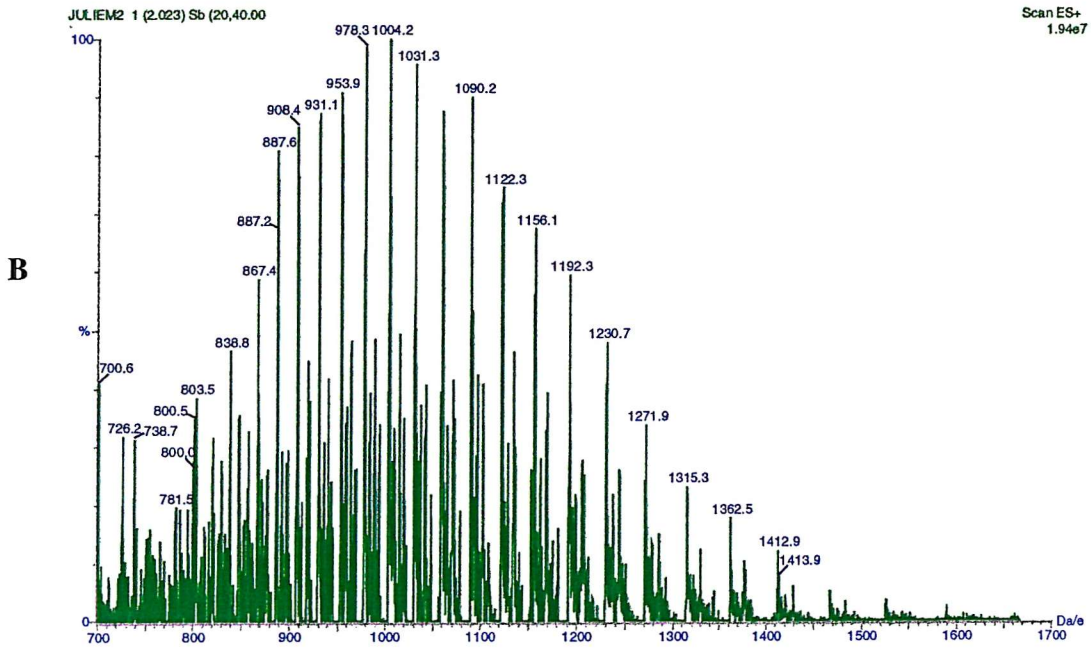


Figure 4.7 A) Maximum entropy deconvoluted mass spectrum of recombinant erythroid PBGD mixed with a five molar excess of its substrate PBG. Species at 38,116.5, 38,322.7, 38,535.9, 38746.3 correspond to enzyme E, ES, ES_2, and ES_3 respectively. **B) Corresponding raw data.**



4.2.3 Kinetic analysis of the recombinant ubiquitous and erythroid PBGDs to determine the K_m and V_{max} values

The deaminase assay for the determination of K_m and V_{max} was carried out as described in section 2.3.2. The data was analysed using an Eadie-Hofstee plot, comparison of the determined values with previously published results are shown in table 4.3. For both the recombinant ubiquitous and erythroid PBGDs the V_{max} values roughly correlate with the previously published results (Anderson and Desnick, 1980; Mazzetti and Tomio, 1993). However, the K_m value for the recombinant ubiquitous PBGD is 15 x greater than the previously reported value of 6μM. The K_m value of the erythroid PBGD differs from the published value by a factor of 1.3.

Enzyme	Source	K_m (μM)	V_{max} (nmole/hr)
Recombinant ubiquitous PBGD	This thesis	45	31
Recombinant erythroid PBGD	This thesis	8	2,250
Endogenous ubiqutious PBGD	Mazzetti and Tomio, 1993	3	33
Endogenous erythroid PBGD	Anderson and Desnick, 1980	6	2,300

Table 4.3 The kinetic parameters of the recombinant and endogenous human PBGDs. There is good correlation between the determined V_{max} values and the published values, but the K_m values vary significantly.

4.2.4 Reaction of the recombinant human ubiquitous and erythroid PBGD with Ehrlich's reagent to check the presence of the dipyrromethane cofactor

The *E. coli* holoenzyme gives a characteristic spectrum when reacted with Ehrlich's reagent (Jordan and Warren, 1987). The initial product has a λ_{max} of 565nm and is due to the reaction of Ehrlich's reagent with the free alpha-position of the dipyrromethane cofactor. The λ_{max} subsequently shifts to 495nm due to the formation of a conjugated system of double bonds.
The recombinant human ubiquitous and erythroid PBGDs were both investigated for the presence of the dipyrromethane cofactor, by reaction with Ehrlich's reagent (section 2.3.3). Spectra, over a range of 380-600nm, were recorded at 1minute and 10 minute intervals. Both isoenzymes gave the same positive result (figure 4.8) confirming the presence of the cofactor in agreement with observed enzyme activity (section 4.2.3) and determined molecular weights (section 4.2.2).

4.3 Crystallography of recombinant human porphobilinogen deaminase

Crystallisation conditions for the ubiquitous isoenzyme were initially determined using a vapour diffusion method known as the hanging drop technique (McPherson, 1976). Protein was screened with Hampton I and II matrix screens at a variety of protein concentrations using 2μL drops (1μL protein plus 1μL well solution). Initial hits were in Hampton II condition 15 containing 0.5M ammonium sulphate, 0.1M sodium citrate and 1.0M lithium sulphate. However, these crystals were multiple and therefore unsuitable for X-ray analysis. Several rounds of optimisation of this initial hit condition produced plate shaped crystals of the erythroid isoenzyme with dimensions 100 x 50 x 40 microns (figure 4.9).

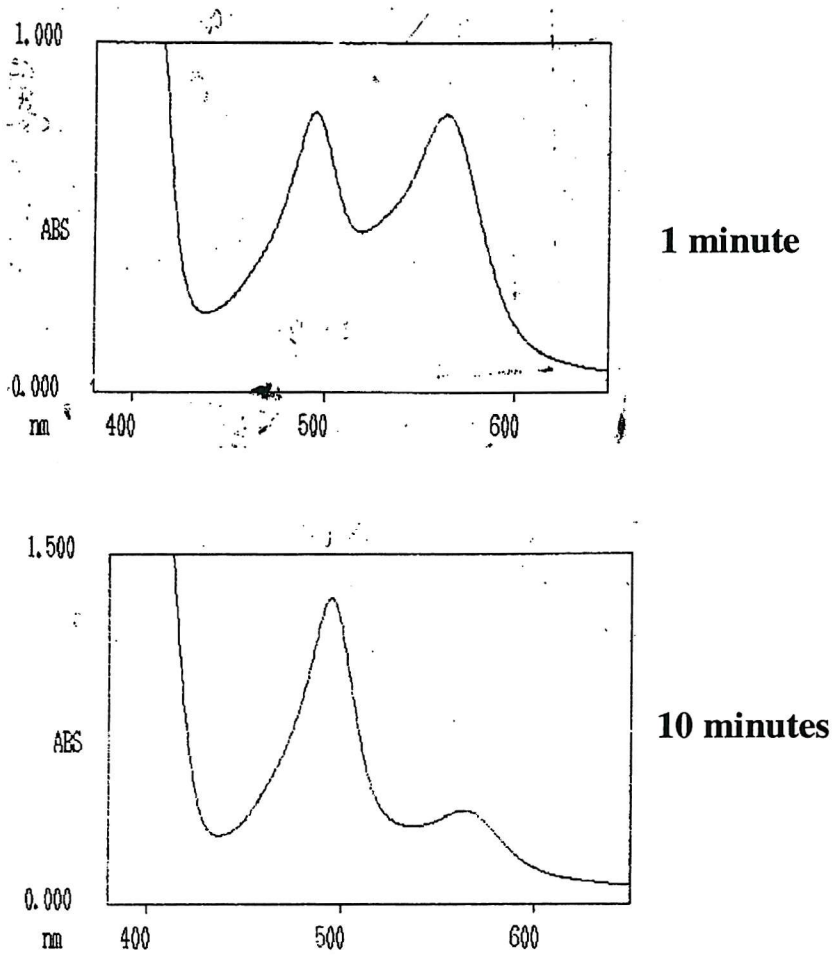


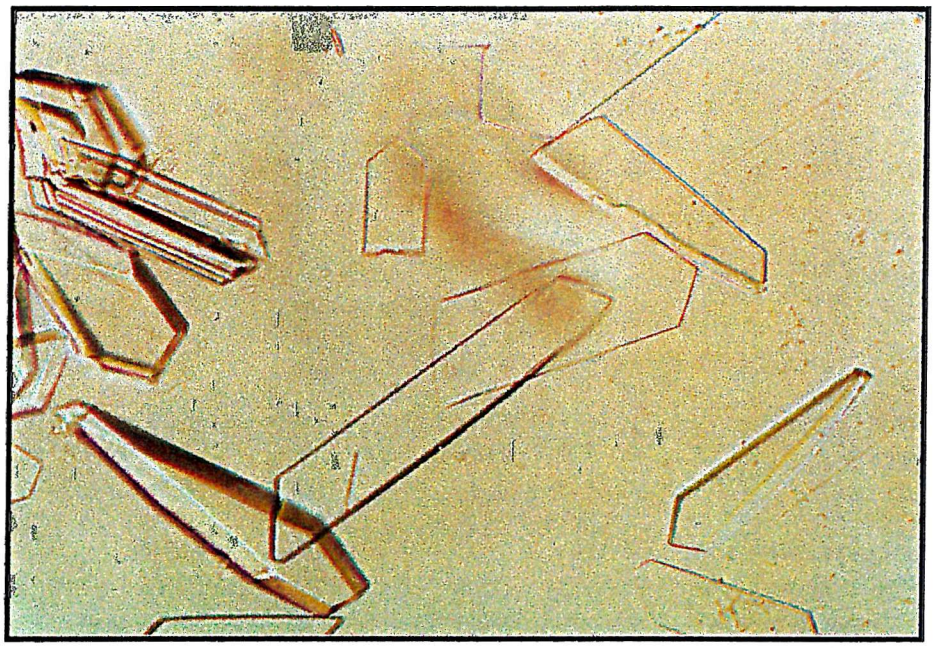
Figure 4.8 Reaction of recombinant human ubiquitous porphobilinogen deaminase with modified Ehrlich's reagent recorded.after 1 and 10 minutes. The initial product has a λ_{max} of 565nm and is due to the reaction of Ehrlich's reagent with the free alpha-position of the dipyrromethane cofactor. The λ_{max} subsequently shifts to 495nm due to the formation of a conjugated system of double bonds.

The optimal crystals were grown at a protein concentration of 5 mg/mL in 1.25M lithium sulphate, 0.6M ammonium sulphate, 50mM sodium citrate pH 5.6 and 5mM DTT, at room temperature. Each well was also filled with gaseous nitrogen prior to sealing with the coverslip to reduce possible enzyme oxidation (Abeer AlDbass) and the crystals were grown in the dark. The single crystals were harvested and immersed in a cryoprotectant, consisting of the mother liquor plus 15% w/v glycerol, before freezing in liquid ethane and liquid nitrogen.

A initial data set was collected on a single crystal of the erythroid isoenzyme at Hamburg but analysis of the data indicated that the crystal was twinned. A range of additives were screened in an attempt to overcome this problem. The presence of 0.5 % v/v dioxane and 15% v/v ethylene glycol enabled single crystals to be grown without twinning. The growth time was between three days and one week.

X-ray diffraction data was also collected on crystals of the ubiquitous isoenzyme, but the diffraction was very weak. Attempts to improve the diffraction of these crystals, by optimising their growth conditions, was unsuccessful.

A



B

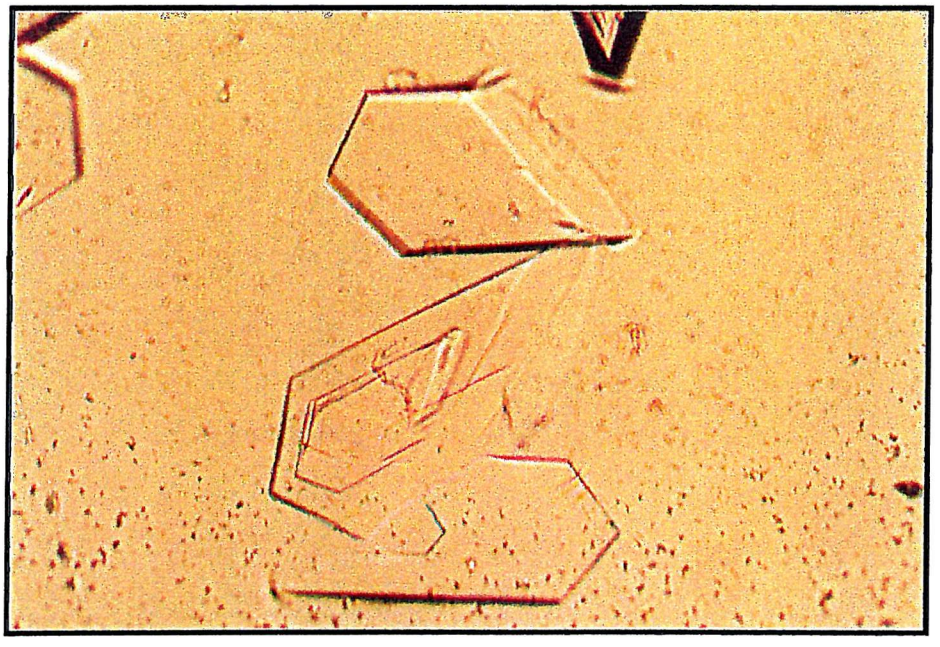


Figure 4.9 A) Crystals of erythroid PBGD. B) Crystals of ubiquitous PBGD (Julie Mosley/Abeer AlDbass).

4.3.1 Data Collection and processing

Synchotron diffraction data was collected on a crystal of the erythroid enzyme, by Dr. Thompson, at the Daresbury Synchotron Radiation Source using a Mar CCDdetector. Crystals were frozen directly in the cryostream in the native mother liquor. A total of 90° of data were collected with an oscillation angle of 1° and exposure time of 20 minutes per degree (figure 1.10). The crystals were anisotropic and diffracted to 2.7Å and 3.5 Å in perpendicular directions. Indexing the raw data with MOSFLM (Leslie, 1991) gave unit cell dimensions a=82.013Å, b=110.205Å, c=199.251Å, $\alpha=\beta=\gamma=90°$, indicating that the crystal belonged to the orthorhombic space group. After indexing, the data was scaled and merged (using the space group P222) before conversion into structure factor amplitudes using TRUNCATE (CCP4, 1994). HKLVIEW (CCP4, 1994) was used to determine systematic absences and indicated 2_1 screw axis along a and b (figure 4.11). Molecular replacement was completed using the recently determined human R167Q mutant ubiquitous PBGD structure (Mohammed, 2001). This indicated that the erythroid PBGD had crystallised in the space group $P2_12_12$ with 4 molecules in the asymmetric unit (figure 4.12). The human R167Q mutant and *E. coli* PBGDs also crystallised in this space group, but with one and two molecules in the asymmetric unit respectively. . Interestingly, the unit cell dimensions of the human erythroid and R167Q mutant PBGDs are almost identical, except c = 199.251 Å for the erythroid crystals, c = 109.6 Å for the R167Q mutant crystals.

