Molecular and Kinetic Characterisation of wild type and mutant Porphobilinogen Deaminase

Elaine Pienaar
PNRELA002

University of Cape Town

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
in fulfilment of the requirements for the degree
MSc (MED) in Medical Biochemistry

Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN

February 2016

Supervisor:
Professor Peter Meissner
UCT Porphyria Laboratory, Groote Schuur Hospital
and Department of Integrative Biomedical Sciences, University of Cape Town

Co-Supervisor:
Professor Edward Sturrock
Department of Integrative Biomedical Sciences, University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

To my grandfather,

For always being our number one cheerleader.

Plagiarism Declaration

I, Elaine Pienaar, declare that this research project is my own original work, except where acknowledgements are indicated. This work, as a whole, or in part, has not been, and will not be submitted to another University for the purpose of receiving a degree.

Citations and references are done according to information for authors from The Biochemical Journal.

I approve and encourage the University of Cape Town to reproduce this content, as a whole or in part, for the purpose of research.

Name: Elaine Pienaar

Student Number: PNRELA002

Signature: Signed by candidate

Date: 16 May 2016

Table of Contents

List of Abbreviations

General

AIP	acute intermittent porphyria
ALA	5-aminolivulinic avid
bp	base pairs
BSA	bovine serum albumin
CAF	Central Analytical Facility
cDNA	complementary DNA
d.H_2O	distilled water
Da	Dalton
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
dsDNA	double-stranded DNA
DTT	1,4-dithiol-DL-threitol
DPM	dipyrromethane
ES	enzyme-substrate complex
EDTA	ethylenediaminetetra-acetic acid
f_F	fraction of folded protein
f_U	fraction of unfolded protein
g	gram(s)/centrifugal force
HMB	hydroxymethylbilane
HMBS	hydroxymethylbilane synthase
h	hour(s)
kb	kilobase pairs
k_{cat}	catalytic constant
kDa	kiloDalton
K_M	Michaelis-Menten constant
L	Litre(s)
LB	Luria-Bertani
(θ)	ellipticity
$\Delta\varepsilon$	molar circular dichroism
$[\theta]_{MRW}$	mean residue ellipticity
M	molar (concentration)
mA	milli Ampere
mdeg	millidegrees
mg	milligram(s)
min	minute(s)
ml	milliliter(s)
mm	millimeter
mmol	milimole(s)
mM	millimolar
mRNA	messenger RNA
NEM	N-Ethylmaleimide
ng	nanogram(s)
nm	nanometer

nmol	nanomole(s)
nM	nanomolar
O/N	overnight
PAGE	polyacrylamide gel electrophoresis
PBG	porphobilinogen
PBGD	porphobilinogen deaminase
PCR	polymerase chain reaction
PDB	protein data bank
pmol	picomol
PMSF	phenylmethylsulfonyl fluoride
RE	restriction endonuclease
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
S	substrate
SDS	sodium dodecyl sulphate
SE	standard error
s	seconds
SNP	single nucleotide polymorphism
SU	Stellenbosch University
T_a	annealing temperature
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
T_m	melting temperature
U	unit
UCT	University of Cape Town
V	volts
V_{max}	maximum velocity (kinetics)
WT	wild type
µg	microgram(s)
µl	microliter(s)
µM	micromolar
^oC	degree Celsius
3’	3 prime (also downstream)
5’	5 prime (also upstream)

DNA nucleotide bases

A	adenine
C	cytosine
G	guanine
T	thymine

Amino Acids

A	Ala	alanine
C	Cys	cysteine
D	Asp	aspartate
E	Glu	glutamate
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methione
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine

Enzymes

ALAD	5-aminolevulinic acid dehydratase
ALAS	5-aminolevulinic acid synthase
CPOX	oproporphyrinogen oxidase
FC	ferrochelatase
PBGD	porphobilinogen deaminase
PPOX	protoporhyrinogen oxidase
$UROS_{III}$	uroporphorinogen III synthase
UROD	uroporphorinogen decarboxylase

Haem Precursors

ALA	5-aminolevulinic acid
CPO	coproporphorinogen
HMB	hydroxymethylbilane
Preuro	preuroporphyrinogen
PBG	Porphobilinogen
Proto	protoporphyrin IX
Protogen	protoporphyrinogen IX
UROIII	uroporphyrinogen III

Porphyrias

AIP	Acute Intermittent Porphyria
ALADP	δ-Aminolevulinate Dehydratase Porphyria
CEP	Congenital Erythropoietic Porphyria
EPP	Erythropoietic protoporphyria
HCP	Hereditary Coproporphyria
PCT	Porphyria Cutanea Tarda
VP	Variegate Porphyria
XLDPP	X-linked Dominant Protoporphyria

List of Figures

List of Tables

Acknowledgments

To my supervisors Professors Peter Meissner and Edward Sturrock. Thank you Prof Pete for being an incredible mentor and supervisor. Thank you for always being available, always answering every one of my countless admin emails, and constantly reassuring me, “don’t panic”. Thank you for the opportunities I have had to learn from you during the last three years. I will treasure all your lessons and teachings. Ed, thank you for your patience and kind heart. Thank you for being the balancing voice amongst all our bouncing ideas during Friday morning meetings. For your support and willingness for me to ‘add-on’ as one of your lab members. It has been a true honour to work with you.

Anne; thank you for your continued interest in my work. For forcing me to take a day off, and for being my ‘second brain’ when I doubted myself. For motivating me to keep going when I felt like I did not want to try again. Thank you for your wisdom, unending love and support, and for teaching me to never give up on anything.

Sylva and everyone in the Sturrock lab. Thank you for making me feel part of the Sturrock lab family, including listening to my weekly venting sessions about some failed experiment. Sylva, thank you for the words of inspiration and always being open to help when I needed.

Marilyn, Sam and everyone in the Department of Surgery. Thank you for always being so obliging to me borrowing things, using equipment, asking advice, and joining in on Friday ‘cake days’. It was so nice knowing I can always pop-in and that you are all close by.

Professor Marina Rautenbach from Stellenbosch University. Thank you for your help and wonderful guidance for the circular dichroism work. For allowing me to take up countless hours of your time just to ask the same question again.

Dr. Jaco Brand and Elsa Malherbe from the NMR Central Analytical Facility at Stellenbosch University, for your help, guidance and support while working with the CD equipment. Brandon Davidson, thank you for your kind and friendly face, and listening to my stories on Friday afternoons. Dee Blackhurst, for the use of your spectrophotometer and laboratory software for our co-factor analysis. The Core DNA Sequencing Facility at Stellenbosch University, for their excellent sequencing results and fantastic service. Helene Bustad, from the University of Bergen in Norway, for taking the time to help with the native PAGE work and generously providing me with your protocol, as well as advice.

Cameron, for knowing almost as much about protein structure and kinetics as I do at this point, and for being a solid support structure throughout this project. My family and friends - thank you for your continued faith in me. Finally, our Heavenly Father, thank you for showing me the small miracles within life.

“There are only two ways to live; either as if nothing is a miracle, or as if everything is.”
~Albert Einstein~

Abstract

The purpose of this dissertation was to provide an overview of acute intermittent porphyria, focussing on the structure and function of the enzyme, porphobilinogen deaminase (PBGD), as well as experimentally demonstrating the use of kinetic, structural and thermodynamic approaches to shed light on the enzyme reaction. The key focus was to investigate the effect of three mutations of the active site lysine 98 residue (K98) on the enzyme's stability and mechanism. Two clinically relevant PBGD mutants, the K98E and K98R were expressed. Both of these mutants have previously been described in patients. We engineered and expressed an additional mutant, K98A, in order to investigate the effect of charge at this residue.

The K98E, K98R and K98A recombinant proteins were successfully engineered, expressed and purified. These mutations were kinetically characterised, and the low enzyme activity supports the fact that the K98E and the K98R are known-disease causing mutations. The negligible activity of the K98A and K98R mutants was predicted as a result of a loss of DPM co-factor binding, which was analysed and proved with a co-factor spectral shift assay. Further attempts to examine the interaction of co-factor binding involved removal of the bound co-factor from wild type enzyme, in order to investigate the possible interaction of the *'apo'*-enzyme with the DPM co-factor. However, no results were obtained to elucidate this interaction, largely due to the highly unstable nature of the generated *'apo'*-enzyme.

Native polyacrylamide gel electrophoresis (PAGE) was performed in order to observe changes in enzyme-substrate complexes between the wild type and the different mutant proteins. The enzyme-substrate complexes for the wild type were clearly shown, however we could not do so in our mutant proteins.

The secondary structure estimations as well as the conformational stability of the mutants were tested with the use of circular dichroism. Far- and near-UV analysis provided insight into the effect of each mutation on the enzyme's secondary and tertiary structure respectively. Results indicate that the different mutations cause marginal alterations in secondary structure, and resulted in changes of aromatic ring conformations in the near-UV analysis. Finally, modelling of each mutation to known crystal structures of the human enzyme was done in order to provide a rationalisation of kinetic and conformational data. Although this provided only a static image and estimation of the structural effect of each mutation, it did allow for some speculation in order to rationalise the kinetic and conformational data obtained. Overall, this work illustrates how the characterisation of expressed, purified, AIP-associated mutant enzymes aids our understanding of the complex structure and mechanism of the PBGD enzyme.

Chapter 1

Literature Review and Development of Dissertation

Introduction

This chapter provides a review of the current literature. The scope of this review includes haem, (its formation and regulation), an overview of the porphyrias (including the classification, diagnosis, and need for molecular analysis) followed by a review of acute intermittent porphyria (AIP) focussing on the molecular basis of the disease related to PBGD structure and function. The study rationale as well as research question and aims for this dissertation follow at the end of this chapter.

Haem

Haem is central to the metabolic processes of most living organisms, including electron transport and enzyme catalysis [1]. In majority of species, haem cannot be obtained through the diet, but rather has to be produced intracellularly. Indeed haem is synthesised in every cell, but the largest proportion (80%) occurs in the erythropoïetic (red blood) cells and a further (15%) in parenchymal cells within the liver [2]. Haem also serves as a prosthetic group for many and various kinds of proteins, including haemoglobin, cytochrome P450s and peroxidases [3,4]. The structure of haem is illustrated in Figure 1.1. It is comprised of a protoporphyrin IX tetrapyrrole molecule, containing ferrous iron (Fe^{2+}).

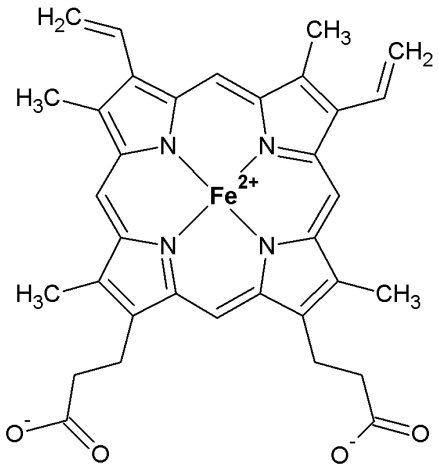


Figure 1.1 The chemical structure of haem (ChEBI_26355), with the ferrous iron (Fe^{2+}) within its tetrapyrrole centre. Image generated by ChemSketch, Version 2.0 for Windows [5]. Molfile of haem imported from Chemical Entities of Biological Interest (ChEBI) database [6].

Haem Biosynthesis

Haem is produced via a pathway consisting of a series of enzyme-catalysed reactions [3]. However, the pathway is slightly different amongst different species. In the first reaction in plants, bacteria and algae, the porphyrin precursor aminolevulinic acid (ALA), is produced from glutamate via the C5 pathway [6,7]. In mammals and fungi, however, the ALA is produced

from from glycine and succinyl coA (the citric acid cycle), by the enzyme aminolevulinic acid synthase (ALAS), followed by translocation into the cytosol to produce the second precursor, porphobilinogen (PBG) by aminolevulinic acid dehydratase (ALAD) [9,10]. PBG is then polymerized in a series of 4 steps by porphobilinogen deaminase (PBGD), forming hydroxymethylbilane (HMB), the first tetrapyrrole in the pathway.

Uroporphyrinogen III synthase (UROS) then converts the HMB to urophoryrinogen III, followed by condensation to coprophyrinogen III by urophoryrinogen decarboxylase (UROD). Coprophyrinogen III is transcolated back into the mitochondria whereby coprophyrinogen III oxidase (CPO) catalyses an oxidative decarboxylation reaction to form protoporphyrinogen IX. Protoporhyrin oxidase (PPOX) allows for the oxidation of protophyrinogen IX to protoporphyrin IX. Finally, ferrochelatase (FC) produces haem with the insertion of the ferrous iron (Fe^{2+}) into the tetrapyrrole centre of the protoporphyrin IX molecules [11,12].

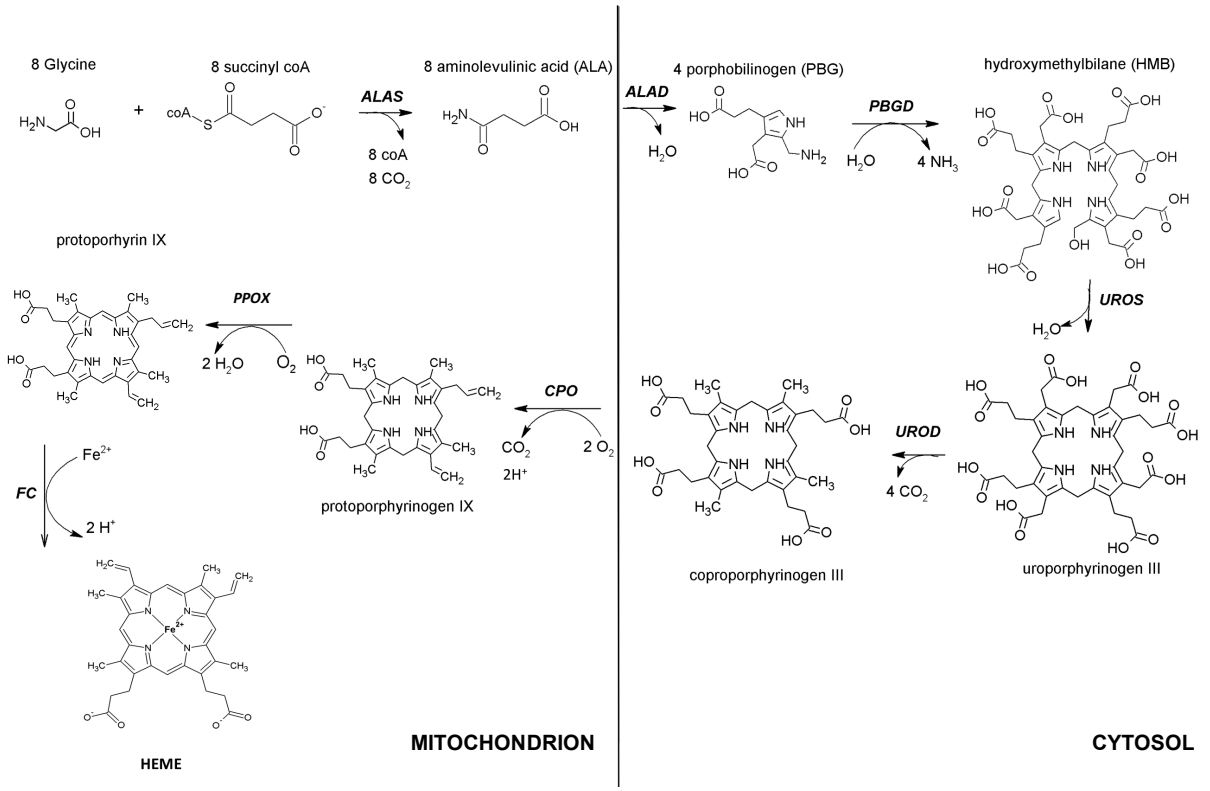


Figure 1.2 The biosynthesis of haem as adapted from [8,13]. Enzyme for each reaction is indicated in bold. Molfiles were imported from Chemical Entities of Biological Interest (ChEBI) database (Image and reaction constructed in ChemSketch, Version 2.0 for Windows [5].

Tetrapyrroles within the haem biosynthesis pathway are referred to as porphyrins, or porphyrin intermediates, whereas the two linear products, ALA and PBG, are known as the porphyrin precursors. Under reducing conditions, the porphyrin intermediates may also be referred to porphyrinogens, whereas upon oxidation, they are called porphyrins [14,15].