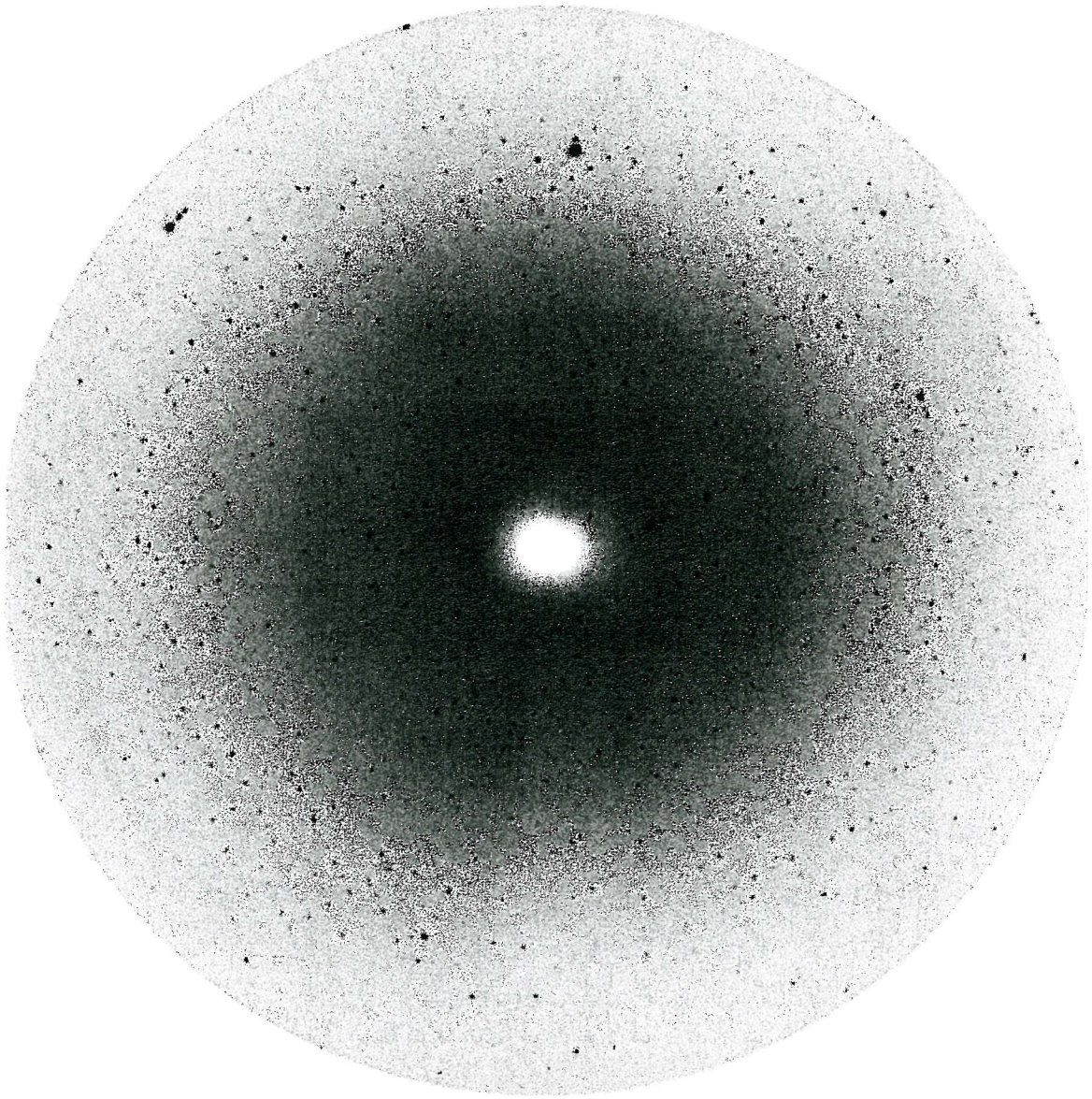


Figure 4.10 A 1° oscillation image from diffraction of a single crystal of recombinant human erythroid PBGD (Dr. Darren Thompson).

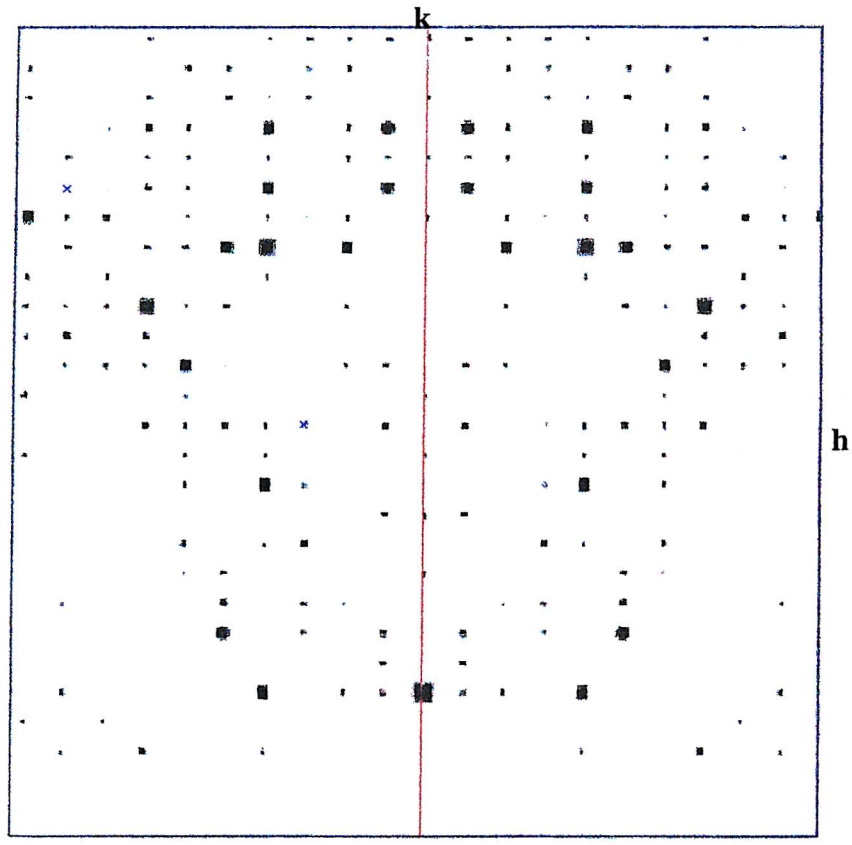


Figure 4.11 The systematic absences along zones 0K0 and H00 indicating a 2_1 screw axis along a and b. Data from 00L zone was completely missing. Picture generated using HKLVIEW (CCP4, 1994) (Dr. Thompson).

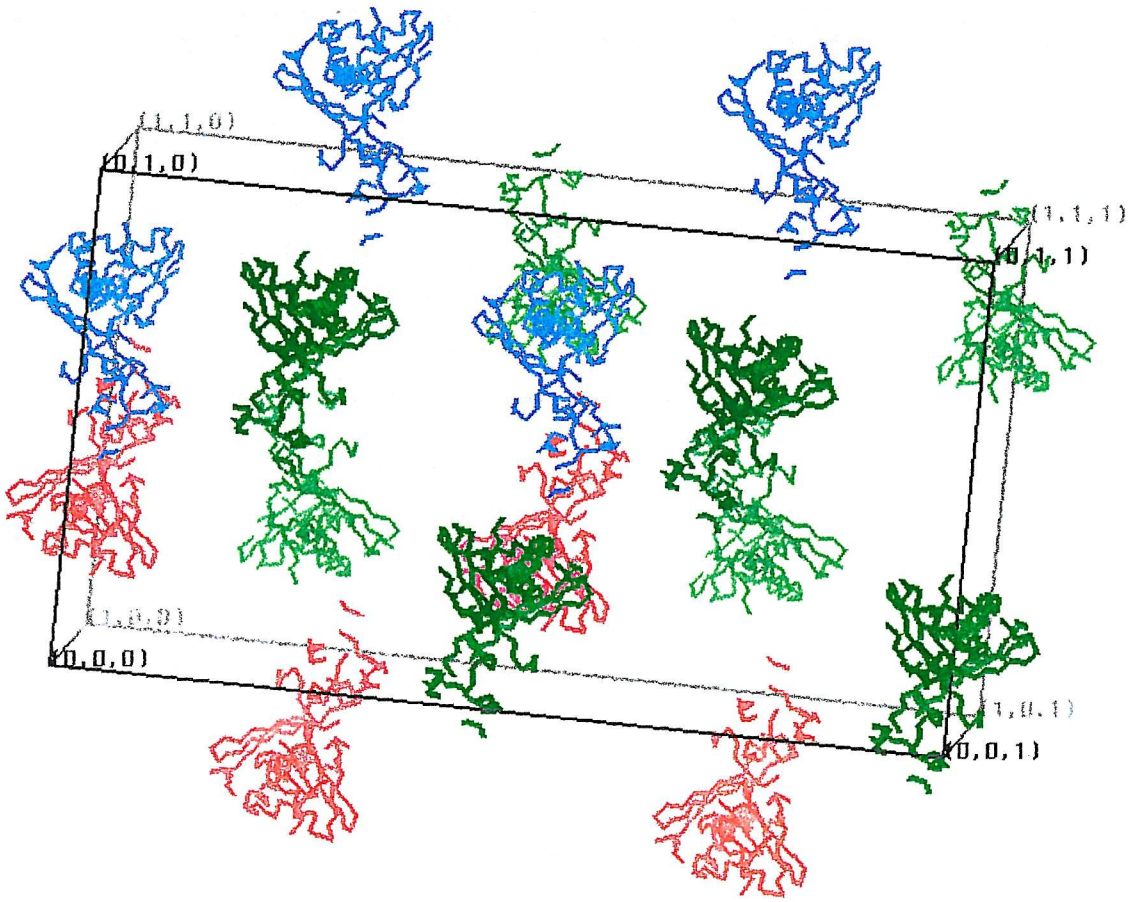


Figure 4.12 The crystal packing of recombinant human erythroid PBGD (Dr. Thompson). The erythroid crystal crystallises in the orthorhombic spacegroup $P2_12_12$ (a=82.013Å, b=110.205Å, c=199.251Å, $\alpha=\beta=\gamma=90°$) with 4 molecules in the asymmetric unit (MOLPACK; Wang *et al.*, 1991).

4.3.2 The structure of recombinant human erythroid porphobilinogen deaminase and comparison to the R167Q mutant and *E. coli* structures

The structure of recombinant human erythroid porphobilinogen deaminase was determined to 3Å using the homologous human ubiquitous PBGD arginine 167 glycine mutant as a search model for molecular replacement. The human R167Q mutant ubiquitous enzyme was purified and crystallised subsequent to the recombinant erythroid and ubiquitous isoenzymes using the methods described in sections 3.3.3 and 4.3. The overall resolution of the mutant structure is 2.8Å (Mohammed, 2001) in comparison to 3.0Å for the erythroid structure and the mutant dataset is more complete.

The recombinant human erythroid PBGD consists of 344 amino acids, only 269 of which are visible in the X-ray structure due to a number of disordered residues. The residues present, along with their associated secondary structural elements are listed in table 4.4.

Unsurprisingly, human erythroid PBGD has an almost identical topology to the mutant and a very similar topology to *E. coli* PBGD. The polypeptide chain is folded into three domains, with dimensions 57 x 47 x 55 Å, identical to the mutant enzyme (figure 4.13). Domains 1 and 2 have a similar topology of five β-sheets with three intervening α-helices packed against each face. Domain three consists of three antiparallel β-sheets with three α-helices positioned on one face. Domains 1 and 2 are linked by two hinge regions (residues 98 – 102 and 196 – 201) and domain 3 is linked to domain 1 by a small fragment (residues 220 – 223). The linking of the three domains, by flexible hinge regions, is also found in the mutant and *E. coli* structures. The N-terminal region (residues 1-8) is absent in the erythroid structure and no electron density was observed for residues 36 – 71, part of which forms a mobile loop (residues 46 - 55) that projects from domain 1 over the active site cleft. Domain one of the R167Q mutant ubiquitous structure is also incomplete in these N-terminal and loop regions, with similar residues 1-18 and 57-78 absent. These

findings are consistent with the observation of a missing loop region (residues 48-58) in domain one of the *E.* coli enzyme. It has been postulated that this region may act as a flexible "lid" to the active site cleft, able to occupy both an open and closed conformation (Louie *et al.*, 1992).

There is no observable density for the four C-terminal amino acids (residues 341 – 344), and three C-terminal amino acids (residues 359-361) in the erythroid and R167Q mutant structures respectively.

The main difference between the two human enzymes and the *E. coli* enzyme is that in the human enzymes there is an additional insert of 29 residues in domain 3, following strand $\beta 3_3$. This is difference is highlighted for the erythroid PBGD structure, by superposition with the *E. coli* structure (figure 4.14). This insert forms a surface-loop, positioned at 281-310 in the erythroid structure and 298-327 in the mutant structure. In both the human structures the insert is almost complete, except for the six residues 290-295 and 307-312 in the erythroid and mutant structures respectively.

Residues	Secondary structure	Residues	Secondary structure	Residues	Secondary structure
1 - 8	disordered	128 – 130	loop	225 – 238	helix ($\alpha 1_3$)
9 - 11	loop	131 – 136	helix ($\alpha 1_2$)	239 – 247	loop
12 - 28	helix ($\alpha 1_1$)	137 – 143	disordered	248 - 255	strand ($\beta 1_3$)
29 - 33	loop	144 – 147	strand ($\beta 2_2$)	256 – 257	loop
34 - 35	strand ($\beta 2_1$)	148 - 153	loop	258 - 266	strand ($\beta 2_3$)
36 – 71	disordered	154 – 161	helix ($\alpha 2_2$)	267 – 272	loop
72 – 74	loop	162 – 167	loop	273 – 280	strand ($\beta 3_3$)
75 – 79	strand ($\beta 3_1$)	168 - 172	strand ($\beta 3_2$)	281 – 289	loop
80 – 82	loop	173 – 178	helix ($\alpha 3_2$)	290 - 295	disordered
83 - 86	disordered	179 – 186	loop	296 – 307	loop
87 – 90	loop	187 – 189	strand ($\beta 4_2$)	308 – 326	helix ($\alpha 2_3$)
91 – 97	strand ($\beta 4_1$)	190 – 201	loop	327 - 329	loop
98 – 102	loop	202 – 208	strand ($\beta 5_1$)	330 – 339	helix ($\alpha 3_3$)
103 - 112	disordered	209 – 211	loop	341 – 344	disordered
113 - 123	loop	212 – 219	helix ($\alpha 3_1$)		
124 – 127	Strand ($\beta 1_2$)	220 – 224	loop		

Table 4.4 Secondary structural elements of recombinant human erythroid PBGD from the N- to the C-terminus. Erythroid PBGD consists of 344 amino acids, only 269 of which are visible in the X-ray structure due to 75 disordered residues.

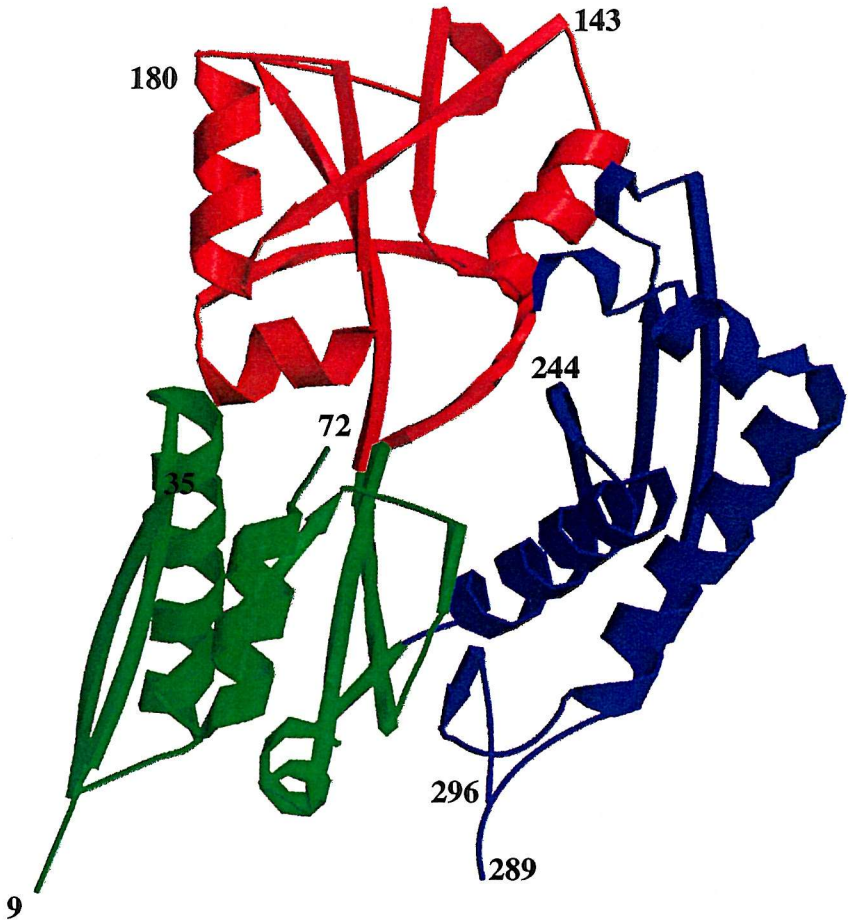


Figure 4 13 The three-dimensional X-ray structure of recombinant human erythroid porphobilinogen deaminase (Dr. Thompson). The structure consists of three domains, domain 1 is green, domain 2 is red and domain 3 is blue. Domains 1 and 2 have a similar topology and flexible hinge regions connect the three domains. The insertion loop in domain three (residues 281 – 289 and 296 – 310) is the main difference between the erythroid and *E. coli* enzymes.

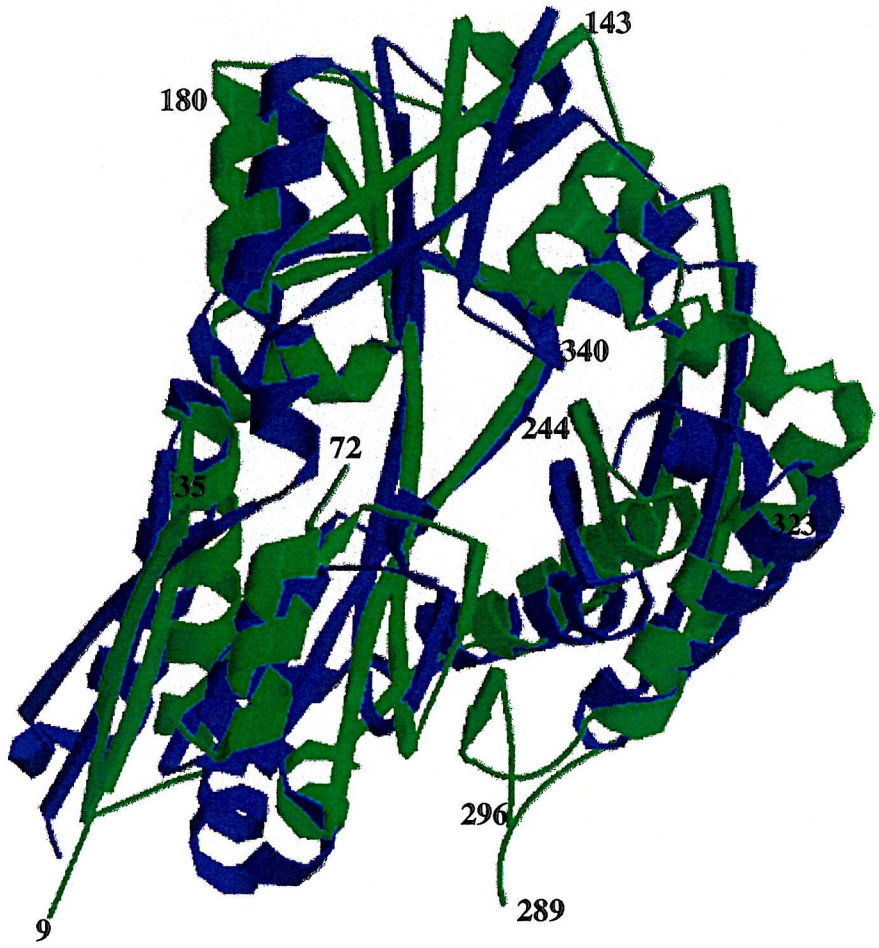


Figure 4.14 The superposition of the human erythroid PBGD structure (green) and the *E. coli* PBGD structure (blue) (Dr. Thompson). The two enzymes have a very similar overall topology, which is expected for two enzymes with 60% homology and 49% sequence identity. The main difference, the insertion of 29 aminoacids (residues 281-310) in domain 3, is emphasised by comparison of the two structures. However, this extra loop region is incomplete, as residues 290-295 have no corresponding electron density.

4.3.2.1 Detail of the active site

The active site of human erythroid PBGD is very similar to the human R167Q mutant PBGD and *E. coli* PBGD enzymes, which is unsurprising considering the large number of invariant residues. The active site cleft, formed at the interface between domains 1 and 2, occupies nearly half the width and depth of the molecule. The "floor" of the cleft is bordered by the carboxyl termini of the strands $\beta 1_1$ and $\beta 3_1$. The amino termini of the helices $\alpha 1_2$ and $\alpha 2_2$ and their connecting loops border the "ceiling" of the cleft.

In the *E. coli* and human R167Q mutant structures the dipyrromethane cofactor is linked to the invariant residue, cysteine 242 and cysteine-261, respectively. This invariant cysteine is located on a loop that projects from domain 3 into the active site cleft. In erythroid PBGD the corresponding residue cysteine 244 is visible, but there is no density for ring C1 of the cofactor and the thiol-linkage with cysteine 244 (figure 4.15).

Ring C2 of the cofactor suggests the cofactor is in the reduced conformation, positioned towards the rear of the active site cleft. The carboxylate groups of the C2 ring form salt links with arginine 178 (195 in the mutant) and serine 79 (96 in the mutant). Lysine 81 (98 in the mutant) also interacts with the carboxylates from the C2 ring (figure 4.16). Mutation of an equivalent residue in *E. coli* (lysine 83) leads to a highly unstable protein that is rapidly degraded. No human mutations have been reported at this position. It is thought that serine 129 (146 in the mutant) interacts with the carboxylates of ring C1 of the cofactor, but this can not be confirmed due to the absence of the C1 ring. The proposed catalytic residue, aspartate 82 (99 in the mutant) interacts with the NH groups of both the C1 and C2 pyrrole rings. Mutation of this residue in humans has been identified as an AIP mutation in Sweden (unpublished).

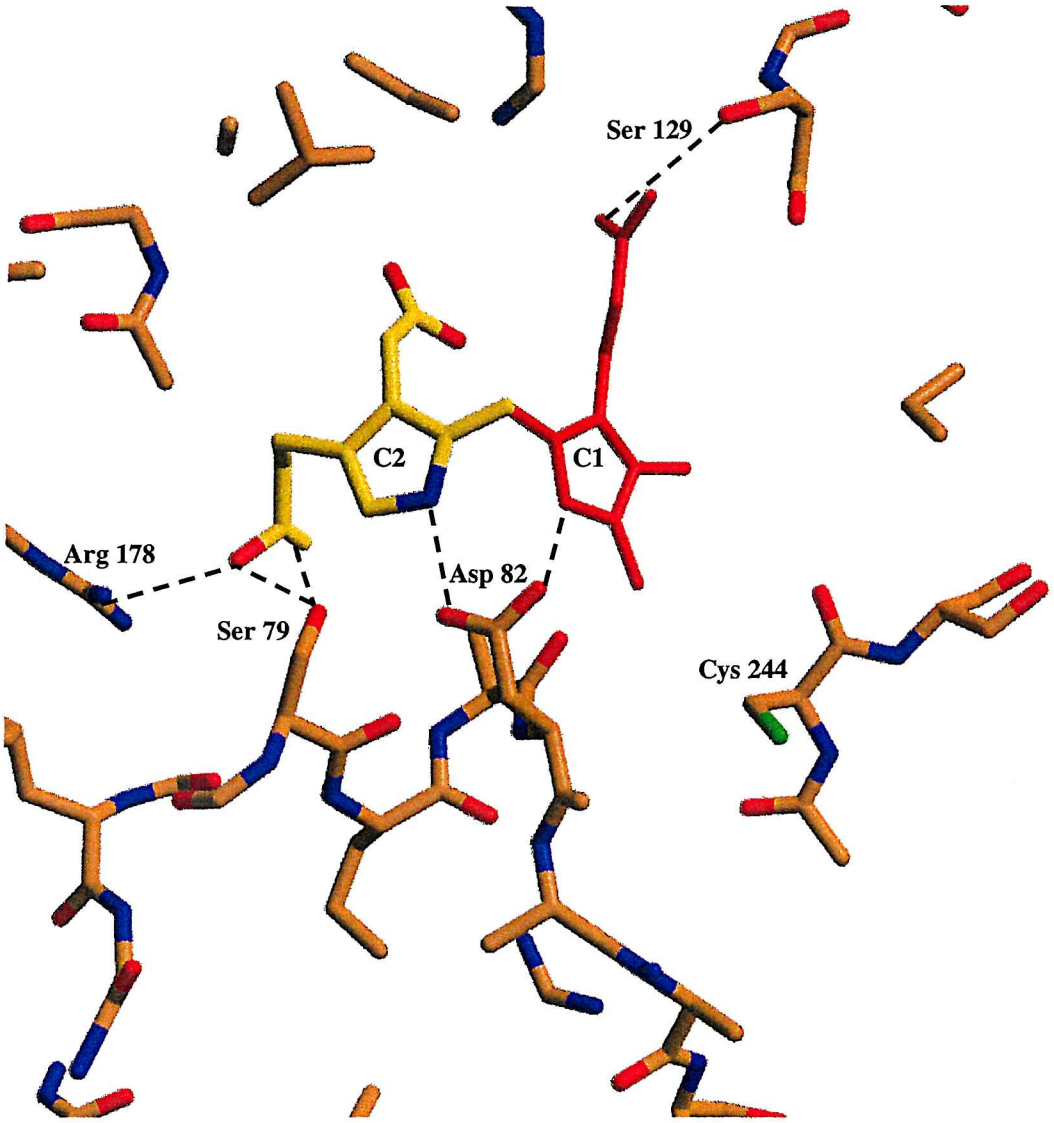


Figure 4.15 Active site of human erythroid PBGD showing the catalytic residue aspartate 82 (Dr. Thompson). There was clear electron density for cysteine 244 and ring C2 of the dipyrrromethane cofactor, however ring C1 of the cofactor (modelled in red) and the thiol linkage to cysteine 244 are missing. The interactions between rings C1 and C2 of the cofactor and the side chains of active site residues lining the cleft are shown in black. The proposed catalytic residue, aspartate 82 (99 in the mutant) interacts with the NH groups of both the C1 and C2 pyrrole rings.

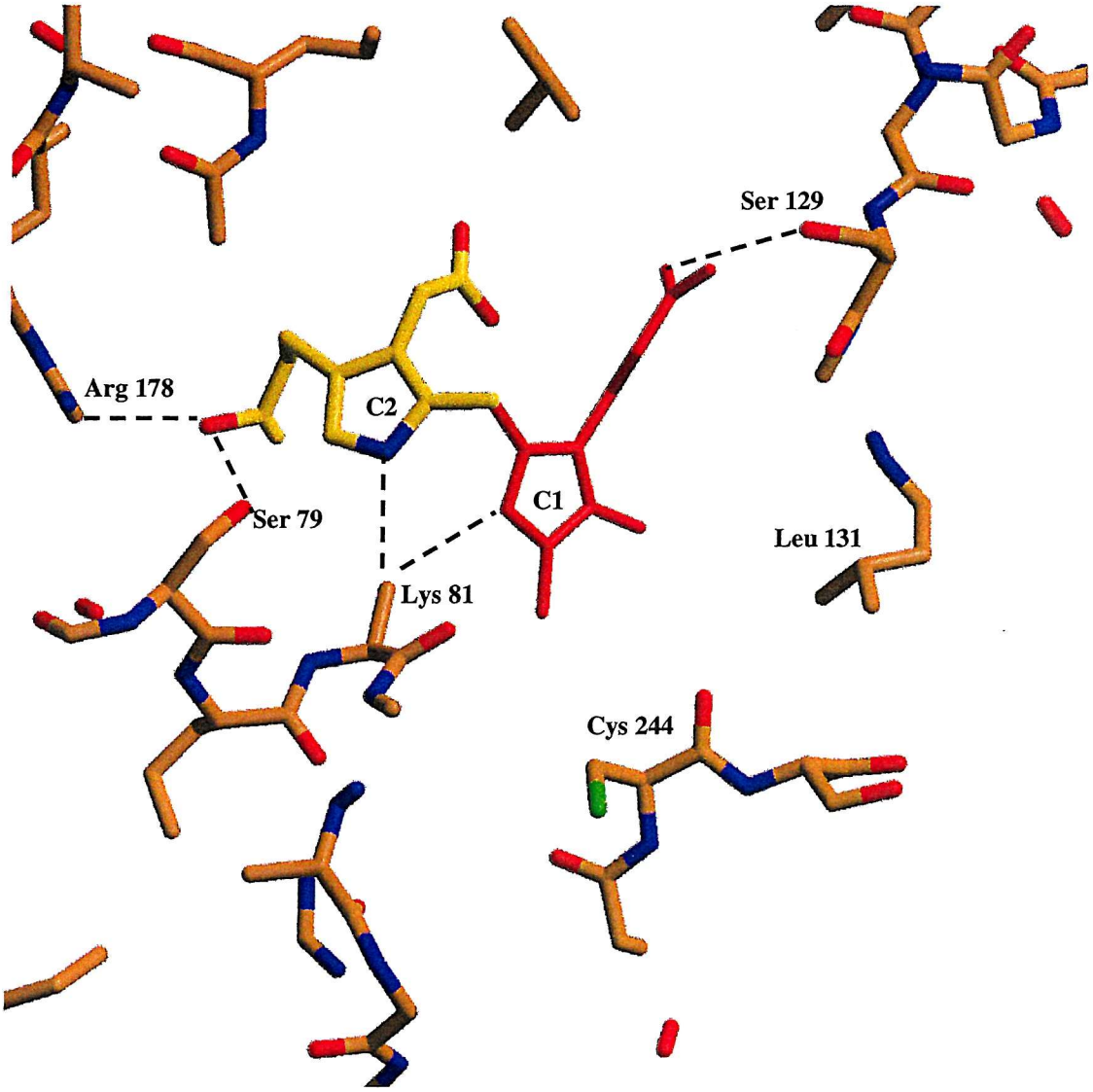


Figure 4.16 Active site of human erythroid PBGD showing the important invariant residue lysine 81 (Dr. Thompson). Lysine 81 form key interactions with the carboxylates from ring C2 of the cofactor. Mutation of an equivalent residue in *E. coli*, lysine 83, results in a highly unstable protein that is rapidly degraded.