Regulation of Haem Biosynthesis

Maintenance of the levels of haem within a cell is crucial, as an excess of haem is toxic to cells [16]. Therefore the pathway needs to have some form of regulation. The formation of ALA in the first reaction is the rate limiting step within the haem biosynthesis pathway and the regulation of this reaction is controlled differently in erythroid and liver cells [17]. The ALAS enzyme has two isoforms, namely the ubiquitous (ALAS1) and the erythroid specific (ALAS2).

In the liver, ALAS1 may be regulated by haem itself, acting as a suppressor on the synthesis and translocation of ALAS1 to the mitochondria, and thus leading to the formation of a negative feedback cycle [18]. In erythroid cells, however, ALAS2 is only induced in cases when the cycle is already active [19]. ALAS2 is regulated transcriptionally by erythroid-specific factors, as well as cellular iron concentrations [11]. The *ALAS2* gene has an iron regulatory element (IRE) located within the 5'- of its mRNA. This allows for the inhibition of the ALAS2 translation when cellular iron concentrations are too low for optimum haem biosynthesis [20].

The Porphyrias

Overview

The porphyrias are a group of metabolic inherited diseases, involving and arising from defects in the enzymes of the haem biosynthetic pathway [21]. As described earlier, under normal circumstances, the biosynthesis of haem is tightly controlled by means of a negative feedback cycle [3]. In the case of the porphyrias, however, genetic mutations within one or more of the enzymes of this pathway lead to a loss of the enzyme's normal activity, and results in the potential failure to produce sufficient haem [22]. This potential failure to control the demand on the cycle leads to up-regulation of the pathway and to the subsequent and further build-up of the porphyrin precursors, ALA and PBG, as well as porphyrin intermediates (or porphyrinogens) further along the pathway [23].

The specific pattern of haem precursor and intermediate accumulation determines the clinical profile with either the presence of skin lesions, acute neurovisceral attacks, or both [2]. The consequent build-up of the porphyrinogens results in accumulation of the oxidised porphyrin products and results in patients suffering from skin lesions [24]. As porphyrins are light sensitive molecules, exposure of the skin, in which excess porphyrins are concentrated and/or circulating, to UV radiation results in the formation of blisters, lesions and excessive fragility [24].

The build-up of excessive ALA and PBG, the porphyrin precursors, however, places patients at risk for the life-threatening neurologic pathology known as the *'acute attack'* [25]. These acute attacks are found in the majority of patients to cause symptoms including severe abdominal pain, nausea, and vomiting, as well as hypertension and tachycardia. In addition, if unrecognised and untreated, it may result in profound sensory and motor neuropathy resulting in paralysis [26,27]. The attacks may be triggered by various factors, including porphyrinogenic drugs that induce the haem biosynthesis cycle, changes in hormonal levels (which is specifically a problem with women and their menstrual cycles), smoking or alcohol, as well as the induction of the haem oxygenase following prolonged periods of fasting or with a fever [28,29]. These factors may up-regulate expression of ALAS directly or result in the further induction of the haem biosynthesis pathway, which, in the case of porphyria, is an already compromised system. This places further stress on the pathway and with the continued build up of ALA and PBG as the system fails to produce haem, the acute attack is triggered. The only currently available option for treatment of these symptoms is the intravenous supplementation of haem with haem arginate. This allows for the bypassing of the cycle, and the supplemented haem may act on ALAS1 to suppress the cycle, thus preventing further build-up of the hazardous and potentially fatal porphyrin precursors, ALA and PBG [24,28]. Liver transplants in patients with AIP or VP, who experienced severe and frequent attacks, have shown to be successful in a few cases. Their ALA and PBG levels returned to normal levels following 24 hours post-transplant, and there were no reports of subsequent attacks [30,31]. However, a liver transplantation is a high-risk procedure, and is only considered as a last resort [32].

Classification of the Porphyrias

The porphyrias are generally classified as acute or cutaneous depending on the presence or absence of acute attacks [33]. The cutaneous porphyrias include, X-linked protoporphyria (XLDPP), congenital erythropoïetic porphyria (CEP), erythropoïetic protoporphyria (EPP), hepatoerythropoïetic porphyria (HEP) and porphyria cutanea tarda (PCT), whereas the four acute porphyrias are: acute intermittent porphyria (AIP), δ-aminolevulinic acid deficiency porphyria (ADP), hereditary coproporphyria (HCP) and variegate porphyria (VP) [34,35].

From a clinical perspective, the classification according to the risk of acute attacks is preferred, however, there are two alternative ways of grouping the porphyrias - either according to the origin of the porphyrin precursors or intermediate build-up (hepatic or erythroid), or by the mode of inheritance (dominant, recessive, autosomal or sex-linked) [32]. From a biochemical point of view, it makes sense to classify the porphyrias simply by the specific enzyme deficiency – thus VP would be PPOX-deficiency porphyria, AIP a PBGD-deficiency porphyria and so on. A

simplified overview of each of the enzymes within the pathway and the respective porphyria is shown in Figure 1.3.

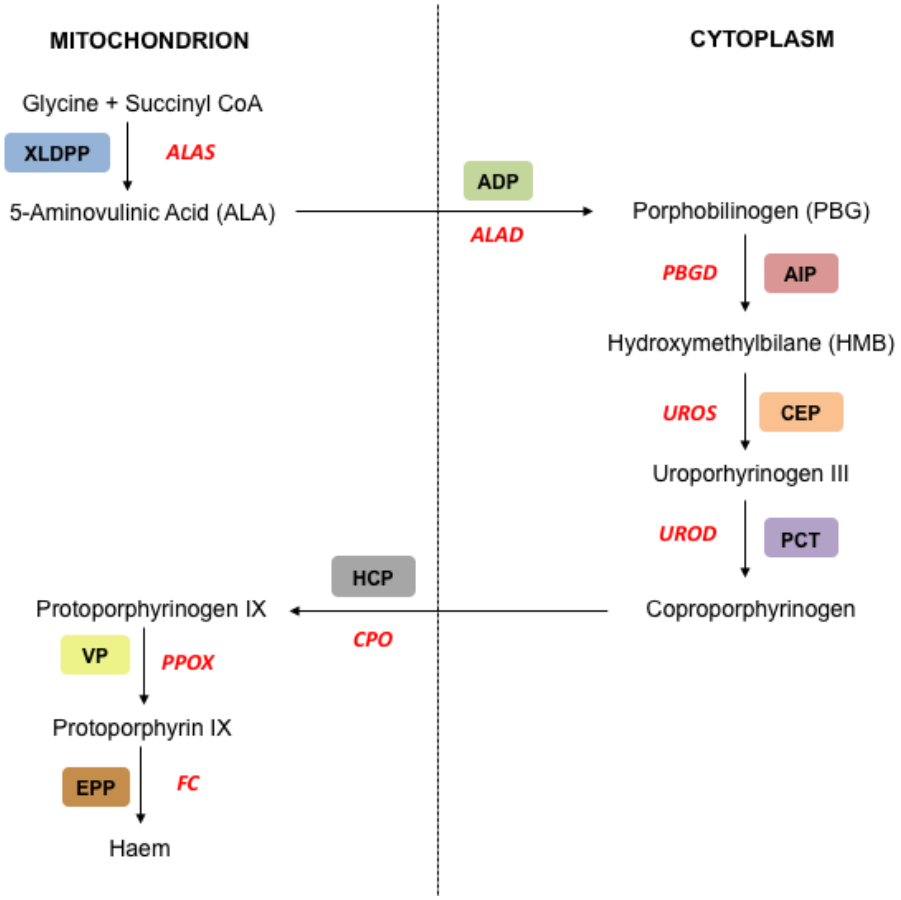


Figure 1.3 A representation of porphyria subtypes and the associated enzyme defects [9]. The enzyme involved in each step is denoted in red. The enzymes are described as follows: (ALAS)-aminovulenic acid synthase; (ALAD) – aminovulenic acid dehydratase; (PBGD) – porphobilinogen deaminase; (UROS) – uroporphyrinogen III synthase; (UROD); uroporphyrinogen decarboxylase; (CPO) – coproporphyrin oxidase; (PPOX) – protoporphyrin oxidase; (FC) – ferrochelatase.

Diagnosis of the Porphyrias

Porphyria is often misdiagnosed in patients, usually for two reasons. Firstly, the disease itself is rare, presenting with a low penetrance (patients presenting with symptoms), and therefore a differential diagnosis might overlook the porphyria symptoms for another, more commonly occurring disease [22,23]. Secondly, the phenotype of patients is not always consistent, and clinical manifestations vary from patient to patient [38]. Symptoms are rarely overtly present in patients with porphyria, and are normally only seen after puberty. Also, a great proportion of patients remain asymptomatic for their entire life [39,40]. As mentioned, each porphyria has a characteristic biochemical porphyrin excretion pattern, depending on which enzyme is defective. Biochemical analysis of urine and or stool samples may provide an indication of porphyria. In acute porphyrias, the levels of PBG are measured in the urine, followed by subsequent urine and or stool analysis in order to determine which specific acute porphyria is

suspected [24,32]. Following a biochemical analysis, DNA may be obtained from patient blood samples for genomic (DNA) screening. This is also very useful for when gene mutation(s) have already been identified in a specific porphyric family, and thus one may usefully employ gene screening as the first line method of choice in establishing a porphyria diagnosis [41,42]. However, the other side of this double-edged sword is, that even when differential diagnoses of suspected cases are made, biochemical analysis from the blood, urine and/or stool often show inconclusive results [43–45]. This illustrates the need for molecular analysis of this rare, yet complex, disease.

The Need for Molecular Analysis

Molecular analysis may be done in two ways. Firstly, as mentioned, by performing genomic (DNA) screening of patients suspected to suffer from porphyria. Here, the biochemical analysis, although perhaps inconclusive as a single test, may serve to aid the genomic screening by indicating the levels of the elevated porphyrin precursors (ALA and PBG) and other intermediates, in order to determine a starting point for which gene within the pathway to screen for mutation(s) [46].

The second method of molecular analysis is at the protein (enzyme) level. Defects within one enzyme lead to the phenotype of a particular porphyria, however, different mutations within the gene, result in different expression, conformations and ultimately efficacy of the enzyme. For example, a simple “T” nucleotide insertion within an exon may result in the generation of a premature stop codon. Consequently, the gene is only partially expressed, resulting in a truncated, non-functional enzyme [47]. In another example, one mutation might affect the enzyme’s kinetic behaviour and affect the ability to produce a product from its substrate (i.e. the enzyme’s specific activity and/or catalytic efficiency). Another different point mutation, located near the protein surface may expose the enzyme’s hydrophobic areas, thus resulting in conformational defects, again with impaired functionality.

Analysis of the protein structure leads to a better understanding of the enzyme mechanism, and thus the disease itself, which potentially allows for *‘bridging of the gap’* between inconclusive biochemical analyses, genomic mutation identification and the often misleading clinical phenotypes [48].

Acute Intermittent Porphyria

Overview

Acute intermittent porphyria (AIP) is one of the acute porphyrias that results from a genetic defect within the third enzyme of the haem biosynthesis pathway, porphobilinogen deaminase (PBGD), also referred to as hydroxymethylbilane synthase *(HMBS)* [45]. AIP is the most commonly occurring acute porphyria world-wide, and also the best described in terms of enzymatic (PBGD) function and characterisation [49].

AIP most often exists as a clinically latent disease, in which patients rarely present with overt symptoms of the disease. In fact, a majority of AIP patients (80%) appear to remain asymptomatic throughout their lifetime [50]. Furthermore, as mentioned, clinical expression of the disease is dependent on multiple factors including nutritional, hormonal, as well as environmental factors that could place pressure on the haem synthetic pathway [43]. Presently, as of January 2016, a total of 403 mutations in the in the PBGD enzyme have been submitted to the by The Human Mutation Genome Database (HMGD®) [51].

The *HMBS* Locus

The *HMBS* gene has been studied extensively since the early 1980s. Raich et al. (1986) first cloned the cDNA of erythroid PBGD from human tissue and deduced the protein size as 344 amino acids [52]. Further studies revealed that two isoforms of the enzyme exist, the erythroid specific form, and the ubiquitously expressed form [53,54]. The human *HMBS* gene was mapped to the long arm of chromosome 11 (11q24.1→q24.2), spanning a total of 15 exons and a genomic interval of 10kb [26,55].

Figure 1.4 The structure and organization of the *HMBS* gene on chromosome 11. Exons are shown as black boxes. The labelling **NE** specifies the non-erythroid promoter for the ubiquitous expression of PBGD, with the transcription start site (ATG) located within exon 1, whereas **E** specifies the erythroid promoter with the erythroid isoform transcription start site (ATG) within exon 3 as adapted from reference [56].

The ubiquitous isoform of the enzyme is the larger, having a total of 361 amino acids, whereas the erythroid isoform has a 17 residue shorter N-terminus, is comprised of 344 amino acids, and is only expressed in erythroid specific cells and tissues [26]. As shown in Figure 1.4, the non-erythroid (ubiquitous) form of the PBGD enzyme is transcribed from the transcription start site found in exon 1, whereas the erythroid form has a downstream erythroid specific promoter within intron 2 and transcription is initiated from exon 3 [53,54,56].

The alternative forms of the PBGD enzyme are produced by alternative splicing of the *HMBS* gene (Figure 1.5). In the ubiquitous (non-erythroid) form (ENST00000278715), the second exon is spliced out and the mRNA transcript is comprised of 14 exons, a total length of 1501 bp and encodes the 361 amino acid enzyme [26]. In the erythroid form (ENST00000392841), however, exon 2 is spliced to exon 3 and transcription is only initiated within exon 3, with the erythroid mRNA thus being comprised of 14 exons, a total length of 1411 bp and encoding the 344 residue form of PBGD [56,57].

Porphobilinogen Deaminase

Porphobilinogen deaminase (EC 2.5.1.61; or hydroxymethylbilane synthase) catalyses the hydrolysis and subsequent condensation of four molecules of PBG to form the linear tetrapyrrole hydroxymethylbilane (Figure 1.5) [14].

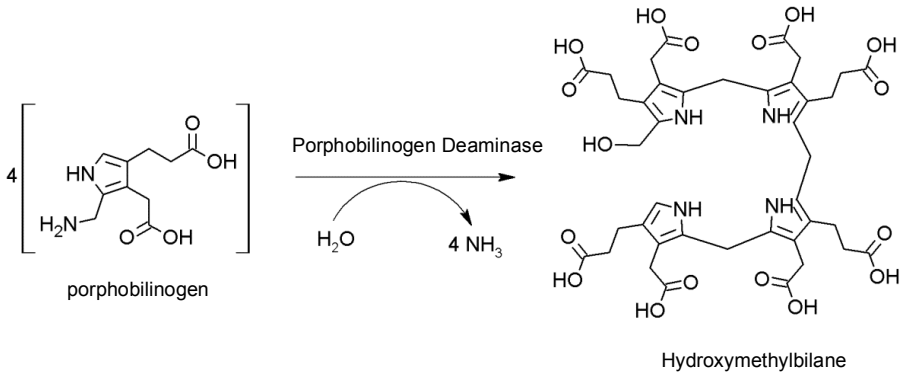


Figure 1.5 The reaction catalysed by Porphobilinogen Deaminase (PBGD) whereby four molecules of porphobilinogen (PBG; ChEBI_17381) are fused together to form the tetrapyrrole hydroxymethylbilane (HMB; CHEMBL273676). Image and reaction generated by ChemSketch, Version 2.0 [5].

One of the characteristic features of the PBGD enzyme is the permanently bound dipyrromethane co-factor, attached to the enzyme via a thioether bond at the cysteine 261 residue within the active site. The dipyrromethane is comprised of the PBG molecules, and is responsible for the stability of the enzyme [58].

Normally, PBGD catalyses the reaction that initiates the formation of the pyrrole ring by hydrolysing, and subsequently condensing four PBG molecules to form the first tetrapyrrole in the pathway, hydroxymethylbilane (HMB) (Figure 1.2). HMB is a linear tetrapyrrole which undergoes cyclization in the next step, to form the prerequisite cyclic porphyrin tetrapyrrole, which is the chemical precursor of haem. In patients with AIP, however, the enzyme usually only has approximately 50% of its activity, placing patients at risk of an acute attack, as failure of the PBGD to convert and utilise the PBG may result in increased levels of PBG in the cell [27,59].