4.4 Conclusions

This first section of this chapter has described the characterisation of the previously purified recombinant ubiquitous and erythroid PBGSs. Experiments to determine the N-terminal sequence and an accurate molecular mass for both isoenzymes agreed with expected results (Raich *et al.*, 1986; Grandchamp *et al.*, 1987) except the N-terminal methionine was clipped from the recombinant ubiquitous enzyme, giving a reduced mass of 39,620 daltons. The identification of the enzyme intermediates by mass spectrometry, combined with evidence of cofactor presence, confirms the isoenzymes are active and that they react in a step-wise fashion forming the intermediates ES, ES_2 and ES_3 (Desnick *et al.*, 1985). The kinetic parameters of the recombinant ubiquitous and erythroid isoenzymes are largely similar to those determined for the endogenous enzymes (Anderson and Desnick, 1980; Mazzetti and Tomio, 1993). The only significant difference was a 15 fold reduction in the substrate affinity of the recombinant ubiquitous enzyme, indicated by an increased K_m value.

The later part of this chapter describes the crystallisation and structure determination of the human PBGDs. The three-dimensional structure of human erythroid PBGD is the first native human PBGD to be determined; previously the human R167Q mutant was resolved to 2.8Å (Mohammed, 2001). The human R167Q mutant protein was purified and crystallised subsequent to the erythroid enzyme, using the same methods developed for the purification and crystallisation of the human ubiquitous and erythroid PBGDs. The erythroid PBGD has a similar overall topology to the *E. coli* enzyme (Louie *et al.*, 1996) the main difference being the large insert in domain three (Dr. Thompson). A previous model of the human PBGD based on the *E.coli* enzyme (figure 4.1) predicted this insertion loop was positioned between residues serine 269 and glutamine 270 in the *E.coli* enzyme, following strand $\beta 2_3$ (Wood *et al* ,1995). However, the actual position of this additional loop relative to the *E.coli* enzyme is between residues alanine 279 and proline 280, following strand $\beta 3_3$. It is proposed that this loop region has a structural role, stabilising the interface between domains 1 and 3 and contributing to the hydrophobic core of domain 3.

The active site of human erythroid PBGD is very similar in architecture and location to the human R167Q mutant and *E. coli* PBGDs. Normally, ring C1 of the dipyrromethane cofactor is covalently linked to an invariant cysteine residue. In the erythroid structure ring C1 and the thiol linkage are absent, so no conclusions can be made about its position and interactions in the active site cleft. The erythroid structure confirms the importance of the invariant active site residue, lysine 81, which is shown to form salt-bridge interactions with ring C2 of the cofactor. Mutations of the corresponding residue in *E. coli* PBGD (lysine 83) disrupts folding of the enzyme, implicating a structural role, involved in orientating the cofactor and hence maintaining structural stability. In the erythroid structure, as predicted, the catalytic residue asparatate 82 forms hydrogen bonds with both pyrrole NH atoms. Disruption of these hydrogen bonds by mutation of aspartate 82 provides a structural explanation for a recently identified Swedish AIP associated mutation (unpublished). Due to the absence of several invariant residues in the active site of the erythroid structure it is not possible to rationalise the structural consequences of other AIP associated mutations occurring in the active site. Several AIP mutations positioned away from the active site can be explained by the analysis of the erythroid structure. One such mutation is L117R which shows the introduction of a charged residue into a hydrophobic core which would disrupt protein folding. The R116W (Solis *et al.*, 1999) and R201(De Siervi *et al.*, 1999) mutations (equivalent to R99 and R184 in the erythroid enzyme) disrupt the salt-briging interactions with glutamate 233 and aspartate 161 that contribute to structural stability. Despite this information the structural consequences of some AIP associated mutations is still poorly understood. Future work, towards increased understanding of AIP, requires the determination of the complete erythroid and ubiquitous PBGD structures. Additionally, the R167Q mutant could be used to study the formation of the enzyme: intermediates, providing essential information about the interacting residues and structural changes that occur during the catalytic cycle.

Chapter 5

Further evidence for flexibility of *E. coli* porphobilinogen deaminase from site-directed site directed mutagenesis of the *E. coli* enzyme at the active site loop and at an interdomain region by mutagenesis of glycine 57 and glycine 264

5.1 Introduction

Porphobilinogen deaminase catalyses a novel polymerisation reaction that involves the addition of four porphobilinogen molecules in a stepwise fashion through enzyme-intermediate complexes, ES, ES_2, ES_3 and ES_4. The X-ray structure of the enzyme indicates that a single substrate-binding site exists that is used repetitively for each addition of the substrate (Louie *et al.*, 1992; 1996). Substantial conformational changes in the vicinity of the active site cleft are thought to occur as the protein adjusts its structure to accommodate the growing polypyrrole chain.

The structure of *E. coli* porphobilinogen deaminase provides some of the most compelling evidence for the ability of the enzyme to undergo conformational changes during the catalytic cycle. Firstly, the enzyme consists of three distinct domains with polar interfaces that are joined by potentially flexible hinge regions. Particularly notable is the double crossover between domains 1 and 2. This is formed by two strands, one of which crosses from domain 1 into domain 2 (residues 100-104) and a second (residues 194-199) that crosses back from domain 2 into domain 1. Secondly there is a strand (residues 218-221) linking domain 2 to the long $\alpha 1_3$ helix of domain 3 (residues 222-237). This helix is composed of a high proportion of polar amino acids suggesting that it is solvent accessible and may be able to adopt various positions between domain 2 and domain 3. Thirdly, the amino acids that make up the dipyrromethane cofactor loop (residues 238-245), including the invariant Cys 242 forming the covalent thioether link with the cofactor, adopt a type I β-bend with substantial

conformational flexibility. Of the amino acids found in this loop (LEGGCQVP) the sequence LEGG is found at the hinge region of immunoglobulins at the junction between the fc and fab domains. Thus Cys 242 is immediately preceded by two sequential glycine residues that occupy relatively little space and hence could allow considerable flexibility in this region.
Two different mechanisms have been proposed to explain the repositioning of the catalytic site to allow the addition of the four substrate molecules during the catalytic cycle (Louie *et al.*, 1996). In one mechanism, the *moving chain* mechanism, the elongating polypyrrole chain could be "pulled through" the catalytic site attached to domain 3 with the major conformational changes occurring at the domain 3 boundaries. An alternative mechanism, the *sliding active site* mechanism, could involve the repositioning of the catalytic machinery to accommodate the growing polypyrrole chain, with the major conformational changes occurring in the vicinity of the active site cleft between domains 1 and 2.

Biochemical evidence in support of the *moving chain* mechanism comes from the observed increase in the inactivation of the *E. coli* enzyme by the thiophilic reagent *N*-ethylmaleimide as a result of the modification of Cys 134 (Warren *et al.*, 1995). Cysteine 134 is buried at the interface between domains 2 and 3 and is largely insensitive to modification in the native enzyme, however, this residue becomes progressively more reactive to *N*-ethylmaleimide as porphobilinogen molecules are added stepwise to form ES, ES_2 and ES_3. This suggests that the distance between domains 2 and 3 increases progressively as the polypyrrole chain is extended, increasing the exposure of cysteine 134 to the modifying reagent.

The alternative *sliding active site* mechanism is supported by the observation that the key substrate binding groups Arg 11, Phe 62 and the catalytic group Asp 84 are all located on domain 1, whereas the cofactor binding residues, Arg 131, 132, 149 and 155 are located on domain 2. Furthermore, large conformational changes are known to occur between domains 1 and 2 on substrate binding in the transferrins and other members of the family of binding proteins (Louie, 1993). Domains 1 and 2 of porphobilinogen deaminase are structurally related to the proteins in this family.

The X-ray structure of porphobilinogen deaminase also implies that there is another flexible region of the active site that may play a key role during the catalytic cycle. This is the active site lid region, a highly conserved, though not invariant, sequence of amino acids from residues 48-58 (figure 5.1). This lid may have several functions: firstly it may form additional binding interactions with the substrate, possibly through the lysine side chains; secondly, it may dock into the active site with substrate to prevent interaction with the external environment during catalysis; thirdly, it may play a role in conformational changes as each substrate binds successively to the enzyme. The importance of the lid is suggested from experiments in which lysine 55 and lysine 59 have been modified by reaction with pyridoxal 5'-phosphate (Hart *et al.*, 1984). Pyridoxylation is substantially prevented by the presence of the substrate pophobilinogen suggesting that the lysines may become accessible during catalysis. Mutagenesis of lysine 55 and lysine 59 to glutamine results in substantial lowering of the specificity of the enzyme (Hadener *et al.*, 1990). However the failure of the mutations to inactivate the enzyme totally suggests a binding role, rather than a catalytic role, is most likely for the lid region.

This chapter describes attempts to gain further evidence in support of conformational changes during the catalytic reaction by the construction of two mutant porphobilinogen deaminases; one to form a conformationally restricting tether between domains 2 and 3 and the other to substitute an invariant glycine in the active site lid.

The former involves mutagenesis of Gly 264 to Cys in domain 3. The X-ray structure of *E. coli* deaminase suggests that if Gly 264 is mutated to Cys, then the distance between the new Cys 264 in domain 3 and the existing Cys 134 in domain 2 will be ideally suited to form a stable S-S bond between domains 2 and 3 (figure 5.2). Since Gly 264 is a variable residue, with alanine, valine and tryptophan known at the equivalent position in other deaminases (figure 5.3), it was not considered likely that the Gly 264 Cys mutation would itself be unduly damaging to the enzyme. The properties of an enzyme in which these two cysteines form a disulphide bond linking rigidly domains 2 and 3 could shed

further light on the importance of conformational changes during the catalytic reaction.

The second mutation involves mutagenesis of the invariant Gly 57 to Ala. Little is known about the structure of this region of the deaminase because no clear electron density for residues 48-58 is available. This relatively conservative change from Gly to Ala would be expected to have a minimal effect on the enzyme structure unless the small size of Gly 57 is essential for conformational flexibility in the active site lid region.

This chapter describes preliminary findings from the mutagenesis experiments that have been carried out using the deaminase enzyme from *E. coli*.

Organism	Residue												
	45												58
Escherichia coli	G	D	V	I	L	D	T	P	L	A	K	V	G	G
Rat	G	D	K	I	L	D	T	A	L	S	K	I	G	E
Mouse	G	D	K	I	V	D	T	A	L	S	K	I	G	E
Human	G	D	K	I	L	D	T	A	L	S	K	I	G	E
Arabidopsis thaliana	G	D	K	I	L	S	Q	P	L	A	D	I	G	G
Pea (Pisum sativum)	G	D	K	I	L	S	Q	P	L	A	D	I	G	G
Pseudomonas aeruginosa	G	D	K	L	L	D	A	P	L	A	K	I	G	G
Euglena gracilis	G	D	K	I	L	D	S	P	L	A	K	V	G	G
Bacillus subtilis	G	D	R	I	V	D	V	T	L	S	K	V	G	G
Chlorobium vibrioformes	G	D	V	L	L	D	S	P	L	S	K	I	G	D
Yeast	G	D	Q	I	Q	F	K	P	L	Y	S	F	G	G
Clostridium josui	G	D	K	I	L	D	K	T	L	D	K	I	E	G
Yersina intermedia	G	D	I	I	L	D	T	P	L	A	K	V	G	G
Erwinia chrysanthemi	G	X	L	L	L	D	T	P	L	A	K	V	G	G
Mycobacteium leprae	G	D	Q	•	•	S	S	A	S	I	D	S	L	G

Figure 5.1 The conservation of amino acids within the surface loop comprising residues 45-57. All the residues showing at least 70% conservation are coloured in blue.

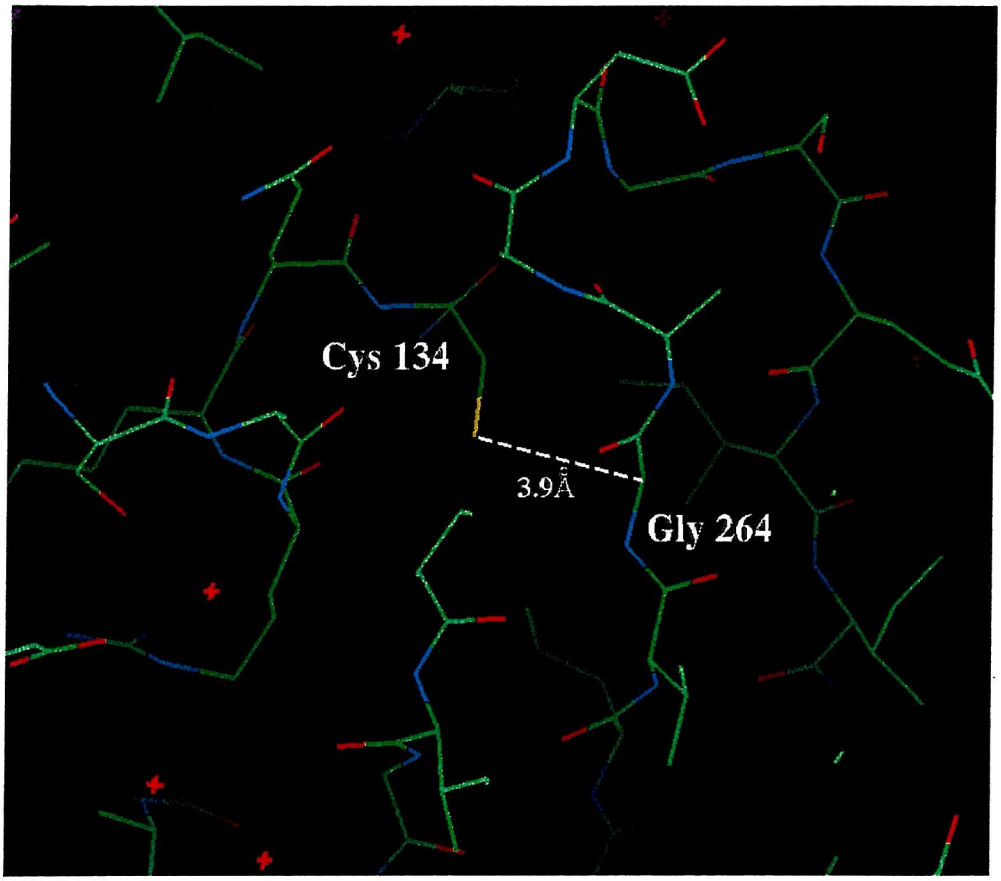


Figure 5.2 The position of the two residues cysteine-134 and glycine-264 in *E. coli* PBGD. The distance between the sulphur group of cysteine 134 and the alpha carbon of glycine-264 is 3.9Å. It is possible that a cysteine residue, introduced by mutation of glycine 264, could be accommodated without disruption of this hydrophobic area. Additionally, cysteine 264 could from a disulphide bond with cysteine 134.

Organism	Residue												
	257												269
Escherichia coli	D	G	E	I	W	L	R	G	L	V	G	A	P
Rat	D	G	Q	L	Y	L	T	G	G	V	W	S	L
Mouse	D	G	Q	L	Y	L	T	G	G	V	W	S	L
Human	D	G	Q	L	Y	L	T	G	G	V	W	S	L
Arabidopsis thaliana	E	G	N	C	I	F	R	G	L	V	A	S	P
Pea (Pisum sativum)	D	G	N	C	L	F	R	G	L	V	A	S	P
Pseudomonas aeruginosa	G	D	Q	L	W	L	R	G	L	V	G	Q	P
Euglena gracilis	D	G	Q	L	R	M	E	A	R	V	G	S	V
Bacillus subtilis	G	Q	D	E	I	M	T	G	L	V	A	S	P
Chlorobium vibrioformes	D	G	T	L	K	L	L	A	F	V	G	S	V
Yeast	T	K	K	L	L	L	K	A	I	V	V	D	V
Clostridium josui	S	E	I	I	L	K	G	L	Y	C	N	E	T
Yersina intermedia	.	.	.	.	.	.	.	.	.	.	.	.	.
Erwinia chrysanthemi	.	.	.	.	.	.	.	.	.	.	.	.	.
Mycobacteium leprae	I	D	G	E	G	R	V	F	E	E	L	S	L

Figure 5.3 The conservation of amino acids in the region of Gly 264 in domain 3. Only valine-266 in this region (coloured in blue) shows at least 70% conservation between species.

5.2 The purification and analysis of the *E. coli* mutants G57A and G264C

The following sections describe the generation, purification and analysis of the glycine mutants.

5.2.1 The generation of the glycine mutants G57A and G264C

Oligonucleotide mutagenic primers were designed to produce the mutation through a single base mismatch at the designated codon. Table 5.1 indicates the oligonucleotide sequence of the mutagenic primers, with the mismatch highlighted.

PRIMER	OLIGONUCLEOTIDE SEQUENCE
WILD-TYPE	GCG AAA GTA GGC GGA AAA G
G57A	GCG AAA GTA GCC GGA AAA G
WILD-TYPE	GGG CTG GTC GGC GCG CCG GAC
G264C	GCG CTG GTC TGC GCG CCG GAC

Table 5.1 Oligonucleotides used for the construction of G57A and G264C porphobilinogen deaminase mutants. The mismatches are highlighted in blue

The mutations of the *E. coli* porphobilinogen deaminase *hemC* gene were constructed by the method of Nakamaye and Eckstein (1986) using the Sculptor™ *in vitro* mutagenesis kit (section 2.2.15.2). After identification of positive clones and checking the sequence fidelity by DNA sequencing (figure 5.4a and b), the mutant *hemC* genes were ligated into the *Bam*HI and *Sal*I restriction sites of pUC18 and the recombinant plasmids were transformed into *E.coli* strain TB1.

5.2.2 Purification of the G57A mutant *E. coli* porphobilinogen deaminase

The expression of the G57A mutant is shown in figure 5.5 and the SDS gel of the purified protein in figure 5.6. The G57A mutant was purified by the same method used for the wild-type enzyme (see section 2.3.13).

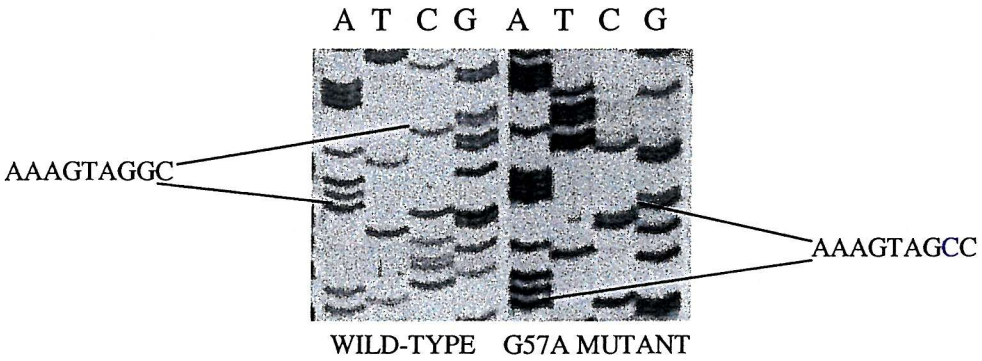


Figure 5.4a Sequencing gels showing the single point mutation in the G57A mutant DNA.
The mutated base is shown in blue and the wild-type sequence is shown for comparison.

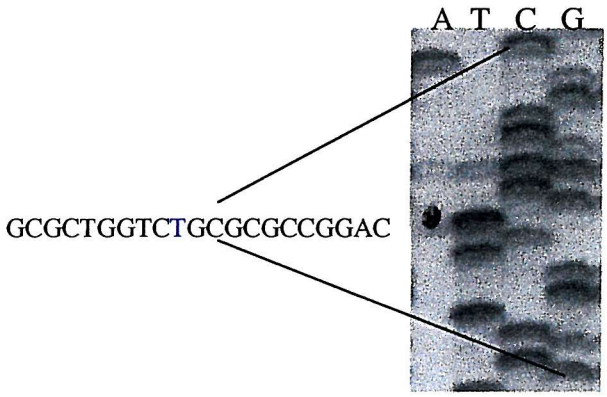


Figure 5.4b Sequencing gel showing the single point mutation in the G264C mutant DNA.
The mutated base is shown in blue.

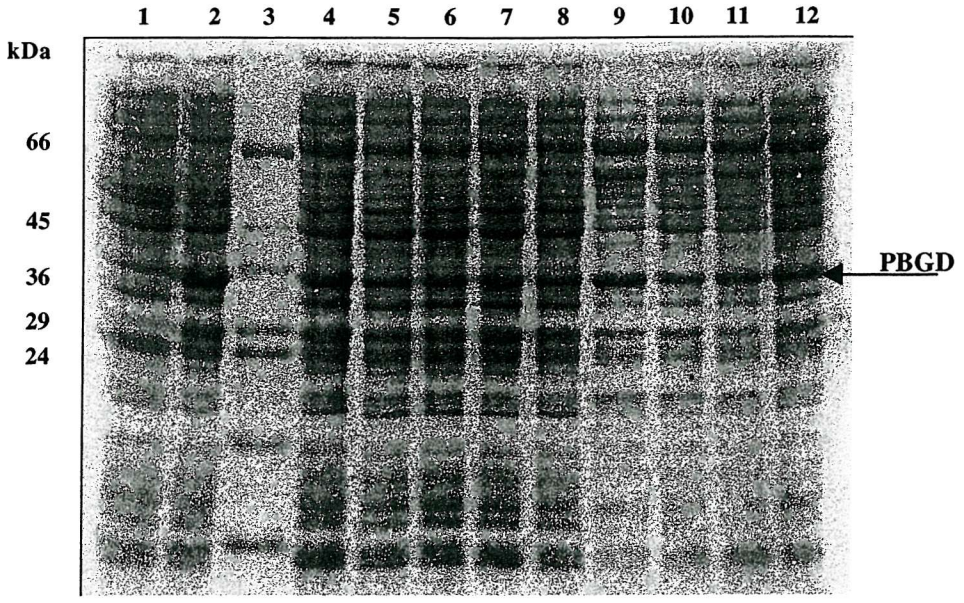


Figure 5.5 SDS PAGE showing the expression of the mutant enzyme G57A in comparison with the wild-type porphobilinogen deaminase. Lane 1 = crude extract of TB1 showing background levels of porphobilinogen deaminase, lane 2 = crude extract of TB1 harbouring the BM3 plasmid showing the overexpression of wild-type deaminase, lane 3 = Dalton VII protein molecular weight standards, lanes 4 to 8 = crude extract from G57A mutant PBGD strain and lanes 9 – 12 = heat treated samples of G57A mutant crude extract.

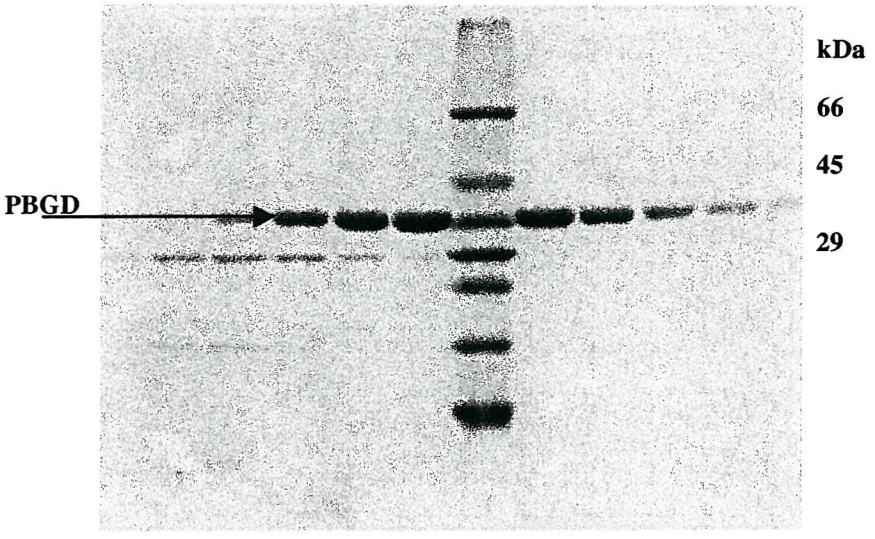


Figure 5.6 SDS gel showing the elution of the G57A mutant *E. coli* PBGD from the DEAE-Sephacel column.

5.2.3 Kinetic analysis of the G57A mutant *E. coli* porphobilinogen deaminase

The specific activity and kinetic parameters were determined and compared to those of the wild-type deaminase (table 5.2). The results indicate that the G57A mutant is only 30% as active as the wild type enzyme. Interestingly, the K_m for porphobilinogen is almost 6 times higher showing that the affinity of the substrate for the enzyme is substantially reduced. The k_{cat} is also reduced by approximately 5 times. Thus a conservative mutation of glycine to alanine has had a surprisingly large effect indicating the sensitivity of this region to any changes.

Porphobilinogen Deaminase	Specific activity ($Umg^{=1}$)	% Specific activity	K_m (μM)	K_{cat} (S^{-1})
Wild-type	42	100	20	0.5
G57A	12.5	30	117	0.11

Table 5.2 Kinetic parameters of the mutant G57A *E. coli* porphobilinogen deaminase in comparison to the wild-type enzyme. Specific activity is measured in μmoles of PBG formed/hour/mg.

5.2.4 Examination of the formation of enzyme: susbtrate intermediates by FPLC

The enzyme: substrate complexes, ES, ES_2 and ES_3, produced during the catalytic cycle, can be isolated by FPLC (Warren and Jordan, 1988). Wild-type and G57A mutant deaminases were mixed with varying amounts of PBG and loaded onto a Mono-Q column as described in section 2.3.13. At low substrate concentrations (1:0.5 enzyme: substrate) the G57A mutant enzyme displays normal enzyme-intermediate complex in comparison to the wild-type enzyme (figure 5.7). When the concentration of substrate was increased an accumulation of ES_3 was displayed, whilst the wild-type deaminase displayed a higher turnover (figure 5.8). The

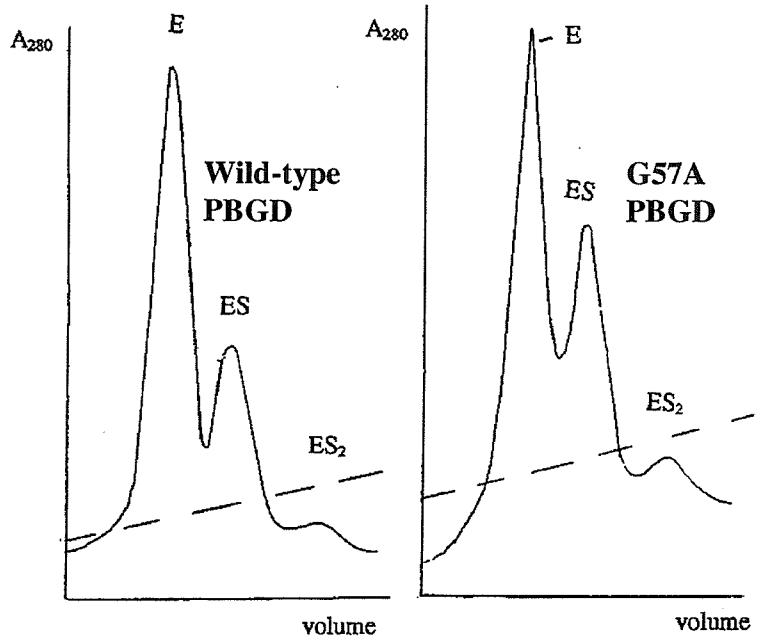


Figure 5.7 The enzyme intermediate complexes formed by the G57A and wild-type deaminases at an enzyme: substrate of 1:0.5

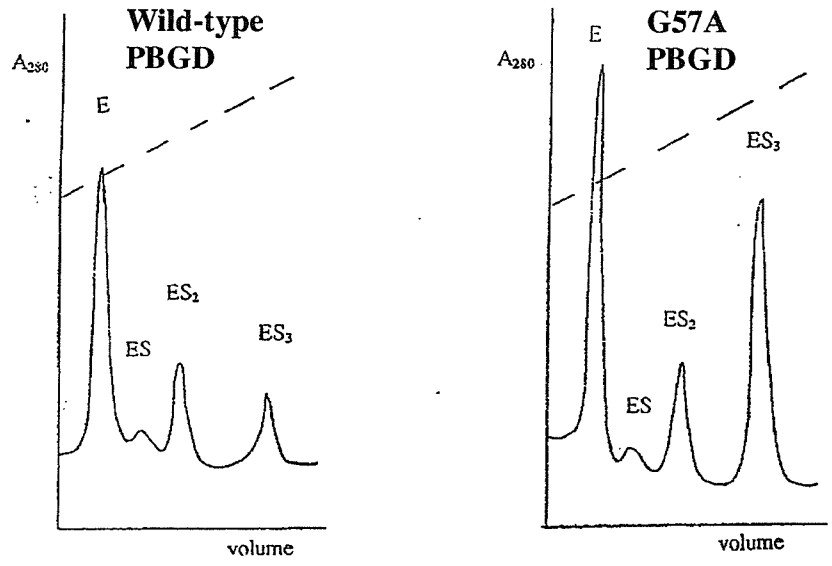


Figure 5.8 The enzyme intermediate complexes formed by the G57A and wild-type deaminases at an enzyme: substrate of 1:13

formation of ES_4 and the release of product seemed to be affected by the mutation. The introduction of alanine at a key position of the surface loop was

therefore concluded to be having a detrimental effect on the mechanism, possibly through reduction of the flexibility of the loop.

5.2.5 Presence of the dipyrromethane cofactor in the G57A mutant deaminase

To determine whether the mutation had affected the assembly of the dipyrromethane cofactor, the purified G57A mutant was reacted with Ehrlich's reagent and compared to the reaction of wild-type enzyme. Figures 5.9 a) b) show the reaction of the G57A mutant which was identical to that of the wild-type enzyme (not shown) indicating that the G57A mutant deaminase existed as a holoenzyme.