Protein Structure

The enzyme exists as a monomer with the non-erythroid monomeric enzyme form having a molecular weight of between 34 and 44 kDa, depending on the species. For the human enzyme, the molecular size is the 44 kDa for the ubiquitous (or housekeeping isoform), and 42 kDa for the erythroid specific isoform [60]. The PBGD monomer is comprised of three distinct domains as shown in Figure 1.6. Several PBGD crystal structures from a variety of species have been solved at high resolution (Table 1.1).

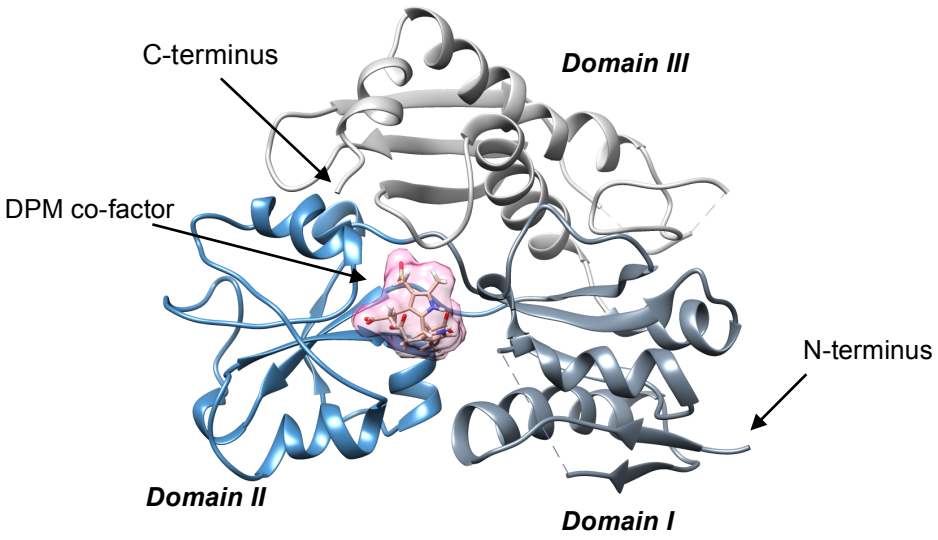


Figure 1.6 The PBGD (EC 2.5.1.61) monomer with its three domains and the dipyrromethane (DPM) co-factor within the enzyme active site. (PDB accession number 3EQ1). Image Generated by UCSF Chimera package [61].

Table 1.1 A summary of the crystal structures available for the PBGD enzyme on the Protein Database (PDB). # Highlighted residues indicated in brackets are the representative residue number in the human enzyme.

Year	Authors	Species	PDB	Å	Detail	Highlighted Residue(s)
1992	Louie et al.	*E. coli*	1PDA	1.76	First reported structure of PBGD, illustrating the 3 domains, the hydrophobic core, hydrogen bonding and crucial interaction of the DPM co-factor [62].	
1998	Helliwell et al.	*E. coli*	1YPN	2.40	Time-resolved structures of hydroxymethylbilane synthase, as it is loaded with substrate in the crystal determined by Laue diffraction [63].	K59Q (K57Q)#
1999	Nieh et al.	*E. coli*	2YPN	2.30	Accurate and highly complete synchrotron protein crystal Laue diffraction data using the ESRF CCD and the Daresbury Laue software [64].	
1999	Hädener et al.	*E. coli*	1AH5	2.40	The crystal structure of the selenomethionine-labelled enzyme ([SeMet]*HMBS*) from *Escherichia coli* solved by means of multi-wavelength anomalous dispersion (MAD) experimental method [65].	
2003	Helliwell et al.	*E. coli*	1GTK	1.66	Time-resolved and static-ensemble structural chemistry of hydroxymethylbilane synthase [66].	
2009	Gill et al.	*Homo Sapiens*	3EQ1	2.80	The structural and functional implications of the R167Q mutation is discussed [60].	R167Q
2009	Song et al.	*Homo Sapiens*	3ECR	2.18	Structural insights into the mechanism of action of PBGD at the molecular level [67].	S96; D99; H120; L238 C261;
2013	Roberts et al.	*Aradopsis thaliana*	4HTG	1.45	Proposed modelling of the intermediate of the elongation process. [68,69].	D95 (D99)#; C254 (C261)#
2014	Azim et al.	*Bacillus megaterium*	4MLV 4MLQ	1.60 1.46	Illustrates how the DPM co-factor becomes progressively oxidised to the dipyrromethene and dipyrromethenone forms, rendering the enzyme inactive [70,71].	

The three domains of the PBGD enzyme are all approximately the same size (Figure 1.6) [70]. The N-terminal domain (domain I) is made up of 3 anti-parallel and 2 parallel beta sheets surrounded by 3 helices. In addition, it consists of numerous loops that form a cap-like structure which covers the active site [67]. Domains I and II are structurally related, and the active site is situated in the cleft between them. The general architecture of domains I and II have been suggested to resemble another protein family known as the periplasmic binding proteins [72]. Such proteins adopt 'open' and 'closed' conformations, in response to ligand binding. This might be a possible explanation for how the enzyme is able to accommodate the growing pyrrole chain within the active site cleft, release a tetrapyrrole product and revert to a 'closed' conformation for the next cycle. Hence, Song et al. (2009) suggested that for the enzyme to be able to incorporate the growing pyrrole chain, there has to be a mechanism involving open and closed conformations. It was predicted that there are three integral residues that allow for the hinge amongst the three domains. These were His 120, Ser 96 and Leu 238 [67]. Roberts et al. (2013) corroborated the hinging mechanism suggesting that the lack of inter-domain connections between domain I and II highlights the possibility of their independent movement. It is important to bear in mind that the largest number of inter-domain contacts are mediated by the DPM co-factor, and so movement of the co-factor may result in movement of the domains.

An interest and focus on the structure of PBGD has continued since the first early considerations of its structure in 1992, until the present. As highlighted in Table 1.1, the structure published in the *Aradopsis thaliana* was the first to be generated from a higher order species of plants. Based on the near atomic resolution of their PBGD crystal, a loop that covers the active site was observed however, in all previous structures, this could not be resolved from diffraction data [68]. Roberts et al. (2013) also focused on the polymerization of the pyrrole rings by the enzyme and its covalently bonded co-factor, and suggested a new thought on an old hypothesis. Louie et al. (1992) suggested two possible theories of the enzyme mechanism based on the *Escherichia coli (E.coli)* crystal structure (which is discussed below) [62,72]. An alternative opinion was discussed by Azim et al. (2014), who published a crystal structure for PBGD in the bacterium *Bacillus megaterium*. The focus of their paper was to report on the oxidation states of the DPM co-factor and to determine 1) how it occurs and 2) what effects and/or consequences it may have (discussed in detail below).

The Active Site

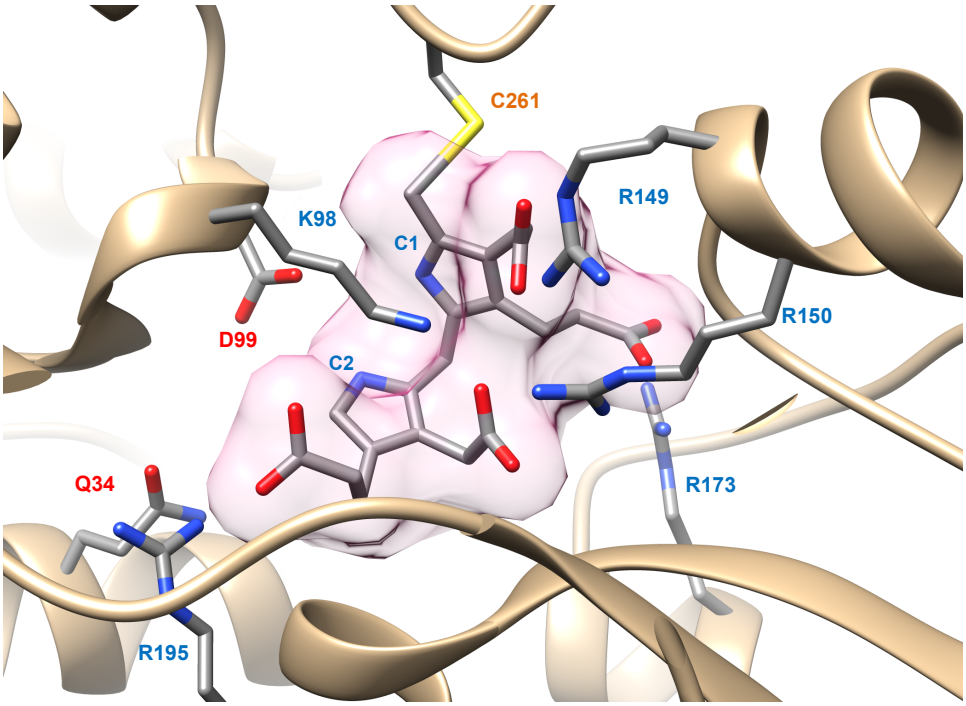


Figure 1.7 An illustration of the architecture within the PBGD enzyme's active site, generated by UCSF Chimera package from the minimised crystal structure 3EQ1 [60,61]. Highlighted residues indicate those relevant in the DPM co-factor binding.

The active site of the PBGD enzyme is very sophisticated, allowing for the permanent addition of the DPM co-factor as well as being able to adapt in order to accommodate the growing tetrapyrrole chain, [66,73]. Several key residues within the active site play an important part in maintaining the active site organization, forming of polar interactions amongst each other, and being essential for enzyme catalysis, as illustrated in Figure 1.7 [67,74]. The majority of the active site is populated with arginine residues, thereby resulting in a largely positively niche, providing an interface for ionic interactions between different residues and chains, the DPM co-factor, as well as incoming substrate (PBG) molecules [74]. Indeed, most reported mutations, if not all, of these residues within close proximity of the active site are known to be associated with the occurrence of AIP in humans [75,76]. The only covalent interactions within the active site are the cysteine 261 (C261) bonded to the DPM co-factor, and the bond formation of the pyrrole chain on each successive terminal ring. The remainder of the interactions are mediated by van der Waal's forces, salt bridges and hydrogen bonds [77]. This is a clear indication why mutations within any of these residues ultimately result in disease as any changes to the network of interactions within this active site cleft is bound to disrupt the required interactions, or alternatively generate different interactions, which may disrupt the enzyme's normal function [60].

The DPM Co-Factor

The DPM co-factor has been extensively studied, since its first discovery in the late 1980s, when Warren and Jordan (1987) provided preliminary evidence for the existence of this dipyrromethane within the active site cleft [78]. In the same year, Hart et al. (1987) demonstrated that the novel group within the enzyme active site (which was initially termed group X), was indeed dipyrromethane [79]. Prior to this discovery, there were several theories as to how the enzyme is actually able to polymerize subsequent PBG molecules.

Previous studies suggested that PBG itself was bound to either lysine or cysteine, but the question remained as to how the enzyme was able to form stable intermediates during this substrate polymerization [78]. With the evidence presented by Hart et al. (1987) it was then clear that the tightly bound pyrrole was a permanent feature of the PBGD enzyme [79,80]. Of course, subsequent questions were raised regarding the origin of this novel binding group as well as the binding of the co-factor to the enzyme. In 1988, two groups published papers at roughly the same time, both of which provided evidence that the co-factor was bound to Cys 261 within domain III [81,82]. Figure 1.8 shows an illustration of the DPM co-factor structure as well as a model from the human enzyme crystal depicting the organization and orientation within the active site.

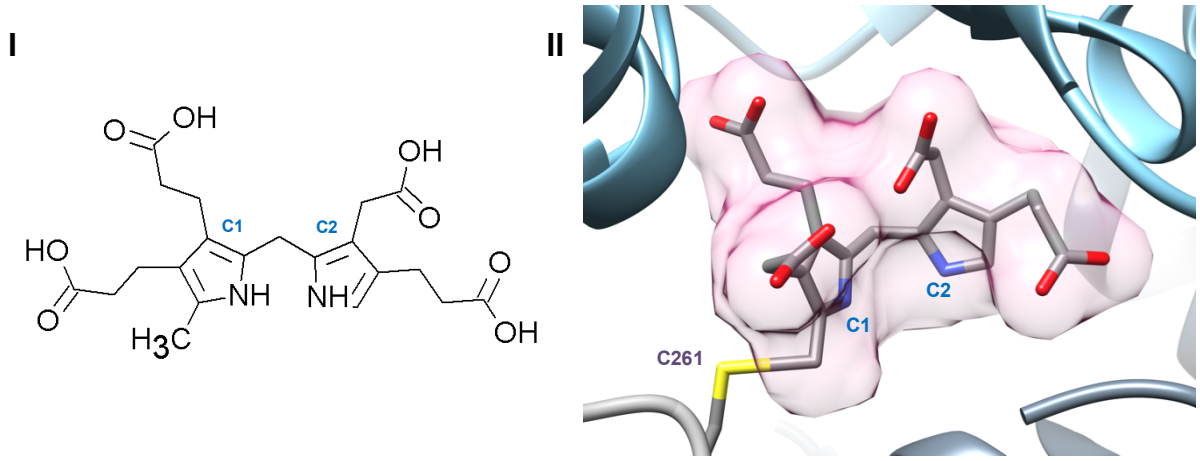


Figure 1.8 An illustration of the DPM co-factor and its position in the PBGD enzyme. **(I)** Shows the chemical structure of the dipyrromethane generated from the ChEBI file for DPM (ChEBI_42121) in Chemsketch Version 2.0 for Windows [5]. **(II)** Shows the orientation and architecture within the PBGD enzyme's active site, generated by UCSF Chimera package from the minimised crystal structure 3EQ1 [60,61].

With regards to the origin of the dipyrromethane co-factor, there are two theories. The first is that the co-factor is generated from the PBG substrate molecules in a first cycle, followed by subsequent polymerization, resulting in a hexapyrrole, with the HMB tetrapyrrole cleaved off and released as a product, whilst the remaining dipyrrole remains attached to the enzyme as the DPM co-factor. Although this has evidence to support this idea, the results also suggest

that process is quite slow [83]. The second theory was put forth by Shoolingan-Jordan et al. (1996) in which they illustrated that the assembly of the DPM co-factor is done by the reaction of the preuorporphyrinogen product with the *'apo'* PBGD enzyme [84]. In their work, they also argue that co-factor assembly is possible with the addition of only the PBG substrate molecules, however, the reaction proceeds very slowly, and the recovery of holoenzyme was very low (between 10 – 40% of normal enzyme activity). Their proposed alternative theory was that the product of the enzyme, HMB, which is also referred to as preuroporphyrinogen, is the ideal precursor for the formation of the DPM co-factor *in vivo* [68,84]. In other words, the preuroporphyrinogen, which is comprised of four tetrapyrroles, acts as the DPM co-factor precursor, attaching to the cysteine 261, thereby illustrating the ES_2 complex, followed by the addition of two further substrate (PBG) molecules to generate ES_4. This is then followed by the release of the ES_4 product, with the dipyrrole remaining permanently attached to the enzyme. The principle is illustrated below in Figure 1.9.

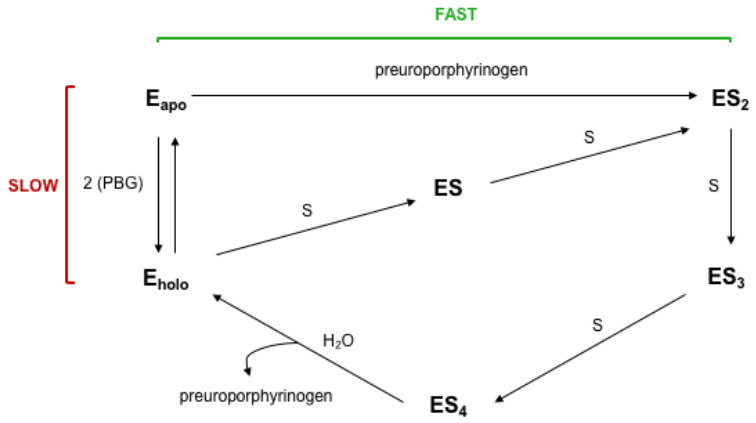


Figure 1.9 The proposed mechanisms by which the PBGD enzyme may acquire the DPM co-factor, as adapted from [84]. Legends as follows: E(apo) – PBGD with no bound co-factor; E(holo) – PBGD with bound co-factor; ES – PBGD with one PBG: ES_2 – PBGD with 2(PBG); ES_3 – PBGD with 3(PBG); ES_4 – PBGD with 4(PBG).