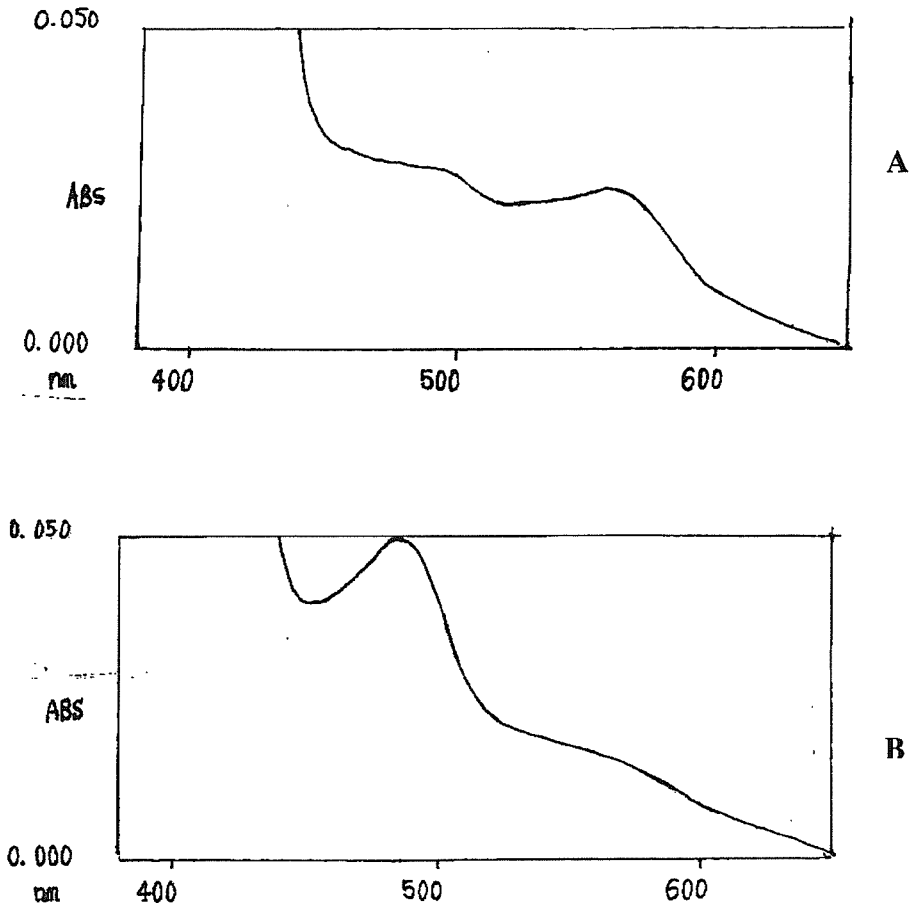


Figure 5.9 Spectra showing the reaction of G57A mutant deaminase with modified Ehrlich's reagent a) after 1minute b) after 10 minutes. The reaction was identical to that of wild-type PBGD indicating the G57A mutant deaminase exists as holoenzyme.

5.2.6 Purification of the G264C mutant *E. coli* porphobilinogen deaminase

The expression of the G264G mutant is shown in figure 5.10 and the purification in figure 5.11. The G264C mutant was purified by a similar method to that used for the wild-type enzyme (see section 2.3.13).

5.2.7 Mass spectrometric analysis of the G264C mutant *E. coli* porphobilinogen deaminase

To check the protein was indeed the G264C mutant, detailed mass spectrometric analysis was carried out to determine whether the expected increase of M_r due to the substitution of glycine for cysteine was evident. The M_r of the porphobilinogen deaminase mutant should be 45 amus higher than the wild-type enzyme with an increase from 34,272 to 34,317. The results shown in figures 5.12 and 5.13 indicate that the M_r had indeed increased by the expected 45 amus although the actual values obtained were approximately 2 amus higher for both enzymes due to slight errors in calibration. Despite this, it is clear that the G264C mutant has been obtained.

Further mass spectrometric analysis was carried out with the G264C mutant following incubation with the substrate porphobilinogen and the results were compared with the same reaction of the wild-type enzyme. The data in figure 5.14 and 5.15 show that additional species of higher Mr were present in both cases however the major species with the wild-type enzyme were E (34278.1), ES (34,483.7) and possibly ES_2 (34,690.8). The species observed with the mutant were E (34318.3) and possibly ES_2 (34735.6). The differences cannot easily be reconciled with the similarities in kinetic results although the instability of ES complexes are such that reproducibility is difficult even with measurements with the wild-type. Overall the results from mass spectrometric analysis suggest that the G264C mutant is behaving parallel to the wild-type in all aspects of function.

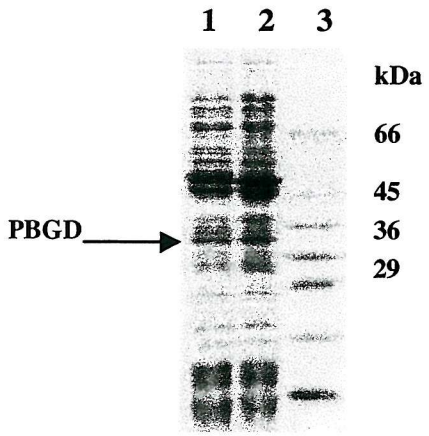


Figure 5.10 SDS PAGE showing the expression of the mutant enzyme G264C in comparison with the wild-type porphobilinogen deaminase. Lane 1 = crude extract from G264C mutant PBGD strain, lane 2 = crude extract of TB1 harbouring the BM3 plasmid showing the overexpression of wild-type deaminase, lane 3 = Dalton VII protein molecular weight standards.

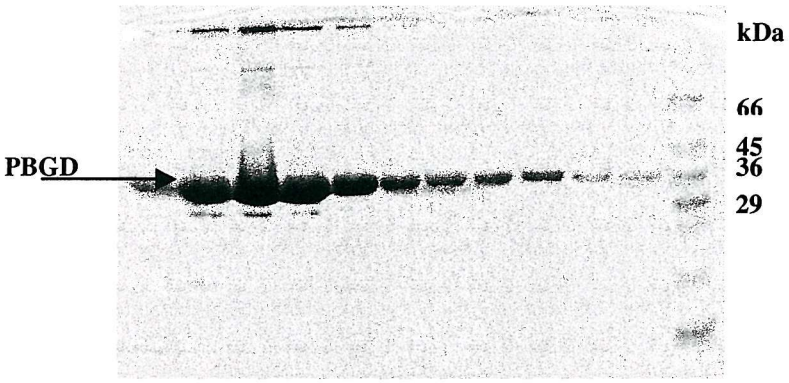


Figure 5.11 SDS-PAGE showing the elution of the G264C mutant *E. coli* PBGD from the DEAE-Sephacel column.

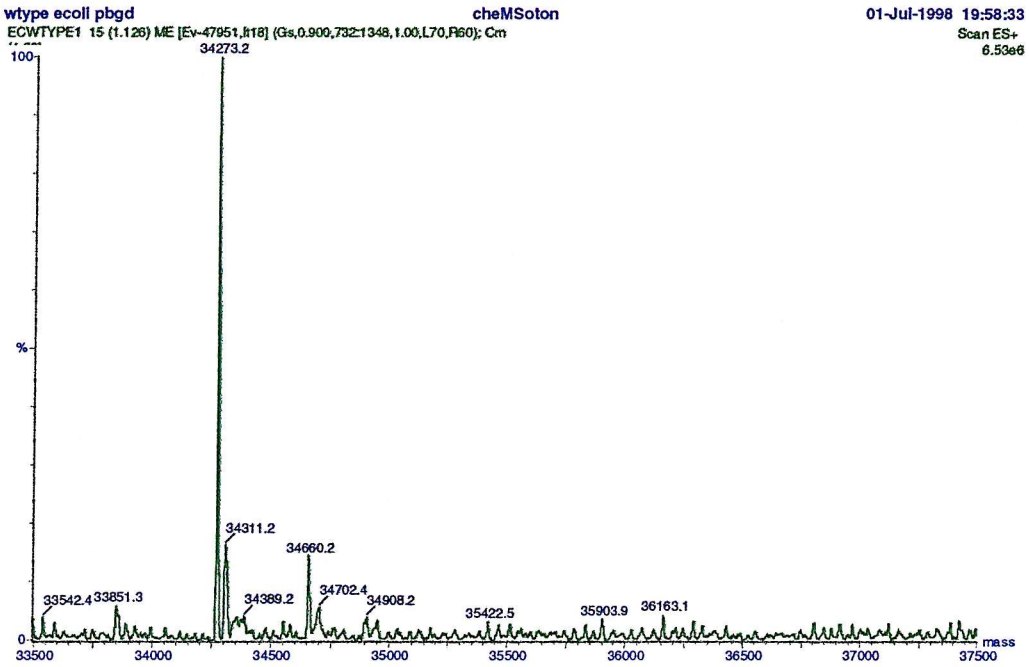


Figure 5.12 Maximum entropy deconvoluted mass spectrum of wild-type *E. coli* PBGD. The main peak at 34,273.2 corresponds to enzyme.

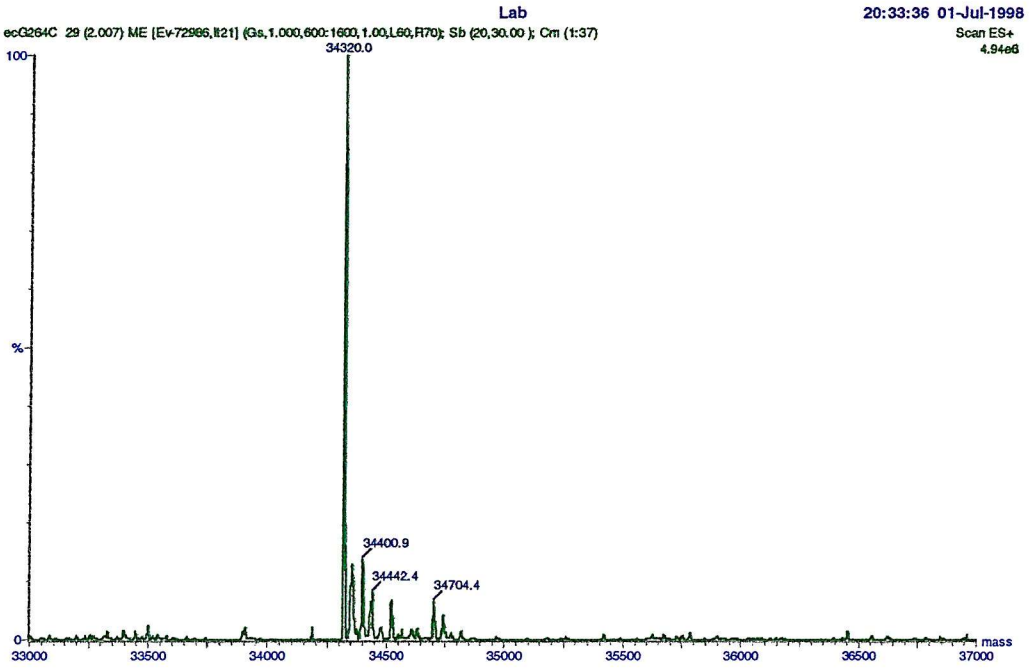


Figure 5.13 Maximum entropy deconvoluted mass spectrum of *E. coli* mutant G264C. The main peak at 34,320 Da is the mutant enzmye.

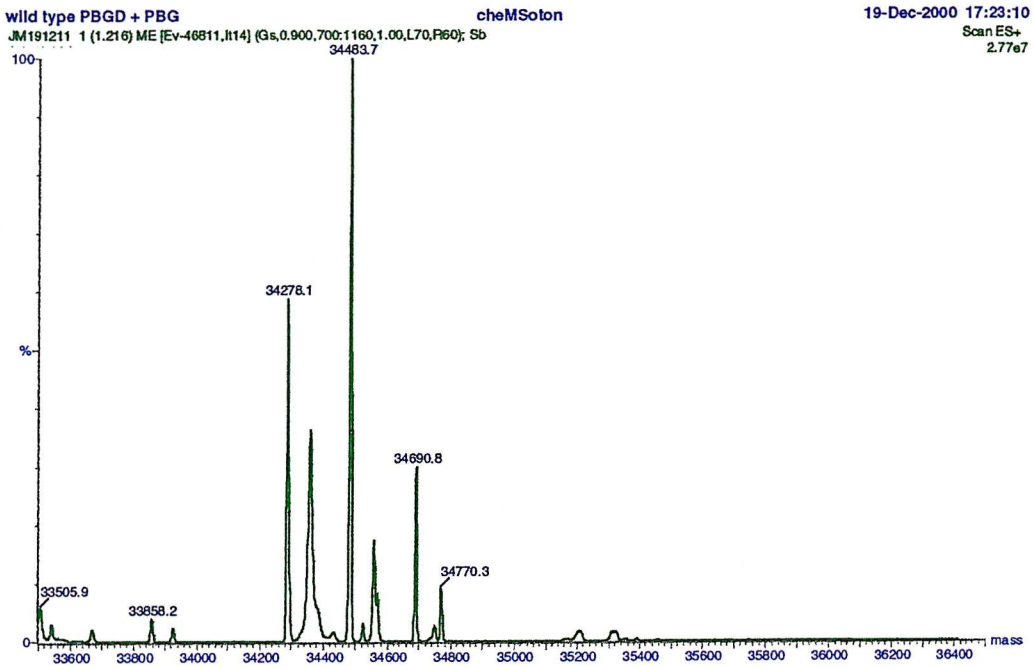


Figure 5.14 Maximum entropy deconvoluted mass spectrum of *E. coli* PBGD mixed with a 5 molar excess of PBG.

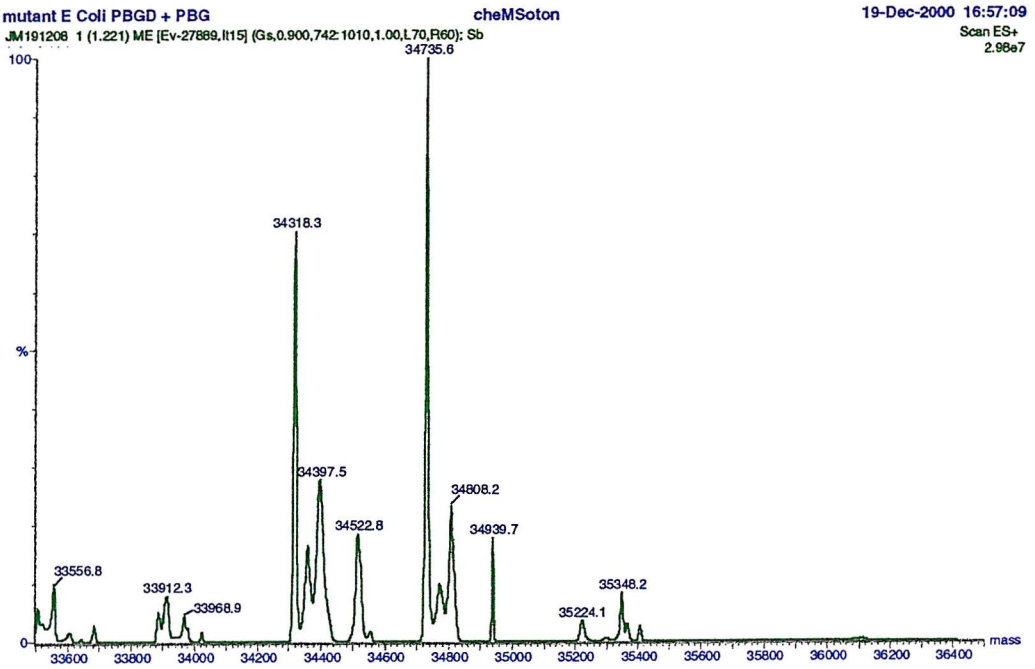


Figure 5.15 Maximum entropy deconvoluted mass spectrum of *E. coli* mutant G264C mixed with a five molar excess of PBG.

5.2.8 Kinetic analysis of the G264C mutant *E. coli* porphobilinogen deaminase

The specific activity and kinetic parameters were determined and compared to those of the wild-type deaminase (table 5.3). The results indicate that the G264C mutant is essentially as active as the wild type enzyme with a similar K_m for porphobilinogen and k_{cat} value. Thus the mutation of glycine to cysteine appears not to have had any appreciable effect on the enzyme activity.

The *moving chain* mechanism demands that domain 3, with its attached dipyrromethane cofactor, moves away from domain 2 during the formation of ES, ES_2 ES_3 and ES_4 in the catalytic cycle. From a detailed study of the three-dimensional model of *E. coli* porphobilinogen deaminase, site directed mutation of glycine 264 to cysteine in domain 3 should generate a sulphydryl group with perfect geometry for S-S bond formation with cysteine 134 in domain 2. If conformational flexibility between domains 2 and 3 was essential then cross-linking domains 2 and 3 with a disulphide bridge would be expected to inhibit polymer formation and catalytic activity completely.

The observation that the purified G264C mutant enzyme appeared in all ways to be very similar to the wild-type deaminase, with identical specific activity and kinetic parameters, suggested that either the mutagenesis had not been successful, despite the rigorous checks, or that the formation of the desired S-S bond had not occurred. It was also possible that if the S-S bond linking domains 2 and 3 had been formed then a *moving chain* mechanism may not be operative.

Porphobilinogen Deaminase	Specific activity (Umg^{-1})	% Specific activity	K_m (μM)	K_{cat} (S^{-1})
Wild-type	42	100	20	0.5
G264C	41	100	19	0.47

Table 5.3 Kinetic parameters of the mutant G264C *E. coli* porphobilinogen deaminase in comparison to the wild-type enzyme. Specific activity is measured in μmoles of PBG formed/hour/mg.

5.2.9 Presence of the dipyrromethane cofactor in the G264C mutant deaminase

To determine whether the mutation had affected the assembly of the dipyrromethane cofactor, the purified G264C mutant was reacted with Ehrlich's reagent and compared to the similar reaction of wild-type enzyme.
The figure 5.16 shows that the reaction was normal indicating that the G264C mutant deaminase existed as a holoenzyme. The reaction of the G264C mutant enzyme was identical to the wild-type enzyme consistent with normal kinetic properties.

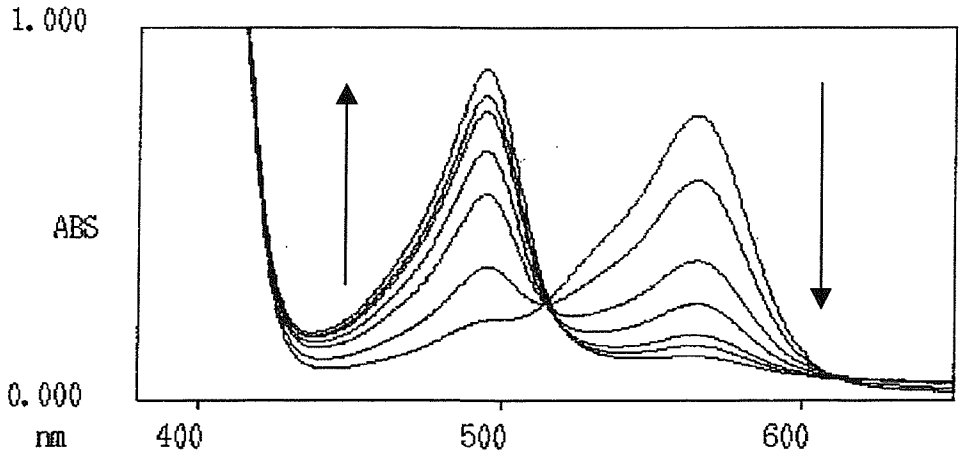


Figure 5.16 Spectra showing the reaction of G264C mutant deaminase with modified Ehrlich's reagent over 10 minutes. The reaction, monitored over 10 minutes, shows a gradual decrease in the peak at 565 nm and subsequent increase in the peak at 495 nm.

5.2.10 Analysis of the G264C mutant PBGD with DTNB

To determine whether the G264C mutation had led to the formation of an interdomain S-S link, the purified G264C mutant and the wild type enzymes were titrated with DTNB, a reagent used for the quantification of thiolate groups. This reaction is essentially one of disulphide exchange. Initially the disulphide linkage of DTNB is subject to nucleophilic attack by the thiolate ion of the protein molecule forming a new disulphide linkage between the protein sulphydryl group and thionitrobenzoate resulting in the release of a thionitrobenzoate anion. This possesses a characteristic absorption maximum at 412 nm with an extinction coefficient of 10,600 $M^{-1}cm^{-1}$ at pH 8.0 (figure 5.17). The number of cysteine residues modified can be quantified by the colourimetric determination of the amount of thionitrobenzoate ion released. This method of sulphydryl group determination is aided by the equilibrium of the reaction being displaced well to the right at pH8.0.

DTNB modification was carried out in the absence of any reducing agents to assess if the G264C mutant PBGD contained a disulphide bridge between the cysteine 264 and cysteine 134 residues. From a consideration of the primary structure of *E. coli* porphobilinogen deaminase, four cysteine residues are present at positions 99, 134, 206 and 242. Since cysteine 242 is linked to the dipyrromethane cofactor only 3 sulphydryl groups are available for titration with DTNB. These are shown in figure 5.18.

Addition of DTNB to both the wild-type and mutant enzyme gave no significant change in the spectrum at 412nm consistent with the expectation from X-ray analysis that the cysteine residues are in buried and inaccessible locations. To overcome this problem, the enzymes were denatured by treatment with SDS prior to titration with DTNB. Under these conditions the DTNB reaction produced an increase in absorbance corresponding to 3 sulphydryl groups in the wild-type protein and 4 sulphydryl groups in the mutant protein. These observations indicate conclusively that the mutant deaminase does indeed contain an additional cysteine residue at position 264 but that, despite its location close to cysteine 134, no S-S bond formation had occurred. The lack of an S-S bond is

consistent with the normal kinetic parameters of the mutant protein and therefore provides no information about whether the *moving chain* mechanism is operative.

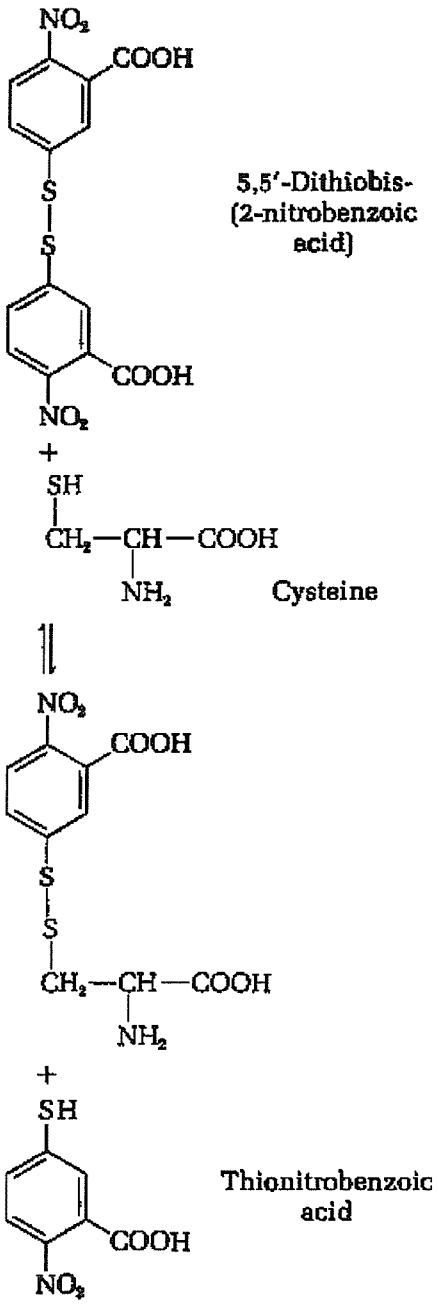


Figure 5.17 The reaction of cysteine with DTNB. This reaction causes the release of a molecule of thionitrobenzoic acid, which at pH8.0 has an intense absorption at 412nm.

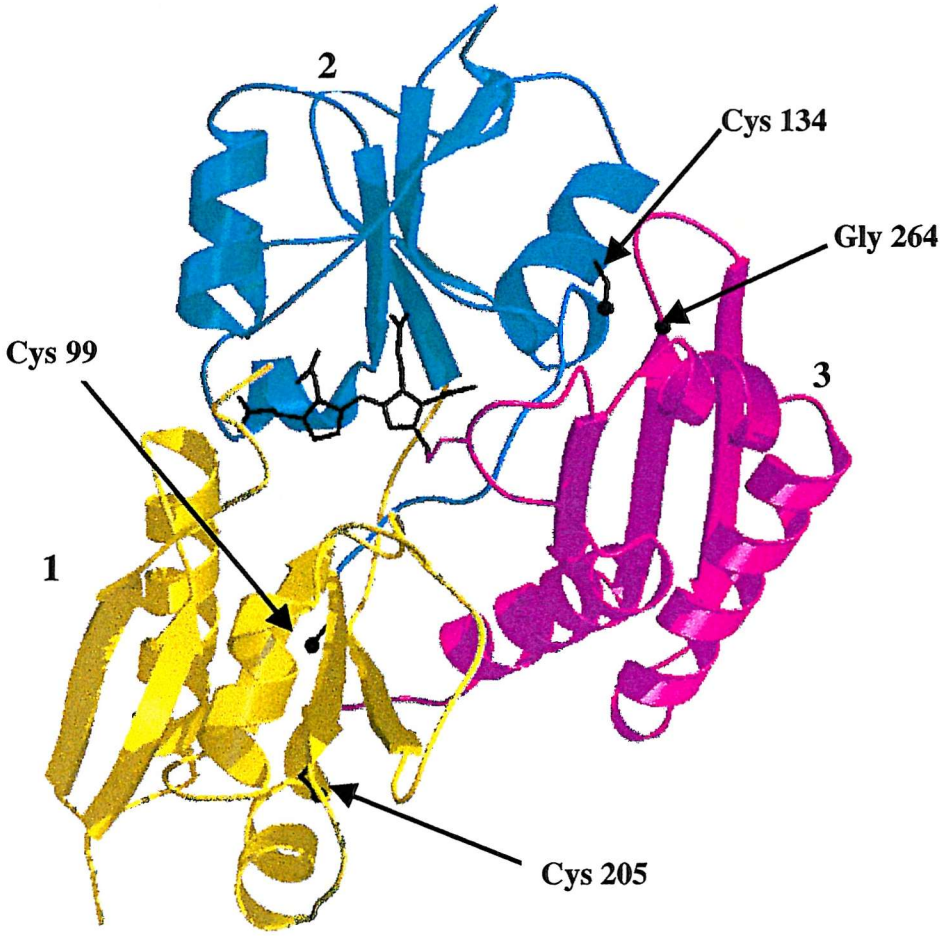


Figure 5.18 The three-dimensional structure of *E. coli* PBGD showing the three cysteines residues (coloured in black) available for reaction with DTNB. Cysteine 99 and cysteine 205 are in domain 1 (coloured yellow). Cysteine 134 is in domain 2 (coloured in cyan). Also glycine 264 is shown in domain 3 (coloured in purple). This residue is substituted with a fourth cysteine in the G264C mutant deaminase.

5.2.11 Attempts to react cysteine 134 with cysteine 264 using cross-linking reagents

Treatment of the mutant porphobilinogen deaminase with oxidising agents is fraught with difficulties because of the sensitivity of the dipyrromethane cofactor to oxidation to a dipyrromethene or dipyromethanone, with a resulting loss of activity. Therefore attempts were made to cross-link cysteines 134 and 264 by reaction with 1,3-dibromopropan-2-one, a bifunctional thiophilic reagent known to cross-link adjacent sulphydryl groups in fatty acid synthases (Stoops and Wakil, 1981) and polyketide synthases (Spencer and Jordan, 1992). When either wild-type or mutant porphobilinogen deaminase were treated with 5μM reagent, no effect on the enzyme activity was observed. Mass spectrometric analysis of the treated enzymes indicated that no modification had occurred to either (figure 5.19).

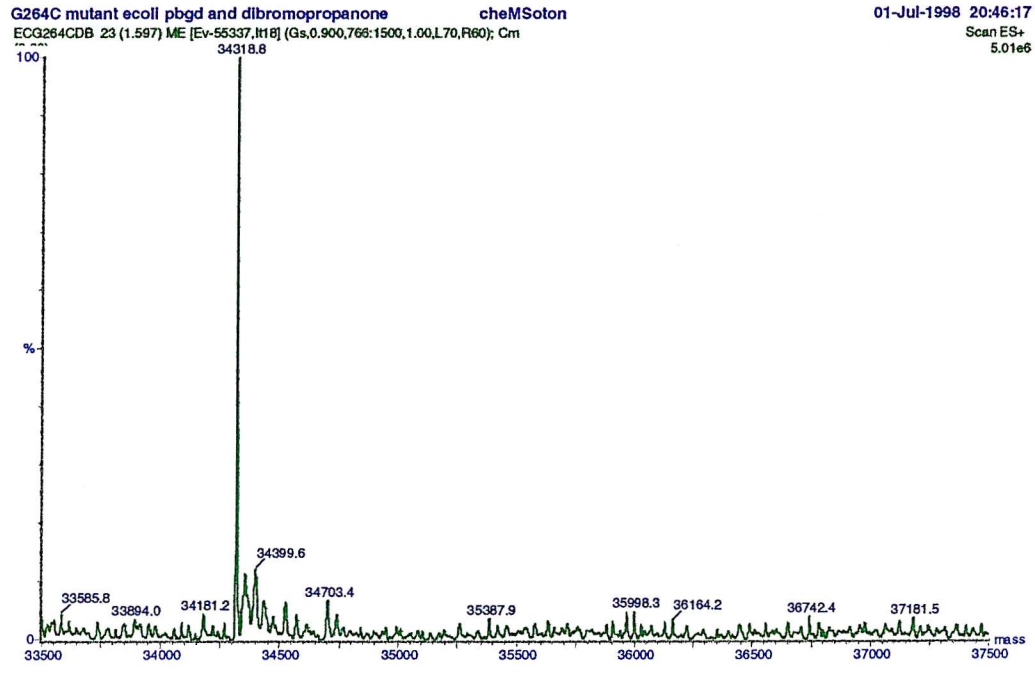


Figure 5.19 Mass spectral analysis of G264C mutant PBGD after reaction with dibromopropanone. There is no change in the mass of the mutant deaminase indicating there is no reaction with dibromopropanone.. Similar results were observed with the wild-type *E. coli* PBGD.