With regards to the oxidation of the co-factor, the characteristic pink colour of the DPM co-factor was initially suggested to be pyrrole intermediates. There was also speculation that it could be the concentration of endogenous PBG in protein samples that leads to the pink chromophore. However, the recent X-ray crystallographic studies have investigated the oxidation states of the co-factor [70,71]. It is known that purified PBGD enzyme loses the bright pink colour and becomes progressively orange over a period of time, resulting in a pale orange-yellow solution [69].

This colour change is accredited to oxidation of the co-factor (Figure 1.10). Indeed it was shown that the α-position of the C2 ring is sensitive to oxidation, and the oxidized dipyrromethane pyrrole, or rather dipyrromethenone, is catalytically inactive, as the carbonyl

oxygen at this point blocks the addition of new pyrroles. In addition, this process is irreversible, and the addition of reducing agents to prevent and/or terminate the oxidation inevitably failed.

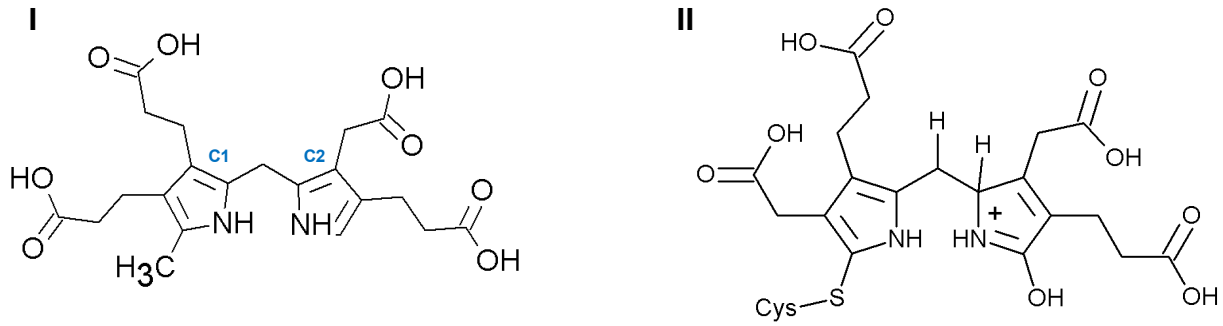


Figure 1.10 DPM co-factor oxidation **(I)** Shows the chemical structure of the dipyrromethane (as before) generated from the molfile for DPM (ChEBI_42121) in ChemSketch Version 2.0 for Windows [5]. **(II)** The proposed bound-oxidised cofactor, dipyrromethenone, illustrating the carbonyl oxygen at the C2 ring.

Highlighted Residues for Co-Factor Interaction

The most important interactions have been suggested to be residues containing pyrrole nitrogen(s) (N) that can form a hydrogen bond with the carbonyl oxygen atoms from the co-factor [68]. For example, the R131L and R132L mutations have been shown to affect the binding of the DPM co-factor. This is as a result of the change in charge from the positive arginine (R) to the neutral leucine (L). Therefore the ionic interaction between the negative carbonyl groups of the co-factor and the positive amine groups in the arginine side chain is lost, rendering the enzyme inactive [77]. Indeed, there are several other residues, including the K98, D99, C261, R150, R173 and the R195, which have been suggested to be involved with the DPM co-factor interaction [67]. Several residues have also shown to be important for both enzyme catalysis as well as co-factor orientation. R149 as well as D99 mutations indicate a loss of co-factor interaction, although they are considered key catalytic residues [85].

PBGD Enzyme Mechanism

PBGD is able to form stable intermediates during substrate polymerization, also referred to as chain elongation [78]. Figure 1.11 provides a simplified illustration of the process, whereby each successive PBG molecule is added to the enzyme-co-factor complex, in a stepwise manner [48,82]. The complexes are generally referred to as ES, with one PBG molecule added, ES_2 depicting two PBG molecules, similarly with ES_3 and ES_4. As illustrated in Figure 1.11, the PBG molecules are added to the enzyme in a head-to-tail fashion, with the free alpha carbon on the C2 ring of the DPM co-factor acting the source for nucleophilic attack in order to form the bond [74].

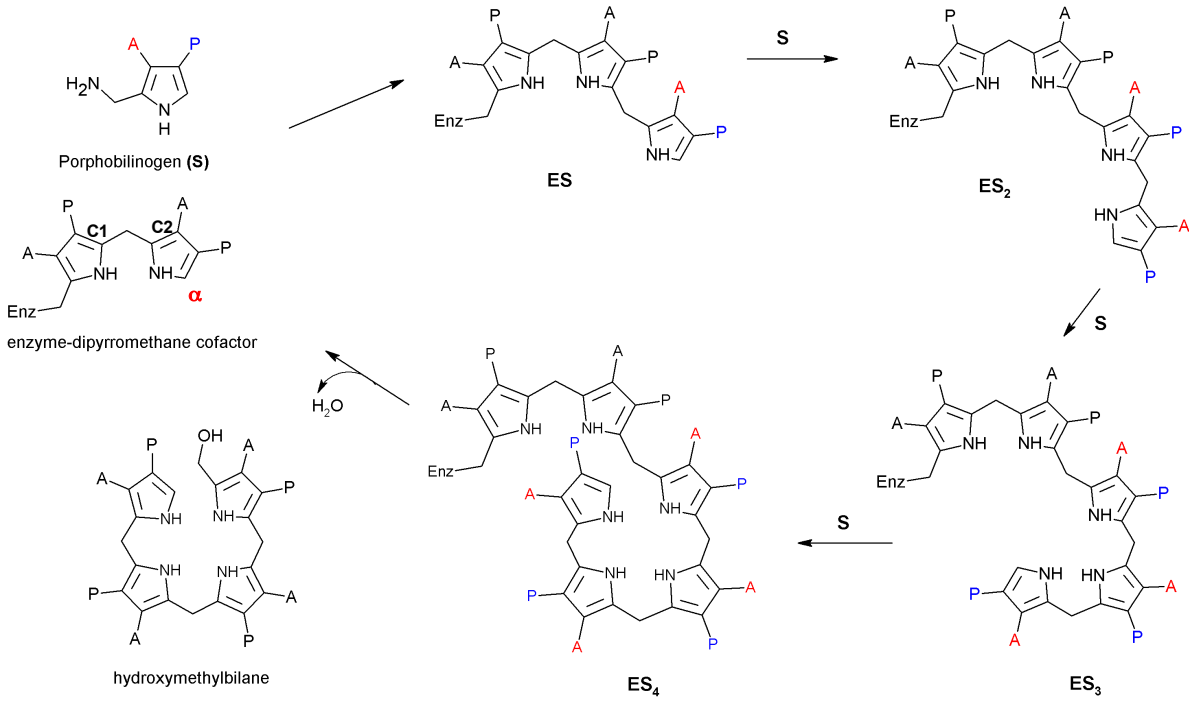


Figure 1.11 An illustration of the PBGD enzyme mechanism, illustrating the polymerization of the PBG substrate in each successive step, as adapted from [74]. Image was generated in ChemSketch Version 2.0 with the use of the Molfiles for PBG and the DPM co-factor, ChEBI_17381 and ChEBI_42121 respectively. The lettering as follows: **A**: acetate (-CH_2CO_2H) and **B**: propionate (-$CH_2CH_2CO_2H$).

There has been much debate regarding the mechanism for this pyrrole chain elongation. As mentioned earlier, the first two proposals were put forward by Louie et al. (1996) following the first successful crystal structure of PBGD in *E.coli*. Their first hypothesis was what is referred to as the *"active site model"*, in which the growing pyrrole chain is accommodated in the active site cleft, and the three domains adjust themselves accordingly in order to facilitate the binding of each subsequent pyrrole, whereas the second hypothesis, referred to as the "*moving chain model"*, suggests that the developing chain is progressively "pulled" across the catalytic site, thereby placing the terminal binding rings within the substrate binding site [72].

Roberts et al. (2013) built on this model, with their proposed "novel mechanism" following observations from their crystal structure [68]. Their mechanism for polymerization included that the co-factor was in its oxidised form in order to incorporate the growing pyrrole chain within the active site cleft. Their argument was that the oxidised conformation of the DPM co-factor acquires a different structural orientation when compared to the reduced form, and this oxidised form allows for additional movement and generation of space in the active site cavity. This concept on its own is true, in that, based on their modelling and predictions, the active site does open up more once the co-factor is oxidised. This was also suggested by Louie et al. (1996), illustrating two conformations within the active site, one with reduced co-factor, and the other with the oxidised form [72]. However, they did not consider what effect the oxidation

state of the co-factor would have on the activity and catalytic mechanism of the enzyme itself. This clearly illustrates one of the limitations of crystallography. Although it provides insight into the structure and architecture of the enzyme and its parts, it remains a static picture of a dynamic system. In addition, the crystallisation conditions are not necessarily equivalent to *in vivo* conditions. This was proven by Azim et al. (2014) who showed that the co-factor is sensitive to oxidation, and that this oxidation renders the enzyme inactive [70]. Furthermore, it is irreversible, and no attempt to 'salvage' the enzyme with re-addition of reducing agents or attempts to prevent oxidation proved viable. This answered the questions raised by Roberts et al. (2013), as their theory also includes the hypothesis that the co-factor may shift between reduced and oxidised states during pyrrole elongation [69].

Bung et al. (2014) published a paper following the suggestions made by Roberts et al. (2013), in which they performed molecular dynamics on the PBGD enzyme, in order to clarify the hypotheses and proposals of the mechanism based on crystal structures [77]. Molecular dynamics involves the computational modelling of the movement of atoms and bonds within an enzyme in order to predict how ligands may interact with a protein [86,87]. Bung et al. (2014) classified the mechanism of the enzyme as operating in three stages. Firstly, chain elongation, forming the tetrapyrrole product. Secondly, the exit of the HMB product from the active site cavity, and finally the relaxation of the enzyme following product release.

With regards to the chain elongation, it was found that the active site loop (residues 42 – 60 in the *E.coli* PBGD) as well as domain II, adjust their conformations to widening the active site cleft, in order to accommodate for the growing pyrrole chain (referred to as the 'open' conformation). Interestingly, they also found that the space occupancy for the entire tetrapyrrole product was adjusted for by the active site loop, and minimal structural changes were observed for the protein as a whole. In addition, when comparing the secondary structure of the domains from the first stage, where no substrate was bound (E), to the final stage that included the bound tetrapyrrole product (ES_4), noticeable changes were observed, specifically between domains I and II. This suggested that there was a widening of the gap between the two domains as the chain elongation progresses.

Highlighted Residues for Enzyme Function

Studies suggest that mutations in the active site arginine residues have an effect on enzyme mechanism [74,88]. The mutations, R11H, R149H, R155H, R176H and R232H all cause disruption in the formation of the HMB product. The alteration of all of these residues to histidine (H) causes steric disruption as rotation of the cyclic bonds in histidine is limited in comparison to bond rotation within arginine (R). It is also possible that the bulky ring from the histidine causes disruption of surrounding residues [74].

In terms of product release/exit, the molecular dynamics and simulations suggest that the HMB product exists between the space of Domain I and II and the active site loop. Finally, structural changes between the last stage following product release and the original form suggested, protein 'relaxation', thus allowing for the active site to re-assume the 'closed' conformation, and prepare for the next cycle [77]. The significance of the molecular modelling is that it provides an analysis of the enzyme's dynamic features [89]. Although the analysis of the PBGD crystal structures have been significant in that it allowed for discovering the residues important for catalysis and for speculation regarding how the structure correlates to the enzyme function, the crystal does not contain the substrate (ES complex). Whereas with the molecular dynamics, it allows for the analysis of each of the stages (as depicted in Figure 1.11) [77].

The R149 residue has been shown to be involved with forming a network of hydrogen bonds with the K98 and D124 within the active site [85]. Thus, mutations of R149 may result not only in failure of catalysis, but disrupt additional residue interactions. Lysine has also proven to be important for the enzyme's catalytic function, as K62 and K59 mutations to glutamine (Q) have been reported to result in disturbances in the enzyme's catalytic efficiency [64,77,90]. Several other residues in the active site vicinity have been shown to be highly conserved in a range of different species, pointing to their potential role in enzyme function and/or confirmation. The D99 residue has been highlighted in several cases as being the major catalytic residue of the active site, and is shown to be involved with each of the polymerization steps during chain elongation [67,85].

Highlighted Residues for the Exit Mechanism

It has been suggested that the movement of the active site loop that falls between domains I and II guided by strong interactions of neighbouring residues, plays a key role in the catalysis, including the exit of the HMB product. In addition, residues R26, Q34, K98 and R196 have been suggested to aid the movement of the product towards the cleft of domains I and II [77]. The D99 residue has been suggested to play a role in the final hydrolysis reaction, whereby the HMB product is cleaved from the co-factor [68].

Lysine 98

The Lysine 98 (or K98) is a highly conserved residue within the PBGD enzyme active site, as illustrated below in Figure 1.12. It has also been shown to be conserved in several other species [67,70]. In addition, it has been suggested throughout the literature as being a key residue for several stages of the enzyme function [67,68,77]. However, in all of these cases the significance of the K98 residue is based on speculation, the vicinity to the active site and also the proximity to nearby residues D99 and R195. No further detail as to why this residue could be important is suggested. Generally there is a failure to mention how exactly this residue interacts further with neighbouring residues in terms of structural orientation, and only speculation in terms of the contribution of this residue to the DPM co-factor interaction [42,45,64].

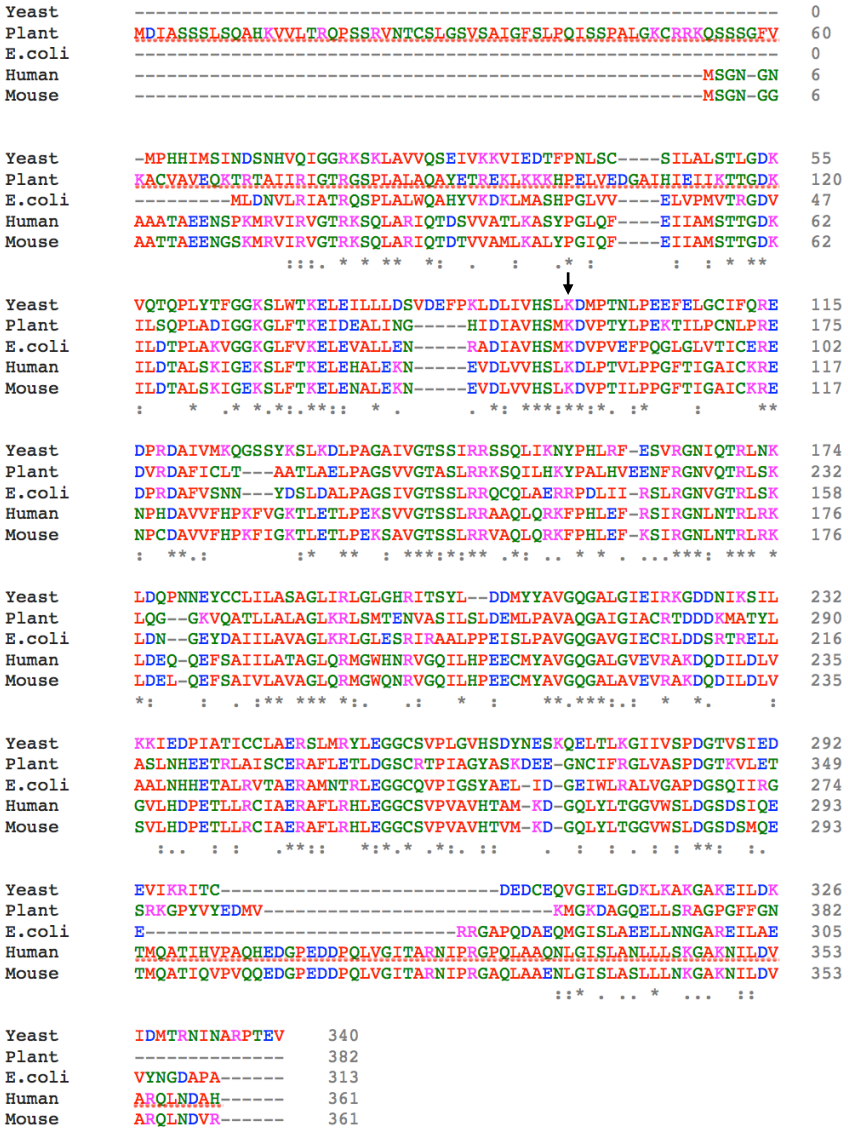


Figure 1.12 The multiple sequence alignment of PBGD in different species. Conserved residues are indicated by the (*). The K98 residue is indicated by the arrow.

Study Rationale

As there is a clear indication that the K98 residue plays a key role in several stages of the enzyme mechanism, we developed an interest in furthering our knowledge around this residue. This may provide insight into our understanding of the role this residue specifically plays within the enzyme active site. Like other studies we believed that we could gain insight by mutating this residue and analysing the effect(s) of such. A previous M.Med student from our laboratory discovered a K98E mutation in a South African AIP family (personal communication, P Fortgens, AV Corrigall and PN Meissner). Furthermore, a K98R mutation has been documented in a patient in 1991. However, no further analysis on this mutation was performed [91]. From our work in 2013, we have preliminary evidence that suggests that mutations in the K98 residue cause disruptions to the enzyme's kinetic behaviour. Thus, apart from an interest in better understanding PBGD at a basic level, our interest in the K98 residue was born out of the clinical realm, and an on-going interest in porphyria disease-causing mutations.

Our aim therefore, and the purpose of this study, was to elaborate on current information regarding the characteristics of the PBGD enzyme, by investigating how mutations within the active site affect the enzyme mechanism as well as conformation. Indeed, further mutations at residues like K98, as well as other structurally identified/important sites, could improve our understanding and may allow us to further elucidate the complex mechanism of the PBGD enzyme.

In chapter 2 of this dissertation, I describe our methodology and approaches to the kinetic and structural characterisation. Included are explanations of the specific protocols developed for the study, as well as optimization strategies that were required for expression and purification of the recombinant enzymes.

Chapter 3 presents and describes the results of this study. In several cases data is represented in both a graphical and table format. This was to allow not only for quantitative presentation, but also a more visual and qualitative overview, when comparing the wild type PBGD to the recombinant enzymes.

Finally, in Chapter 4, I provide an in-depth discussion of each section of results, and an integrated discussion and depiction of each of the different mutations. We conclude and provide an overview of future directions for this work.

Specific Research Question

Do PBGD Lysine 98 (K98) mutants affect enzyme function and if so, what are the likely underlying molecular mechanisms and explanations of such dysfunction?

Aims

1. **To engineer a series of K98 mutant PBGD enzymes by site directed mutagenesis.**
   Objectives:
   - Design primers that will incorporate the desired mutation within the *HMBS* gene.
   - Perform a site directed mutagenesis PCR reaction followed by a *DpnI* restriction endonuclease digest in order to eliminate any template DNA.
   - Transform mutated sequenced DNA into JM109 supercompetent *E. coli* cells.
   - Plate transformed cells onto prepared Agar plates and select colonies for inoculation.
   - Incubate selected colonies in order to extract DNA for sequencing.
   - Verify which selected colonies have incorporated the desired mutation by means of direct sequencing.
2. **To perform an activity and kinetic characterisation on the engineered, expressed and purified PBGD K98 mutants.**
   Objectives:
   - Inoculate large scaled batched of media in order to express and purify the selected mutations.
   - Establish purification protocol for newly generated mutant. Purify each of the mutants in parallel with wild type.
   - Set-up and analyse the kinetic profile of each of the mutants and correlate to the wild type protein.
3. **To perform a native polyacrylamide gel electrophoresis (PAGE) analysis of purified PBGD K98 mutants in order to observe enzyme-substrate (ES) complexes.**
   Objectives:
   - Purify wild type and each of the K98 mutants.
   - Establish a protocol that allows for successful visualisation of wild type ES bands on native PAGE gels.
   - Perform a native PAGE and verify whether any of the mutations result in changes in the electrophoretic behaviour of the enzyme complexes.

4. **To analyse the effect of mutating K98 in terms of the enzyme's DPM co-factor.**

   Objectives

   - Set up the spectrophometric analysis to observe co-factor binding.
   - Correlate mutation co-factor binding to wild type enzyme.

5. **To attempt to remove the bound dipyrromethane co-factor from wild type enzyme.**

   Objectives:

   - Set up and perform an assay to remove the bound dipyrromethane from the native wild type enzyme.
   - Confirm the successful removal of the enzyme's co-factor (as in 4. above).
   - Set up and establish a reconstitution assay, to test the enzyme's ability to accept co-factor versus the PBG substrate.

6. **To gain structural insights into wild-type PBGD and K98 mutants using conformational and thermodynamic characterisation assays.**

   Objectives:

   - Purify each of the respective mutants in parallel with wild type enzyme.
   - Set up and establish a protocol for measuring the Far- and near- UV circular dichroism spectra for each of the mutants.
   - Determine the secondary structure estimations for each of the mutants by means of several structural algorithms (CDPro, CDNN, DSSP and STRIDE).
   - Set up and perform a thermal assay for each of the samples.
   - Fit data to a two-state unfolding plot and determine the melting temperature for each of the mutants and compare to wild type.
   - Determine the conformational stability of each mutant by measuring the fraction of unfolded protein of each mutant and comparing it to wild type.

7. **To model the K98 mutant proteins from a structural perspective and attempt to reconcile the structural analysis with the kinetic and thermodynamic data.**

   Objectives:

   - To generate a collective enzyme structure from available PDB crystal structure files, in order to have an accurate depiction of enzyme structure (including missing loops).
   - Minimizing the superimposed-aligned structure of the enzyme in order to have the lowest energy calculations.
   - Model each of the mutations to the prepared crystal structure of the PBGD enzyme and observe possible alterations and/or explanations that correlate to the kinetic and thermodynamic data.

Chapter 2

Materials and Methods

Introduction

This chapter details the materials and methods employed for the experimental work. In general, this takes the form of presenting the principles on which the method is based followed by the procedure used in the lab. Where further detail is required the reader is referred to an appropriate appendix.

Site-Directed mutagenesis

Principle

QuickChange® employs supercoiled, double-stranded plasmid DNA and two designed, synthetic primers containing the desired mutation in order to generate a mutation by means of the highly effective DNA polymerase *PfuTurbo*. Figure 2.1 below is an illustration of the mutagenesis principle, adapted from the Stratagene *QuickChange*® user manual.

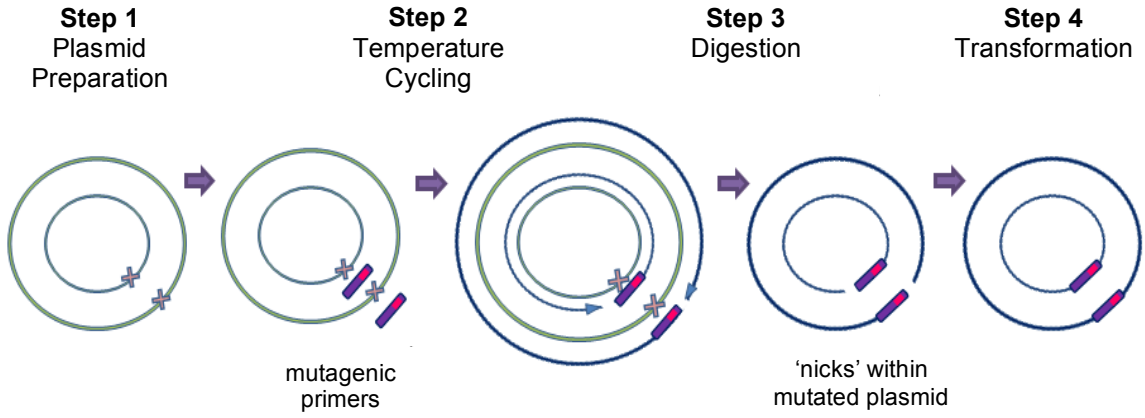


Figure 2.1 An illustration of the principle of *QuickChange*® Site Directed Mutagenesis. **Step 1**) The plasmid containing the gene of interest and the target site (**X**) for the mutation. **Step 2)** Denaturing of the plasmid to allow for annealing of the mutagenic primers ( ) containing the desired mutation. Following denaturation, allow the *PfuTurbo* DNA polymerase to extend and incorporate mutagenic primers. **Step 3)** Digest the wild type parental template plasmid with *DpnI*. **Step 4)** Transform the circular, nicked dsDNA into JM109 supercompetent cells. These cells will then repair the nicks in the mutated plasmids.

Procedure

Wild type *HMBS* DNA was extracted from an overnight growth of an *E.coli* plasmid culture using the Wizard® Plus SV Minipreps DNA Purification kit (Promega Corporation, Madison, USA) (detail given in Appendix 5). Mutagenic primers (also referred to as oligonucleotides) were designed complementary to the opposite strands of the plasmid, and extended, by the action of *PfuTurbo* during the temperature cycling. The products were then subjected to *DpnI* restriction endonuclease (RE) digestion, which cleaves and digests the parental template strand. The *DpnI* recognises the methylated sequence (5'-Gm_6ATC-3') found within DNA isolated from *E.coli*. This allows for the selection of only mutated plasmids. The mutated plasmid DNA was then transformed into supercompetent cells (JM109). The details of the

protocol and precise methodology for this study are outlined in Appendix 9. Post transformation and plating, eight colonies were selected after an overnight incubation at 37ºC and inoculated into 6 ml LB media (Appendix 3) containing ampicillin. Samples were cultured in a shaking incubator at 37ºC overnight. The selected samples were also re-plated, prior to LB inoculation, on another sterile agar plate, and incubated at 37ºC overnight. Thereafter 30% glycerol stock solutions were made and stored at –80ºC (Appendix 4).

Screening Clones

In order to screen the entire *HMBS* gene (Appendix 1), we designed two sets of primers (forward and reverse) using Primer Designer (Software Packaging version 2, Scientific and Education Software). The primers were designed to stretch from the vector into the *HMBS* gene insert and also from the insert to the vector (detailed in Appendix 2). This was done in order to enable screening for the desired mutation and also to ensure that no additional mutation(s) had inadvertently been introduced. The analysis software from Integrated DNA Technologies (IDT), Oligo Analyzer, was used to analyse the properties of the designed primers as follows: percentage GC content, melting temperature (T_m), ΔG, self-dimer, and hetero-dimer (Table 2.1). Primers were manufactured by Integrated DNA Technologies (IDT). Further details are outlined in Appendix 7.

Table 2.1 Outline of the parameters set for the primer design.

Parameter	Minimum	Maximum	Optimum
GC (%)	50	60	55
T_m (ºC)	70	80	74

Gradient PCR was employed using wild type DNA to determine the optimum annealing temperatures and ensure specificity of the *HMBS* fragments (as outline in Appendix 2). The PCR was performed as described in Appendix 10. The plasmid DNA from overnight cultures of the engineered mutants was extracted using the Wizard® Plus SV Miniprep DNA Purification system (Appendix 5) and quantified (Appendix 6). The amplified PCR products were then loaded on a 6% non-denaturing PAGE gel (Appendix 11) in order to check the quality of the products. The online bioinformatics tool, Webcutter 2.0 (http://rna.lundberg.gu.se/cutter2/), was utilised to ascertain if a restriction site would be created/abolished by the engineering of the mutants. If so, a relevant restriction analysis was performed for easy screening of potential mutants. Only one of the engineered mutations, K98R, created a *Bsm FI* recognition site. Details for the restriction analysis are outlined in Table 2.2.

Table 2.2 *Bsm FI (FaqI)* restriction analysis.

Reagent	Final Concentration	Volume (µl)
PCR Product (non-purified)		5
Buffer (10X)	1X	2
Bsm FI (2U/µl)	1	0.5
d.H_2O		12.5

Following digestion, products were loaded on a 6% non-denaturing PAGE gel (Appendix 11). As no cutting site was introduced/abolished by the engineering of the K98E and K98A mutations, 4 selected clones were sent for direct sequencing. Due to the high success rate of engineering mutations previously experienced in the UCT Porphyria laboratory it was not deemed necessary to design a primer that would create a suitable restriction enzyme cutting site.

Confirmation of Mutated Sequences

PCR products were purified by the Illustra^{TM} GFX^{TM} PCR DNA Purification kit (Appendix 12). DNA Sequencing was done by the Core DNA sequencing facility of the University of Stellenbosch using the BigDye Terminator sequencing kit, Version 3.1 (Applied Biosystems, Brachberg, USA) on either a ABI3130xl or ABI3730xl Automated Genetic Analyzer. Purified PCR samples were sent in 8 ng/µl concentrations and primers at 1.1 pmol/µl. In order to ensure proper screening of the entire fragment, sequencing was performed both in the forward and reverse direction. The mutated sequence was aligned to the known wild type fragment(s) (Appendices 1 & 2).

Expression of wild type and mutant PBGD

Wild type PBGD clones with a 6X-His tag plasmids have been used in our laboratory prior to this study. The PBGD containing pTrc His-A Vector plasmid we utilised was a gift from Professor Harry Dailey (University of Georgia, Athens, Georgia, USA). For PBGD expression, a 1 ml stock (Appendix 4) of PBGD was inoculated in 1L LB media (Appendix 3) containing ampicillin and incubated at 37^{o}C for 22 hours in a shaking incubator (225 rpm). In the case of our low expressing mutants, we inoculated a double-batch, thereby having 2L of inoculated media for K98R. For the K98A, the growth period was increased to 24 hours.

Purification of wild type and mutant PBGD

The TALON™ metal affinity resin protein purification system was used for the purification of both wild type and mutant PBGD.

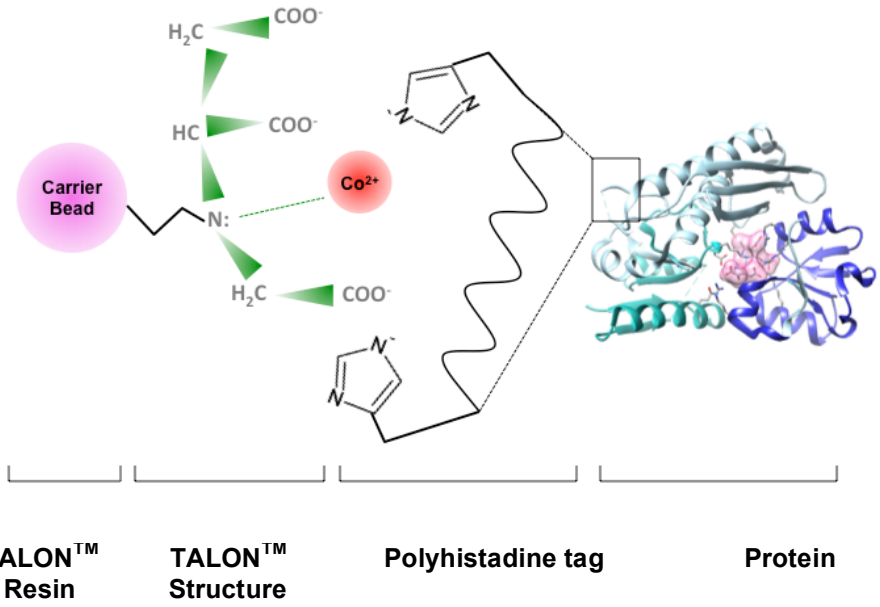


Figure 2.2 An illustration of the principle of the TALON™ metal affinity chromatography. The carrier bead is attached to the talon structure containing the Cobalt metal ion (Co^{2+}). Image is based on the concept from the Clonetech website [http://www.clontech.com][92,93]. Histidine (ChEBI:15971) residue was generated by ChemSketch, Version 2.0 [5].

Principle

This purification system is based on the interaction and degree of affinity of histidine (with a high electron dense region) for Cobalt (Co^{2+}), within the TALON™ resin. The positive oxidation state of Cobalt provides a means of interaction for the electron histidine-rich residues within the 6X-His tag (Figure 2.2). The His-tagged PBGD 'sticks' to the resin, allowing other non-specific proteins to flow through. A wash buffer, with a low concentration of imidazole, is used to remove any non-specifically bound proteins. Finally, an elution buffer with a high concentration of imidazole is used to competitively elute the purified PBGD, as illustrated below in Figure 2.3.