5.3 Conclusions

The results presented indicate that the desired mutations were generated in both cases. In the G57A mutant, the substitution of a methyl group led to a significant reduction of enzyme activity, however, further work is needed to investigate the effects of substitution by amino acids with larger side chains and the possible effect of charged groups. More importantly, further attempts are needed to freeze the active site loop from residues 48-56 so that this important part of the active site structure can be mapped accurately on the 3-dimensional model. Only then will the significance of this part of the enzyme be elucidated. Mutations of other amino acids in this active site loop may possibly be useful as an aid to its visualisation in the crystal structure.

In the case of the G264C mutant, further work is needed to find a method to oxidise the sulphydryl groups at positions 134 and 264 so that an S-S bond can be formed. The difficulty of obtaining access to buried groups by an oxidising agent is a challenge that needs s further investigation. Small oxidising agents or the use of electrochemical techniques may enable the S-S bond to be established so that its effects can be determined. If this can be accomplished, then a decision between the *moving chain* and *sliding active site* mechanisms may be possible. Further investigations at points where the enzyme may pivot on domain linking strands may also prove of value in investigating the nature of any conformational changes during the catalytic cycle. In this connection the double cross-over residues between domains 1 and 2 (residues 100-104 and residues 194-199) deserve special consideration.

References

Anderson,P.M. and Desnick,R.J. (1980) *Journal of Biological Chemistry*, **255**, 1993-1999.

Bagust,J., Jordan,P.M., Kelley,M.E.M. and Kerkut,G.A. (1985) *Neuroscience letters*, **21**.

Battersby,A.R. (1978) *Experientia*, **34**, 1-13.

Battersby,A.R., Fookes,C.J.R., Matcham,G.W.J. and McDonald,E. (1979) *J.Chem.Soc.Chem.Commun*, 539-541.

Beale,S.I. and Castelfranco,P.A. (1973) *Biochemical & Biophysical Research Communications*, **52**, 143-149.

Beale,S.I., Gough,S.P. and Granick,S. (1975) *Proc.Natl.Acad.Sci.U.S.A*, **72**, 2719-2723.

Berry,A., Jordan,P.M. and Seehra,J.S. (1981) *FEBS Lett.*, **129**, 220-224.

Bogorad,L. and Granick,S. (1953) *Proc.Natl.Acad.Sci.U.S.A*, **39**, 1176-1188.

Bradford,M.M. (1976) *Analytical Biochemistry* , **72**, 248-254

Braithwaite,O. and Russell,C.S. (1987) *Federation proceedings*, **46**, 2264.

Brownlie,P.D., Lambert,R., Louie,G.V., Jordan,P.M., Blundell,T.L., Warren,M.J., Cooper,J.B. and Wood,S.P. (1994) *Protein Sci.*, **3**, 1644-1650.

CCP4 Collaborative computational Project, Number 4 (1994), Acta. Cryst. D50. 760.

Chretien,S., Dubart,A., Beaupain,D., Raich,N., Grandchamp,B., Rosa,J., Goossens,M. and Romeo,P.H. (1988) *Proc.Natl.Acad.Sci.U.S.A*, **85**, 6-10.

Correa,G., Sr., Rossetti,M.V. and Batlle,A.M. (1991) *Z.Naturforsch.[C.]*, **46**, 1017-1023.

Dailey,H.A., Jones,C.S. and Karr,S.W. (1989) *Biochimica et Biophysica Acta*, **999**, 7-11.

Davies,R.C. and Neuberger,A. (1973) *Biochemical Journal*, **133**, 471-492.

De Siervi,A., Mendez,M., Parera,V.E., Varela,L., Battle,A.M. and Rossetti,M.V. (1999) *Human Mutation*, **14**, 355.

Desnick,R.J., Ostasiewicz,L.T., Tishler,P.A. and Mustajoki,P. (1985) *J.Clin.Invest*, **76**, 865-874.

Dunbar,B. and Wilson,S.B. (1994) *Analytical Biochemistry*, **216**, 227-228.

Elder,G.H. (1993) *Journal of Clinical Pathology*, **46**, 977-981.

Elder,G.H. (1998) *Clinics in Dermatology*, **16**, 225-233.

Elder,G.H., Hift,R.J. and Meissner,P.N. (1997) *Lancet*, **349**, 1613-1617.

Evans,J.N., Burton,G., Fagerness,P.E., Mackenzie,N.E. and Scott,A.I. (1986) *Biochemistry*, **25**, 905-912.

Ferreira,G.C., Vajapey,U., Hafez,O., Hunter,G.A. and Barber,M.J. (1995) *Protein Science*, **4**, 1001-1006.

Gorchein,A. (1997) *Br.J.Clin.Pharmacol.*, **44**, 427-434.

Grandchamp,B., De Verneuil,H., Beaumont,C., Chretien,S., Walter,O. and Nordmann,Y. (1987) *Eur.J.Biochem.*, **162**, 105-110.

Hadener,A., Matzinger,P.K., Malashkevich,V.N., Louie,G., Wood,S.P., Oliver,P., Alefounder,P.R., Pitt,A.R., Abell,C. and Battersby,A.R. (1993) *Eur.J.Biochem.*, **211**, 615-624.

Hadener,A., Alefounder,P.R., Hart,G.J., Abell,C. and Battersby,A.R. (1990) *Biochemical Journal*, **271**, 487-491.

Hart,G.J., Abell,C. and Battersby,A.R. (1986) *Biochem.J.*, **240**, 273-276.

Hart,G.J., Leeper,F.J. and Battersby,A.R. (1984) *Biochem.J.*, **222**, 93-102.

Higuchi,M. and Bogorad,L. (1975) *Annals of the New York Academy of Sciences*, **244**, 401-418.

Hunter,G.A. and Ferreira,G.C. (1999) *Biochemistry*, **38**, 12526.

Jones,R.M. and Jordan,P.M. (1994) *Biochem.J.*, **299**, 895-902.

Jordan,P.M. and Berry,A. (1981) *Biochem.J.*, **195**, 177-181.

Jordan,P.M. and Seehra,J.S. (1979) *FEBS Letters*, **104**, 364-366.

Jordan,P.M. and Shemin,D. (1973) *Journal of Biological Chemistry*, **248**, 1019-1024.

Jordan,P.M., Thomas,S.D. and Warren,M.J. (1988**a**) *Biochem.J.*, **254**, 427-435.

Jordan,P.M. and Warren,M.J. (1987) *FEBS Lett.*, **225**, 87-92.

Jordan,P.M., Warren,M.J., Williams,H.J., Stolowich,N.J., Roessner,C.A., Grant,S.K. and Scott,A.I. (1988**b**) *FEBS Lett.*, **235**, 189-193.

Jordan,P.M. and Woodcock,S.C. (1991) *Biochem.J.*, **280**, 445-449.

Juknat,A.A., Dornemann,D. and Senge,H. (1994) *Planta*, **193**, 123-132.

Kannangara,C.G., Gough,S.P., Bruyant,P., Hoober,J.K., Kahn,A. and von Wettstein,D. (1988) *Trends in Biochemical Sciences*, **13**, 139-143.

Kappas,A., Sassa,S., Galbraith,R.A. and Nordmann,Y. (1989) *The porphyria 6th ed, (Scriver, C.R., Beaudet, A.L., Sly, W.S.and Valle, D.,eds)*, **McGraw-Hill: New York**, 1305-1365.

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

Lander,M., Pitt,A.R., Alefounder,P.R., Bardy,D., Abell,C. and Battersby,A.R. (1991) *Biochemical Journal*, **275**, 447-452.

Lee,J. and Anvret,M. (1991) *Proc.Natl.Acad.Sci.U.S.A*, **88**, 10912-10915.

Leslie, A.G.W (1991) Molecular data processing in crystallographic computing (Moras, D., Podjarny, A.D. and Thierry, J.C. Eds) pp 50 – 61, IUC Oxford University Press, Oxford.

Louie,G. (1993) *Curr.Opin.Struct.Biol.*, **3**, 401-408.

Louie,G.V., Brownlie,P.D., Lambert,R., Cooper,J.B., Blundell,T.L., Wood,S.P., Malashkevich,V.N., Hadener,A., Warren,M.J. and Shoolingin-Jordan,P.M. (1996) *Proteins*, **25**, 48-78.

Louie,G.V., Brownlie,P.D., Lambert,R., Cooper,J.B., Blundell,T.L., Wood,S.P., Warren,M.J., Woodcock,S.C. and Jordan,P.M. (1992) *Nature*, **359**, 33-39.

MacFerrin,K.D., Terranova,M.P., Schreiber,S.L. and Verdine,G.L. (1990) *Proc.Natl.Acad.Sci.U.S.A*, **87**, 1937-1941.

Mathewson,J.H. and Corwin,A.H. (1963) *J.Am.Chem.Soc.*, **83**, 135-137.

Mazzetti,M.B. and Tomio,J.M. (1988) *Biochim.Biophys.Acta*, **957**, 97-104.

Mazzetti,M.B. and Tomio,J.M. (1997) *Biochemistry and Molecular Biology International*, **42**, 685-692.

Mazzetti,m.m. and Tomio,j.m. (1993) *An.Asoc.quim.Argent*, **81**, 381-389.

May,B.K., Bhasker,C.R., Bawden,M.J. and Cox,T.C. (1990) *Mol.Biol.Med.*, **7**, 405-421.

Mazzetti,M.B. and Tomio,J.M. (1990) *Biochem.Int.*, **21**, 463-471.

McPherson,A. (1976) *Methods of Biochemical Analysis*, **23**, 249-345.

Messing,J. (1983) *Methods in Enzymology*, **101**, 20-78.

Mgbeje,B. (1989) Ph.D. Thesis. University of Southampton.

Mohammed,F. (2001). Ph.D. Thesis. University of Southampton.

Moore,M.R. (1993) *Int.J.Biochem.*, **25**, 1353-1368.

Nakamaye,K.L. and Eckstein,F. (1986) *Nucleic Acids Research*, **14**, 9679-9698.

Namba,H., Narahara,K., Tsuji,K., Yokoyama,Y. and Seino,Y. (1991) *Cytogenet.Cell Genet.*, **57**, 105-108.

Pluscec,J. and Bogorad,L. (1970) *Biochemistry*, **9**, 4736-4743.

Radmer,R. and Bogorad,L. (1972) *Biochemistry* , **11**, 904-910.

Raich,N., Romeo,P.H., Dubart,A., Beaupain,D., Cohen-Solal,M. and Goossens,M. (1986) *Nucleic Acids Res.*, **14**, 5955-5968.

Raux,E., Schubert,H.L. and Warren,M.J. (2000) *Cellular & Molecular Life Sciences*, **57**, 1880-1893.

Russell,C.S. and Rockwell,P. (1980) *FEBS Letters*, **116**, 199-202.

Sano,S. and Granick,S. (1961) *J.Biol.Chem.*, **236**, 1173-1180.

Sassa,S. and Kappas,A. (2000) *Journal of Internal Medicine*, **247**, 169-178.

Schon,A., Krupp,G., Gough,S., Berry-Lowe,S., Kannangara,C.G. and Soll,D. (1986) *Nature*, **322**, 281-284.

Scott,A.I., Clemens,K.R., Stolowich,N.J., Santander,P.J., Gonzalez,M.D. and Roessner,C.A. (1989) *FEBS Lett.*, **242**, 319-324.

Scott,A.I., Roessner,C.A., Stolowich,N.J., Karuso,P., Williams,H.J., Grant,S.K., Gonzalez,M.D. and Hoshino,T. (1988) *Biochemistry*, **27**, 7984-7990.

Sharp,P.M., Cowe,E., Higgins,D.G., Shields,D.C., Wolfe,K.H. and Wright,F. (1988) *Nucleic Acids Res.*, **16**, 8207-8210.

Shemin,D. and Russell,C.S. (1953) *J.Am.Chem.Soc.*, **75**, 4873-4875.

Shioi,Y., Nagamine,M., Kuraki M and Sasa,T. (1980) *J.Biol.Chem.*, **255**, 1993-1999.

Shoolingin-Jordan,P.M. (1998) *Biochemical Society Transactions*, **26**, 326-336.

Shoolingin-Jordan,P. and Cheung,K.M. (1999) *Comprehensive Natural Products Chemistry*, **4**, (*Barton, D., Nakanishi, K. and Meth-Cohn, O., eds*), **Elsevier:-Amsterdam**, 61-108.

Smythe,E. and Williams,D.C. (1988) *Biochemical Journal*, **251**, 237-241.

Solis,C., Lopez-Echaniz,I., Sefarty-Graneda,D., Astrin,K.H. and Desnick,R.J. (1999) *Molecular Medicine*, **5,** 664-671.

Spano,A.J. and Timko,M.P. (1991) *Biochim.Biophys.Acta*, **1076**, 29-36.

Spencer,J.B. and Jordan,P.M.(1992), *Biochemical Journal*, **288**, (3), 839-846.

Spencer,P. and Jordan,P.M. (1995) *Biochemical Journal*, **305**, 151-158.

Stoops,J.K. and Wakil,S.J. (1981) *Journal of Biological Chemistry*, **256**, 8364-8370.

Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorf,J.W. (1990) *Methods Enzymol.*, **185**, 60-89.

Thomas,S.D. and Jordan,P.M. (1986) *Nucleic Acids Res.*, **14**, 6215-6226.

Thunell,S. (2000) *Scandinavian Journal of Clinical & Laboratory Investigation*, **60**, 509-540.

Ullmann,A., Jacob,F.and Monod,J. (1967) *Journal of Molecular Biology*, **24**, 339-343.

Waldenstrom,J. (1956) *Acta Genet.Stat.Med.*, **6**, 122-131.

Wang,D.W., Driessen,H.P. and Tickle,I.J. (1991) *Journal of Molecular Graphics*, **9**, 50-52.

Warren,M., Jay,M., Hunt,D.M., Elder,G. and Rohl,J.C. (1996) *Trends in Biochemical Sciences*, **21**, 229-234.

Warren,M.J., Gul,S., Aplin,R.T., Scott,A.I., Roessner,C.A., O'Grady,P. and Shoolingin-Jordan,P.M. (1995) *Biochemistry*, **34**, 11288-11295.

Warren,M.J. and Jordan,P.M. (1988) *Biochemistry*, **27**, 9020-9030.

Williams,D.C. (1984) *Biochem.J.*, **217**, 675-678.

Williams,D.C., Morgan,G.S., McDonald,E. and Battersby,A.R. (1981) *Biochem.J.*, **193**, 301-310.

Wood,S.P., Lambert,R. and Jordan,P.M. (1995) *Molecular Medicine Today*, **1**, 232-239.

Woodcock,D.M., Crowther,P.J., Doherty,J., Jefferson,S., DeCruz,E., Noyer-Weidner,M., Smith,S., Michael,M.Z. and Graham,M.W. (1989) *Nucleic Acids Research*, **17**, 3469-3478.

Woodcock,S.C. and Jordan,P.M. (1994) *Biochemistry*, **33**, 2688-2695.

Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103-119.