Procedure

Briefly, cells were cultured from a 1L sample of medium, harvested by centrifugation and resuspended in a sonication buffer. Cells were then lysed by sonication and lysate ultra-centrifuged to yield a clear suspension. The TALON™ resin was equilibrated and sample loaded onto the column at a flow rate of 15 ml per hour. Flow through was collected (the void) and pooled. The column was then washed with a minimum of 10ml wash buffer, before eluting

the bound protein with the final (elution) buffer. Phenylmethyldsulfonylfloride (PMSF) was added in a final concentration of 1 µg/ml to all purification buffers directly before addition to column, in order to inhibit serine proteases [94]. Further details of the protocol are outlined in Appendix 13.

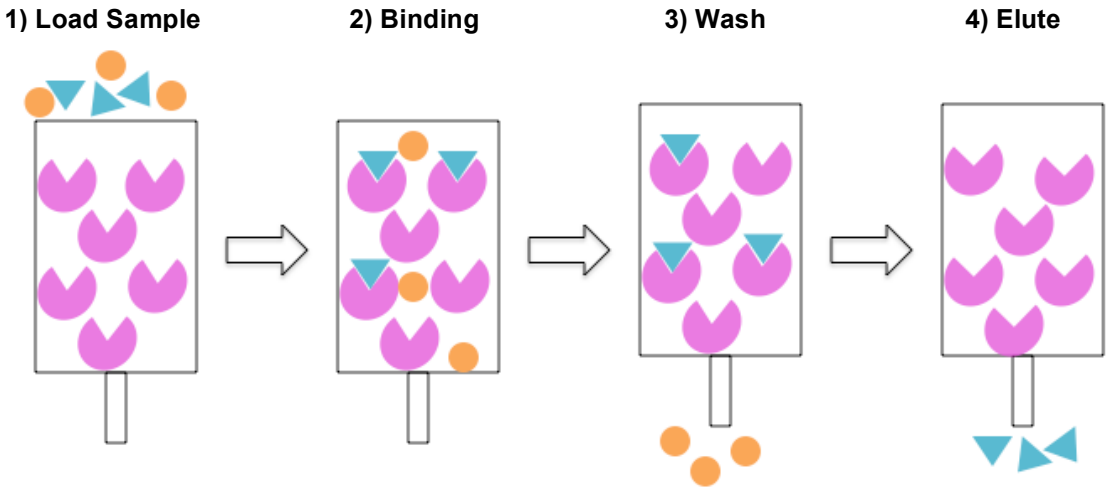


Figure 2.3 A simplified schematic of the affinity chromatography. Steps are as follows: **1)** The sample, containing desired protein amongst various other proteins, is loaded onto column**. 2)** The polyhistidine-tagged protein binds to the resin with high affinity, whereas other proteins flow through resin and column. **3)** Non-specific proteins are washed off, leaving only the proteins binding to the resin with higher affinity than to the wash solution. **4)** Finally, the polyhistidine-tagged protein is eluted from the column with a solution that transcends the affinity of the protein to the resin.

Optimization

Although this affinity chromatography method has been utilised previously in our lab with success, it has been previously reported that occasionally, with engineered mutations, the stability and purity of these proteins are affected. Therefore, several optimization strategies where employed, specifically with the low expressing mutants, K98R and K98A, in order to 1) improve the percentage purity in the eluted protein, without having drastic effects on protein yield, and 2) to possibly improve the stability of protein samples. Each one of these strategies was performed separately, in order to determine which is the most feasible.

Alternative Buffer Composition

The manufacturing company of the resin provide alternatives to the sonication buffer. It is possible to add imidazole to the sonication buffer in a concentration of between 5 – 10 mM. The rationale is that by having a competing element within the sonication buffer, it is easier to prevent non-specific binding, as these by-products usually bind with low affinity to the column. Another alternative is to reduce pH of the wash buffer to 6.8, which is suggested to promote the removal of contaminants [92,93].

Protease Inhibitors

A variation of protease inhibitors was used to test whether this could improve the purification results and prevent protein degradation (as in the case of the unstable mutant proteins). However, both the reducing agent, Dithiothreitol (DTT) and the metalloprotease inhibitor, Ethyl-enediaminetetra-acetic acid (EDTA) often used in these protease inhibitor cocktails, are incompatible with the metal affinity resin. Therefore, we made up a cocktail containing 7.2 mM Benzamidine, 5.0 mM N-Ethylmaleimide (NEM), and 4.2 mM PMSF.

Cold Room

Previous work in our lab with PPOX mutants were performed in the cold room at 4°C, which improved protein purity and also prevented immediate protein precipitation [94].

Concentration & Glycerol

It was also suggested in the Clontech Laboratories manual, as well as in other studies, that lower protein concentration would reduce the degree of protein precipitation. Between 10-20% sterile glycerol was also added to freshly purified mutant protein solution in order to prevent protein precipitation [95].

Assessment of Purity and Size of wild type and mutant PBGD

Following purification, the protein samples were loaded on a 7.5 – 17.5% gradient SDS-PAGE gel (Appendix 14) in order to determine if the protein eluted is of the correct size and to assess purity at each step of the purification. The optimization strategies were also verified using these SDS-PAGE gels.

Protein Quantification

The collected protein samples, including the load, void, wash and purified eluted protein, were quantified using the BioRad® (Bio-Rad® Laboratories Ltd, Johannesburg) micro-assay for protein quantification (Appendix 16). A standard curve was set up using a bovine serum albumin (BSA) sample range of known concentrations, and the concentration of samples extrapolated from the curve based on their absorbance at 595 nm.

PBGD Activity Assay

The activity of wild type and mutant PBGD was assessed by measuring formation of uroporphyrin I from porphobilinogen (substrate). Although PBGD catalyses the formation of hydroxymethylbilane (HMB) (Figure 1.5), the produced HMB can be converted non-enzymatically to uroporphyrinogen I (Figure 2.4), which can be converted to uroporphyrin I by

spontaneous oxidation in the presence of ultraviolet light [9]. Uroporphyrinogen I, unlike uroporphyrinogen III, does not re-enter the mitochondrion as part of haem biosynthesis, and in this assay conditions are unfavourable for any downstream further processing of the uroporphyrin I. The production of uroporphyrin I is thus considered an accurate reflection of the production of HMB by the PBGD enzyme, as illustrated by Figure 2.5. BSA, of equivalent enzyme concentration, was used as a negative control. The protocol for this PBGD assay is detailed in Appendix 17.

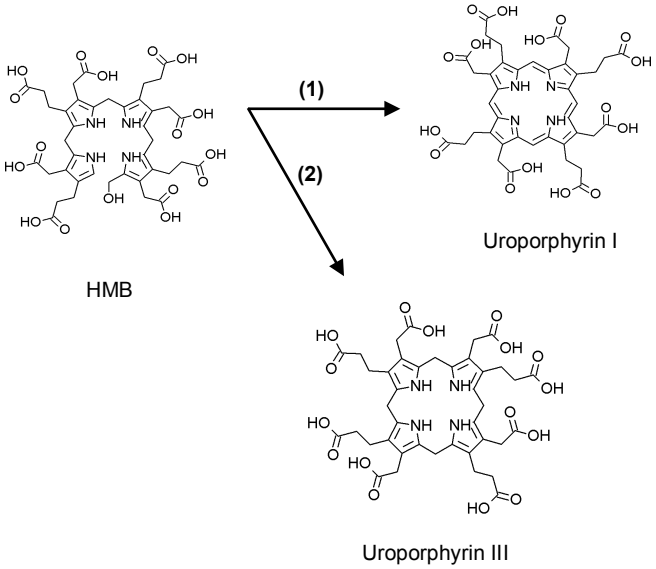


Figure 2.4 Illustration of the possible conversions of HMB (ChEBI_16645; or preurorphyrinogen), as adapted from [96]. Figure description as follows: **(1)** The conversion to uroporphyrin I (ChEBI_27484) non-enzymatically. **(2)** The conversion to uroporphyrin III (ChEBI_15436) via uroporphyrinogen synthase [9]. Image was generated by ChemSketch Version 2.0 [5].

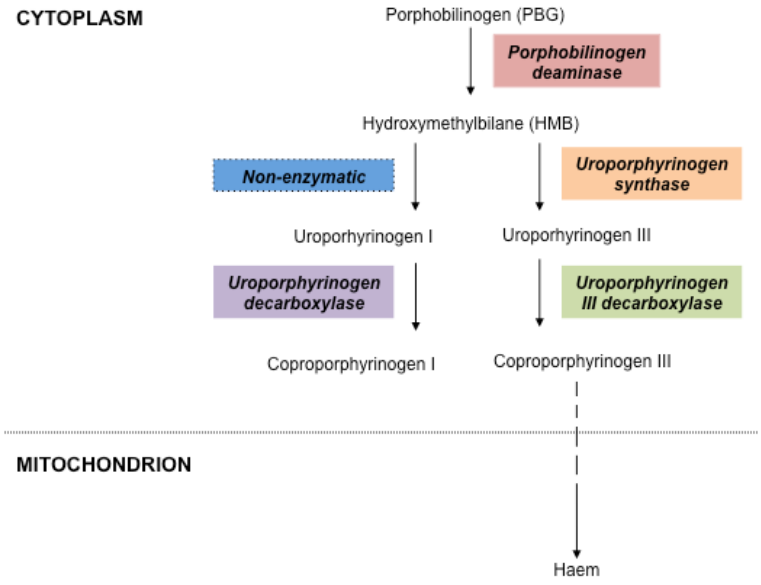


Figure 2.5 The reaction catalysed by porphobilinogen deaminase (PBGD) leads to the formation of HMB, that may be converted to uroporphyrinogen I non-enzymatically, as adapted from [9].

Kinetic Characterisation

In order to determine the extent to which the engineered mutants effect PBGD's kinetic function, the kinetic parameters, namely the substrate affinity (also Michaelis constant) (K_M), and the maximal velocity (V_{max}) were determined and compared to that of wild type enzyme. Data was fitted for a Michaelis-Menten kinetic curve using iterative curve fitting, done by GraphPad Prism, Version 6.03, for Windows (GraphPad Software, California, USA) using a least of square fit with a maximum of 1000 iterations. In the assay, increasing amounts of PBG (substrate) were used to a level of substrate excess. Thereafter, the enzyme turn-over rate k_{cat} as well as the efficiency of the enzyme, expressed as the ratio of k_{cat}/K_M was determined (Appendix 18).

Native PAGE Analysis

Principle

As a result of the step-wise polymerization of the PBG substrate by the enzyme, it has previously been demonstrated that it is possible to view differences in enzyme complex states, namely ES, ES_2, ES_3 and ES_4, each corresponding to the number of PBG molecules being added onto the formation of the pyrrole ring on a non-denaturing 10% PAGE gel [48,85,97]. The high concentration of acrylamide allows for the separation of the different complexes (Figure 1.11).

Procedure

The protocol for Native PAGE is similar to that of SDS-PAGE gels, except that the gel was a 10% acrylamide gel not a gradient gel, and the denaturing/reducing agents SDS and β-mercaptoethanol were not included in the sample loading dye. Protein samples were run on a small Bio-Rad Gel electrophoresis system (courtesy of UCT Department of Surgery's Research Lab), not the older systems used for SDS-PAGE gels routinely in our Porphyria Lab. The methodology was based on recently published work, and the protocol for the Native PAGE gel was kindly sent from Ms Helene Bustad, following personal correspondence [98]. Further details of the protocol is described in Appendix 25.

DPM Co-Factor Analysis

The co-factor assay is based on the unique reaction of the DPM co-factor with Ehrlich's reagent, which consequently results in the tautomerisation of the dipyrromethane to a pyrromethane [58,84]. This may be tested chemically by the observation of a change in absorbance spectra over a time period [79,80]. Thereby, in order to test for the presence of the DPM co-factor, equal volumes of 0.5mg/ml of purified wild type, K98E or K98R mutant

PBGD, and a modified Ehrlich's aldehyde were incubated in a 96 well plate, at room temperature. A spectral scan ranging from 440 – 650 nm was immediately performed and repeated after a 20-minute interval. The scan was examined for a spectral shift (see Appendix 19 for details).

DPM Co-Factor Removal

This methodology is based on the work of Hart et al. 1988 [80]. The aim of this was to remove the bound DPM co-factor from wild type PBGD, and following this to reconstitute the enzyme activity by the addition of the substrate (PBG), or, additionally in our case, the DPM co-factor itself.

Procedure

In brief, wild type protein was subjected to hydrolysis with concentrated HCl (1M) in order to cleave the DPM from the bound Cysteine 261 residue (Figure 1.7) [67]. Following 25-hour incubation in the dark at room temperature, protein was refolded using a series of dialysis steps. Finally, the *'apo'* protein, lacking the DPM co-factor was concentrated and kept at 4°C. The detailed protocol is outline in Appendix 20.

PBGD reconstitution assay

Principle

This protocol is based on the principle of Hart et al. (1998) whereby PBG was added to the generated *'apo'* PBGD enzyme in order to 'rescue' and attempt to restore activity[80]. For our interest, we wanted to test whether we could repeat this method, and, if so, if it was possible to test whether the '*apo'* - enzyme would accept, and show preference for its DPM co-factor versus the PBG substrate [84]. The rationale for this is further detailed in the discussion and not outlined here.

Procedure

The assay is based on our previously detailed assay for PBGD activity. The methodology was slightly altered in order to incorporate testing of the DPM co-factor, either independently or in combination with the substrate. Details for this assay are outlined in Appendix 21.

DPM Solubility

The DPM co-factor is commercially available and was purchased from Frontier Scientific (USA address and then represented by Labretoria in South Africa). However, very little information is available regarding its solubility in organic solvents other than methanol. We therefore had to determine whether the DPM cofactor was soluble in solutions comprised of different percentages of 1M imidazole and methanol (results shown in Chapter 3).

Circular Dichroism Analysis

The PBGD wild type and mutant enzymes were subjected to circular dichroism studies in order to investigate possible changes in protein secondary structure, and conformational stability as a result of the different mutations. This work was undertaken at the Central Analytical Facility (CAF) at Stellenbosch University, under the guidance of Dr. Jaco Brand (CAF) and mentoring from Professor Marina Rautenbach (Department of Biochemistry).

Principle

Circular Dichroism is based on the principle of polarised light and the differences in absorption of optically active (chiral) molecules within a chemical bond, peptide as well as proteins [99]. As shown in Figure 2.6 polarised light is light that travels in a single direction.

With circular dichroism, light is circularly polarised, allowing for either the rotation of light in a clock- or counter-clock wise direction [99]. When asymmetric molecules, like proteins, interact with this circularly polarised light, they may absorb the right- and left-handed light a different extent, as shown in Figure 2.7 [99,100].

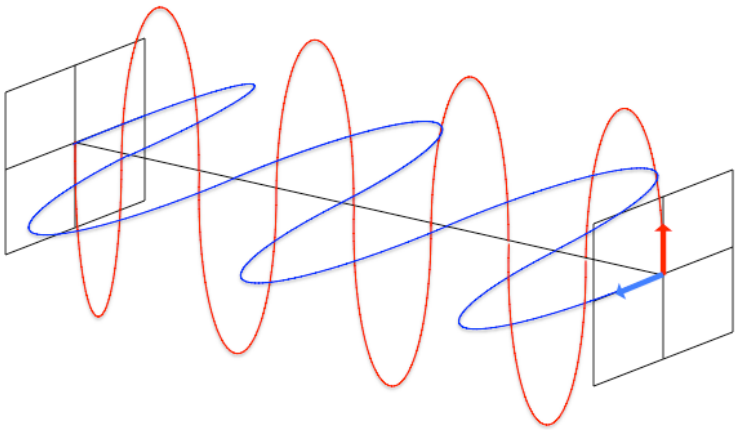


Figure 2.6 A simplified illustration of different polarised light. Each light beam has been polarised to allow only for the transmission of light in the indicated direction (by arrows).

Procedure

The methodology for the far-UV analysis of the wild type and mutant PBGD enzymes is based on the work done by Bustad et al. (2013) [98], except our work was performed on a Chirascan Plus CD Spectrophotometer (Applied Photophysics Limited, United Kingdom)

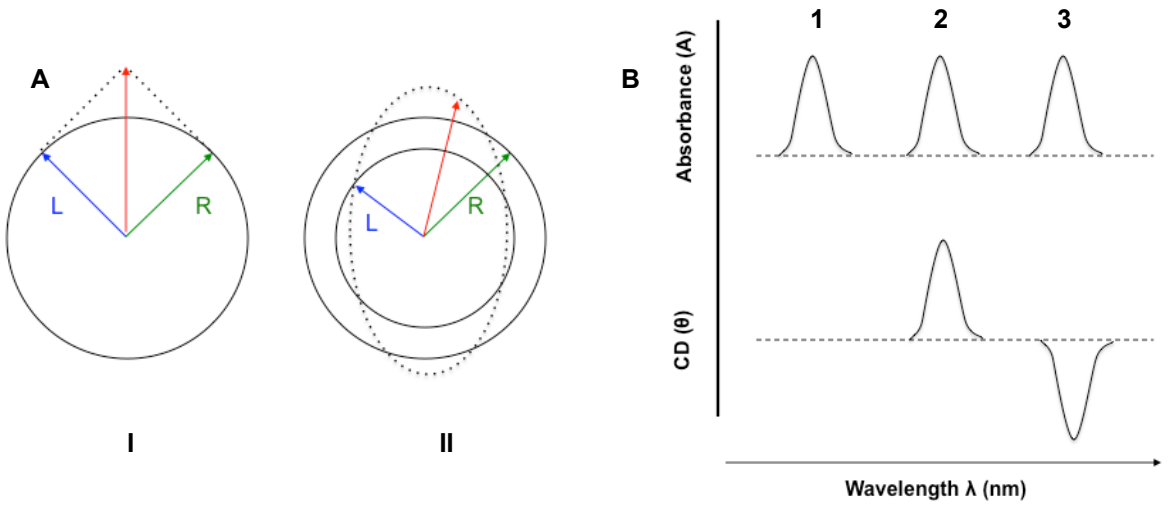


Figure 2.7 An illustration of the origin of the CD effect. 1. Figure descriptions as follows: **(A)** Circularly polarised light components, left (L), and right (R). Seen in **(I),** When the L & R components have the same amplitude, it generates plane polarised radiation when combined. **(II)** If however, the components have different magnitudes, the subsequent light is elliptically polarized (dashed line). **(B)** Illustration of the relationship between ellipticity and absorbance. Band 1 has no CD spectrum, due to an achiral chromophore; Band 2 has a positive CD spectrum with L absorbed more than R; Band 3 has a negative CD spectrum, as more R is absorbed than L [99].

Freshly purified protein sample was exchanged to a new buffer (CD Buffer) which is not optically active with the use of PD-10 desalting columns [100]. Details for the buffer exchange are discussed in Appendix 22. Protein was then diluted to a final concentration of 7μM. Scans were set up in a range of 190 – 300 nm, with the temperature controlled at 25°C by a far-UV analysis was done in the range of 190 – 260 nm, to measure the secondary (α-helical and β-sheet) content of wild type and mutants. The near-UV analysis (260 – 300 nm) was done to measure the 'tertiary footprint' of the protein, by the presence of the aromatic rings, phenylalanine, tyrosine and tryptophan peaks at between 255 – 270 nm, 275 – 282 nm, and 290 nm respectively [99]. Protocol further detailed in Appendix 23.

Data Fitting and Analysis

CDPro

Data was corrected for the buffer (blank) and normalized to the final five readings (295 – 300nm) in order to have a more accurate representation of the circular dichroism data. The normalization of the data allowed for adjustment of the scale without affecting the meaning of the dataset. Data was further normalized by calculating the mean residue ellipticity ($[\theta]_{MRW}$), according to the following formula illustrated below (Figure 2.8) [99]. This ensures the data for each sample is independent of the protein residue composition [101].

$$[\theta]_{MRW} = (10^6 \times \theta) / c \times l \times n$$

Figure 2.8 The formula for the conversion of ellipticity (θ) in millidegrees (mdeg) to mean residue ellipticity $[\theta]_{MRW}$ where c is the concentration of the protein in μM, (l) the path length in millimetres (mm), and (n) the number of residues in the protein (361 for non-erythroid).The multiplication with 10^6 allows for final answer expressed in ($deg.cm^2.dmol^{-1}residue^{-1}$).

Data was analysed using CDPro, by the CONTINLL [102,103], SELCON3 [104,105] and CDSstr algorithms [106]. The CDPro software, however, is only able to estimate secondary structures within the range of 190nm to 240 nm, as well as data in molar circular dichroism ($\Delta\varepsilon$). The conversion from mean residue ellipticity ($[\theta]_{MRW}$) is shown below in Figure 2.9.

$$\Delta\varepsilon = [\theta]_{MRW} / 3298$$

Figure 2.9 The relationship of molar circular dichroism and the mean residue ellipticity $[\theta]_{MRW}$ as noted from [99].

Ten reference sets were screened against each sample and the secondary protein estimation calculated from the mean of the top 5 reference set for each algorithm. The more converged the results are, the higher the confidence in the protein estimation of secondary structure [107]. One-Way ANOVA statistical tests were performed from the mean of the final result of the three different algorithms (CDStr, CONTINLL and SELCON3), in order to test if changes amongst mutants were significant in comparison to wild type.

This methodology for data analysis and choice of CDPro algorithms is based on the work of Professor Marina Rautenbach (SU), with whom this work was done. A more detailed description of the different algorithms can be found on the Dichroweb website [http://dichroweb.cryst.bbk.ac.uk/html/references.shtml].

CDNN

The CDNN algorithm is another software used for the estimation of the protein secondary structure. Although it is not as robust as the CDPro, it does allow for a graphical comparison of the wild type to different mutants, as it includes random coils and β-turns [108].

Thermal Denaturation Profiles

Procedure

The methodology for the thermal analysis of the wild type and mutant PBGD enzymes is also based on the work done by Bustad et al. (2013) [98]. Protein samples were prepared similarly

to the far-UV scans (Appendix 23). The ellipticity was recorded at 208 nm and 222 nm as a function of temperature between 20 and 90°C. Stepped temperature ramping was allowed for a total of 300 seconds at each temperature point – this was done in order to allow for the cuvette to reach as close to the desired 'set' temperature. Independently purified samples were repeated twice, with adaptive sampling on the Chirascan Software (Applied Photophysics Limited, United Kingdom). Detailed protocol further discussed in Appendix 24.

Data Fitting and Analysis

Data from the thermal spectrums were fitted to this a two-state unfolding model as detailed in Figure 2.10. During a two-state unfolding of a protein, there are only two species present in considerable amounts, namely the folded (or native) protein (N) and the unfolded protein (U). Therefore, the following is true for the total account of folded and unfolded protein:

$$\mathbf{f_N + f_U = 1} \quad (1)$$

So, it may be deduced that, at any point during the transition from native to unfolded protein, the observed signal is as follows:

$$\mathbf{\theta_t = \theta_N f_N + \theta_U f_U} \quad (2)$$

By rearranging this equation (2), the fraction of folded protein may be expressed as follows:

$$\mathbf{f_N = (\theta_t - \theta_U) / (\theta_N - \theta_u)} \quad (3)$$

Figure 2.10 The derivation of the fraction of unfolded protein, as determined from the circular dichroism data from thermal curves, as described previously. **(1)** The fraction of protein species, where the sum of native (folded) protein (f_N), and unfolded protein (f_U) are equal to 1. **(2)** The representation that the measured/observed ellipticity (θ) at any given point (i.e. each temperature point) is the sum of the fraction of native protein and unfolded protein. **(3)** The final equation for the fraction of folded protein.

The use of equation 3 allowed for the conversion of data, as presented in millidegrees, to be converted to fraction of folded protein for each respective temperature point. Results were plotted in GraphPad Prism Version 6.02 for Windows, expressed in millidegrees [98,109,110]. Our interest was not to further investigate the thermodynamic principles of the protein folding and refolding; hence we did not further extrapolate van Hoff's formula for free energy (ΔG). The melting Temperatures (T_m) for wild type and each mutant was determined from the fitted data at 222 nm, that was plotted on a non-linear regression, and fitted to the Levenberg–Marquardt algorithm using the Origin Pro Software (OriginLab, Northampton, MA) [110].

Protein Structural Analysis

As highlighted in Chapter 1, several crystal structures have been published for the PBGD enzyme (Table 1.1). The most characterised to date is the structures in *E.coli*. However, we preferred to work with the human enzyme, as we were expressing two clinical human mutations.

The two crystal structures available for the human enzyme both lack several loop regions, including the active site loop as mentioned previously [60,67]. However, it is possible to perform homology studies amongst the different crystal structures, and thereby have a modelled enzyme, based on the human form, with the 'missing loops' modelled from other crystal structures. The online bioinformatics tool and database, Expasy, hosts a range of homology systems that provides an online interface for homology predictions [111].

We utilised CPH models Server 3.0 as is suggested for protein homology [112]. The sequence from the Protein Data Bank (PDB) file from the 3ECR crystal was used, as this crystal structure did not contain a mutation, like the 3EQ1 [60,67]. The CPH Server can access all the data of known published crystal structures for the sequence that it recognises. Thereby, the active site loop could be 'modelled' from the crystal in the higher plant species *Aradopsis thaliana* [68]. In addition, it is possible to submit the protein sequence with the desired mutation prior to the homology modelling. This allowed for the incorporation of the mutation into the homology modelling, instead of just altering the amino acid manually.

The final modelled structure(s) (wild type and mutants), based on the human enzyme form (3ECR), were minimised with 150 steps of steepest decent, followed by 100 steps conjugated descent in order to allow for lowest energy configuration. The molecular graphics and analyses were performed with the UCSF Chimera package. Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311) [61].

Chapter 3

Results

Introduction

This chapter details all of the relevant results obtained for this dissertation. It follows the same order as that described in Chapter 2 for the methodology.

Site-Directed Mutagenesis

To assess the contribution of PBGD active site residues towards the enzyme's activity and co-factor interaction, single amino acid conversions (K98E, K98R and K98A) were introduced and the mutants expressed in JM109 supercompetent *E.coli* cells using the QuickChange® methodology.

Screening of the entire *HMBS* gene (sequence given in Appendix 2) revealed that no additional mutations had been introduced (data not shown). Following validation that each engineered sample was a true representation of the desired mutation, we pursued the expression, purification and characterisation of each mutant protein product, in terms of kinetics, co-factor interaction as well as structural analysis.

Restriction Analysis to confirm K98R sequence

The K98R mutant allowed for a restriction analysis of a new identified restriction enzyme site by *Bsm FI* and showed successful mutation for all eight colonies screened (Table 3.1 and Figure 3.1). An additional cutting site was generated, indicated by the presence of bands at 114bp and 258bp.

Table 3.1 *Bsm FI* restriction analysis for screening of the K98R.

Mutation	Size (bp)	Restriction Endonuclease	Recognition Sequence	Digest Products (bp)	
				Wild Type	Mutant
K98R	677	*Bsm FI*	5'-GGGAC$(N)_{10}$-3'	305; 372	114; 258; 305

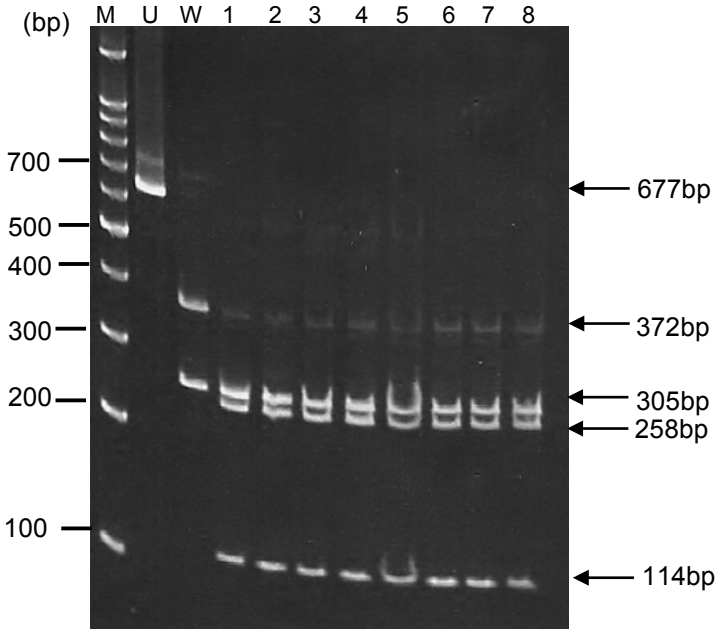


Figure 3.1 The *Bsm FI* restriction digest of the K98R mutation. Lanes are labelled as follows: **M)** Molecular Marker **(U)** Uncut PCR product **(W)** Wild type PBGD; **(1- 8)** Selected K98R colonies.

Confirmation of Mutated Sequences

Direct sequencing showed successful engineering of the mutants. Figure 3.2 shows a partial direct sequencing for each mutant and Table 3.2 outlines the nucleotide change(s) in each respective mutant. As noted in Chapter 2, samples were sequenced in both the forward and reverse direction, in order to 1) confirm the mutation at the right position, and 2) to observe that no additional mutations were introduced. Samples were compared to the Genbank sequence for human *HMBS*, as noted in Appendix 1.

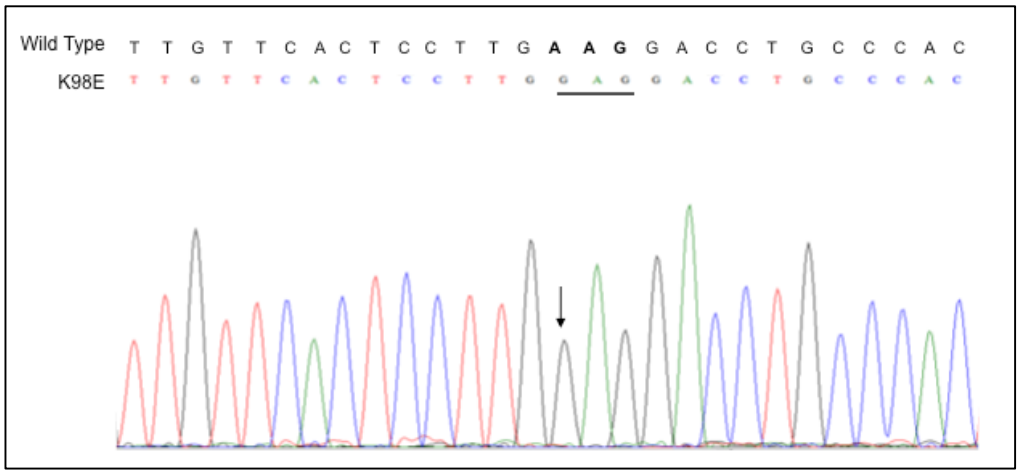
For the K98A mutation, two base pairs were altered from the wild type lysine (AAG), to GCG for alanine. As detailed in Appendix 9, the PCR for the mutagenesis was adapted slightly in order to increase the cycle time from 12 cycles to 16 cycles, as suggested by the QuickChange® protocol.

Table 3.2 The base-pairs of residue lysine 98 of the PBGD enzyme, illustrating the respective nucleotide change(s) in each mutation.

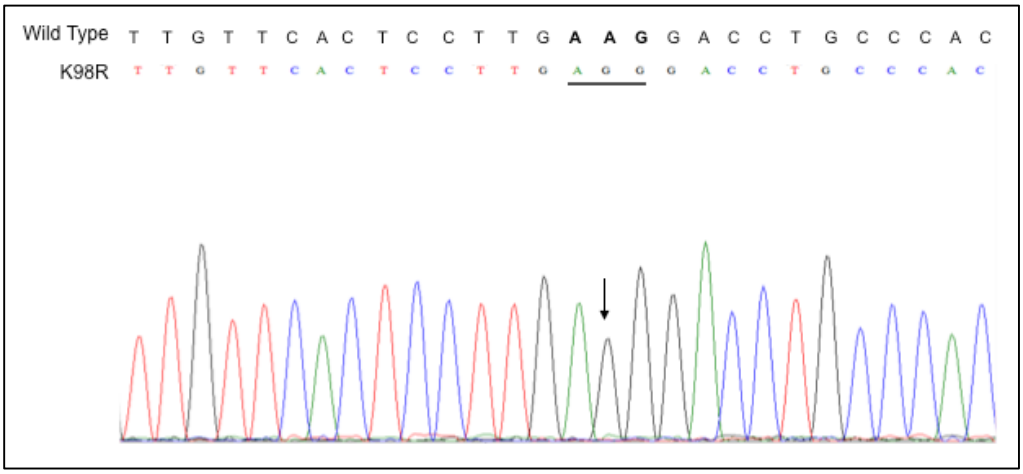
	Wild Type	**K98E**	**K98R**	**K98A**
Residue 98	AAG	GAG	AGG	GCG

.

I



II



III

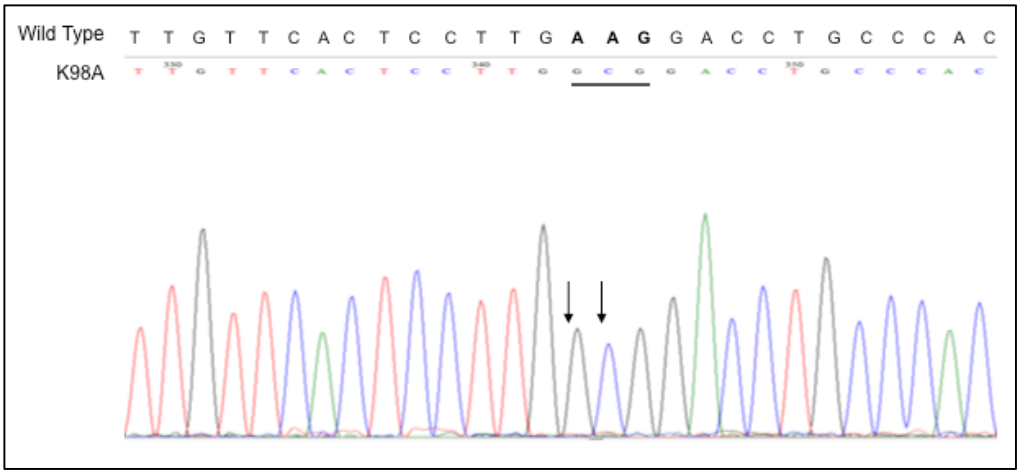


Figure 3.2 Partial direct sequencing of the K98E, K98R and K98A mutants using a forward primer. The wild type sequence is shown above the mutant read sequence. The base pair change(s) is identified by an arrow and the underlined nucleotides signify the 98 residue. **(I)** K98E mutation showing the base change from AAG to GAG **(II)** The K98R mutation showing the base change from AAG to AGG and **(III)** The K98A showing the base changes from AAG to GCG.

Expression and Purification of PBGD

The successfully expressed wild type and mutant proteins were subsequently purified using TALON™ metal affinity chromatography, as described in detail in Chapter 2. Previous work in our laboratory on PPOX in the same expression system showed that IPTG did not significantly improve protein expression, and the PBGD expression levels we obtained without IPTG induction were adequate for the intended purpose [94,113]. Table 3.3 illustrates the purification table for the PBGD wild type and mutant enzymes. The K98E had very high expression levels compared to the others. However, the activity of the eluted protein was still significantly lower in comparison to wild type.

Table 3.3 A representative purification table combining the data obtained in a number of protein runs of the wild type and mutant PBGDs showing protein yield, activity and percentage recovery.

	Volume (ml)	[Protein] (µg/ml)	Total Protein (µg)	Specific Activity (pmol/µg/h)	Total Activity pmol/h	Recovery (%)
WT						
Load	25	14.7	368	26	9568	100
Void	23	13.2	304	2.5	760	8.0
Wash	10	0.6	6.0	24	144	1.5
Eluate	1.5	6.0	9.0	555	4995	52
K98E						
Load	30	14.4	432	94	40608	100
Void	30	12.7	381	1.7	648	1.6
Wash	11	0.7	7.7	40	308	0.8
Eluate	1.4	17.7	25	297	7425	18
K98R						
Load	30	9.6	288	4.2	1210	100
Void	28	8.9	249	2.5	623	50
Wash	11	0.4	4.4	0.3	1.3	0.1
Eluate	1.5	3.4	5.1	0.5	2.6	0.2
K98A						
Load	28	6.2	174	1.2	209	100
Void	25	6.5	163	0.7	114	55
Wash	10	0.2	2.0	0.2	0.4	0.2
Eluate	1.5	4.3	6.5	0.6	3.9	1.9

Table 3.4 shows just the specific activities of the pure eluted proteins, providing a comparison of the mutants to wild type, expressed as a relative percentage to the wild type PBGD, which was normalized to 100%.

Table 3.4 The specific activities for the wild type, K98E, K98R and K98A and the relative specific activity (expressed as pmol/µg/h) and the fraction (shown as %) to wild type specific activity.

	Specific Activity	
	pmol/µg/h	**Relative to Wild Type (%)**
Wild Type	555	100
K98E	297	54
K98R	0.5	0.1
K98A	0.6	0.1

To check the purity of the protein and confirm the correct size of eluted protein, specifically 44 kDa for the non-erythroid form, a gradient SDS-PAGE was performed for each sample as illustrated in Figure 3.3. The images presented are illustrative of the degree and nature of purifications achieved for each expressed mutant.

A useful feature of any form of PBGD that has its DPM co-factor bound is that it exhibits a characteristic pink colour (see Chapter 1, Figure 1.8). This feature was utilised in monitoring our PBGD purification. However, for the K98R and the K98A mutants, several fractions of eluted protein had to be collected in the region of expected elution, since the protein had no characteristic pink colour (as in wild type and K98E). These additional fractions were loaded onto the gel in order to deduce which fraction contained the purified PBGD protein. Following confirmation of correct protein size and purity on the SDS-PAGE as well as the Bradford Protein Assay (as described in Chapter 2), the eluted protein samples of adequate concentration were pooled and concentration of pooled sample determined.

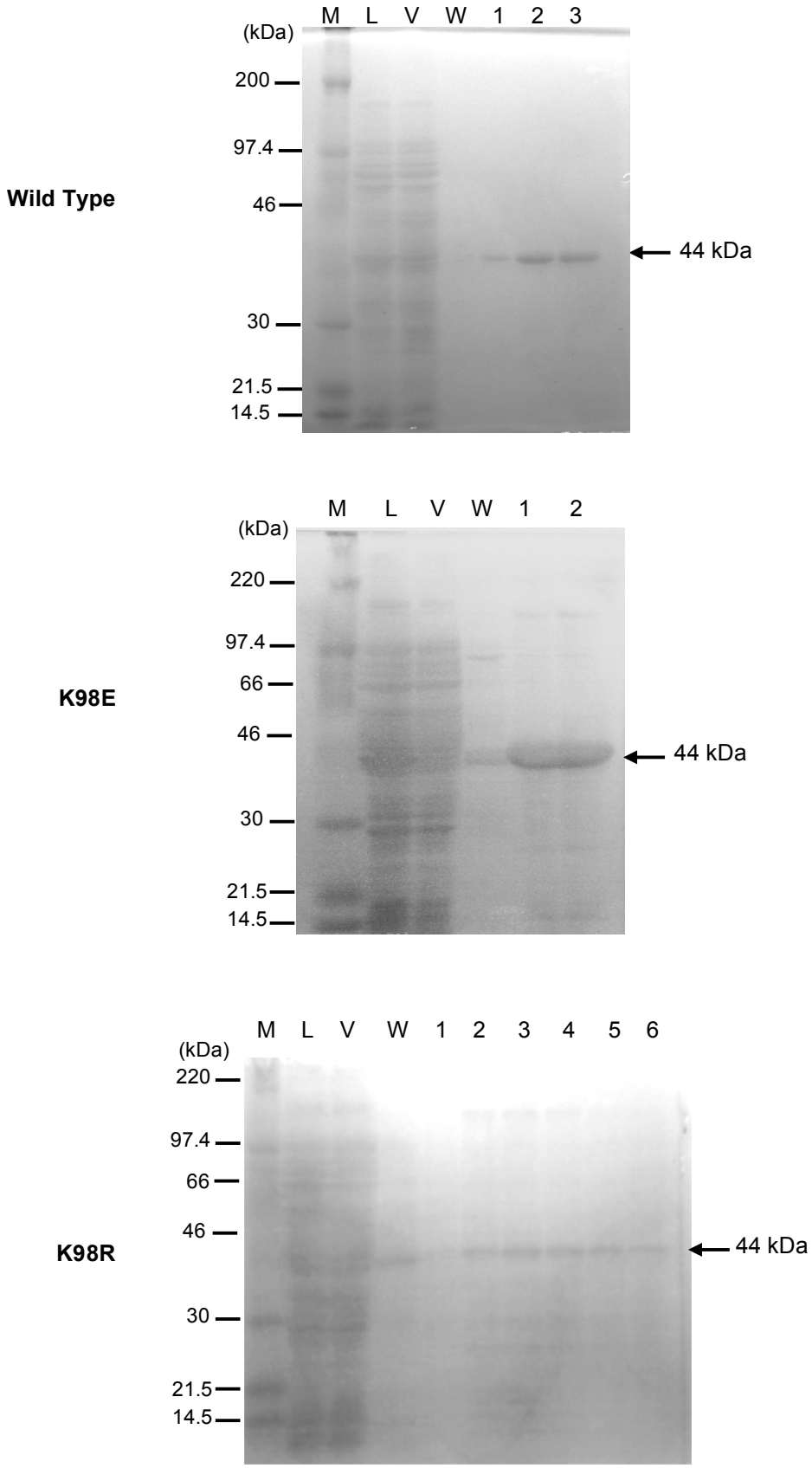


Figure 3.3 The 7.5 – 17.5% gradient SDS-PAGE of wild type and mutant PBGD. Lanes are labelled as follows: **(M)** Rainbow Marker, **(L)** Load, **(V)** Void, and **(W)** Wash. With regards to the eluted proteins – Wild type: **(1-3)** Purified eluted protein. K98E: **(1&2)** Purified protein, and for K98R several fractions were collected **(1-6)**, of which the first fraction **(1)** contained no protein.

Optimization of Mutant Expression

Initial purification of the K98A revealed issues pertaining to the stability of the protein as well as to the high level of non-specific proteins that were binding to the TALON™ resin and thus eluting with the protein of interest (Figure 3.4).

In order to address these issues several alternatives were attempted (as detailed in Chapter 2), finally reaching a satisfactory pure protein by adding a total of 10 mM imidazole in the sonication (1st buffer) (Figure 3.5), reducing the resin volume used to 1 ml, and performing the entire purification at 4°C (Figure 3.6).

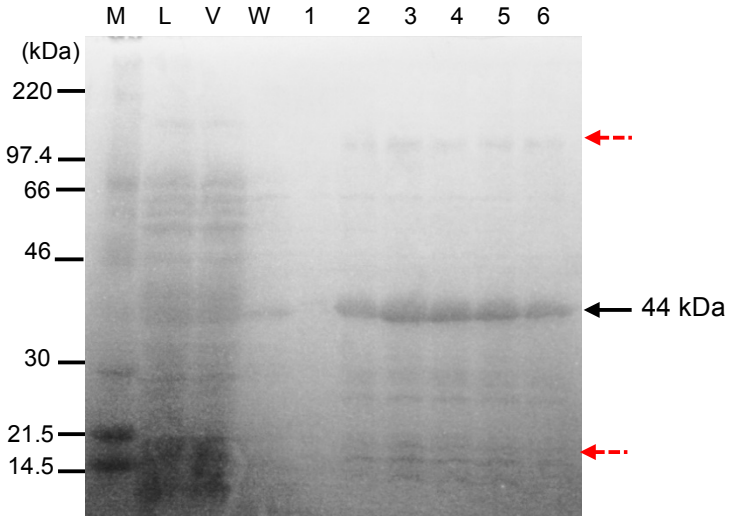


Figure 3.4 The initial 7.5 – 17.5% gradient SDS-PAGE of the K98A mutant PBGD protein. Lanes are labelled as follows: **(M)** Rainbow Marker **(L)** Load **(V)** Void **(W)** Wash **(1-6)** eluted fractions. Arrows highlighted in red illustrate some of the contaminants that co-eluted with the K98A protein of interest.

As illustrated in Figure 3.5, the addition of the 5 mM imidazole together with the reduction of resin volume of 1.2 ml enabled us to eliminate a proportion of the non-specific proteins whilst at the same time maintaining the yield of desired protein. However, this strategy did not eliminate all of the non-specific protein fractions.

We then further optimised the protocol for purification of the K98A mutant protein, which included the reduction of the TALON™ resin to 1.0 ml, a total of 10 mM imidazole added to the sonication buffer, as well as the purification performed at 4°C (Figure 3.6). Although this slightly affected protein yield, the total protein obtained was still satisfactory to perform downstream assays and characterisation. Although we had several other optimization strategies including alterations of the wash buffer (as discussed in Chapter 2), these strategies proved unsuccessful and are not shown.

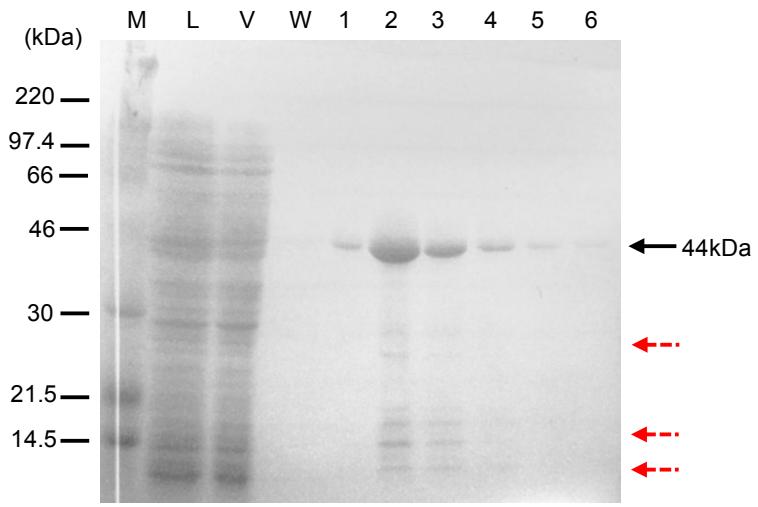


Figure 3.5 The 7.5 – 17.5% gradient SDS-PAGE of the K98A mutant PBGD protein of the purification done with the reduced TALON™ resin volume of 1.2 ml together with the 5 mM imidazole in the sonication buffer. Lanes are labelled as follows: **(M)** Rainbow Marker **(L)** Load **(V)** Void **(W)** Wash **(1-6)** eluted fractions. Arrows highlighted in red illustrate some of the contaminants that co-eluted with our K98A protein of interest.

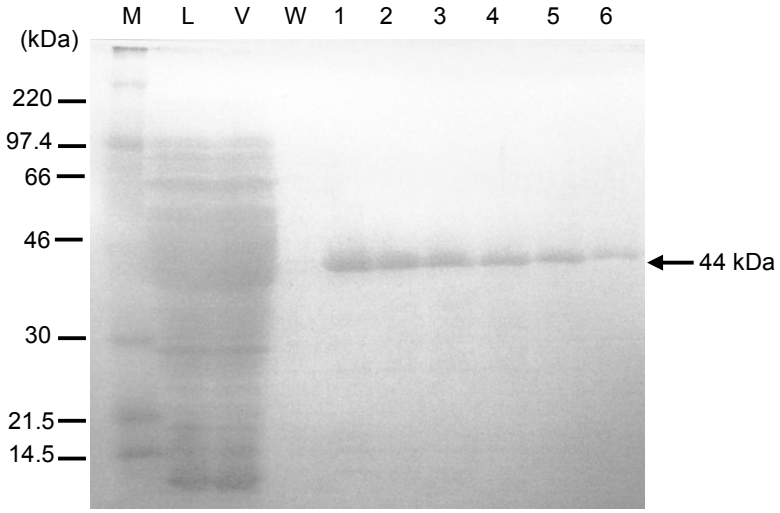


Figure 3.6 The 7.5 – 17.5% gradient SDS-PAGE of the K98A mutant PBGD protein of the purification done with the further reduced TALON™ resin volume of 1.0 ml together with the 10 mM imidazole in the sonication buffer. Lanes are labelled as follows: **(M)** Rainbow Marker **(L)** Load **(V)** Void and **(W)** Wash and **(1-6)** eluted fractions.

With regards to the impurities for the K98R and K98E mutants, the yield for the K98E of the desired protein was satisfactory and the comparison of the concentration of the PBGD versus the impurities was very low, hence the protein was considered of adequate purity. With regards to the K98R, the same optimization strategy as was done for the K98A was performed, and it was found that the protein purity could be improved to some extent, as shown in Figure 3.7 below.

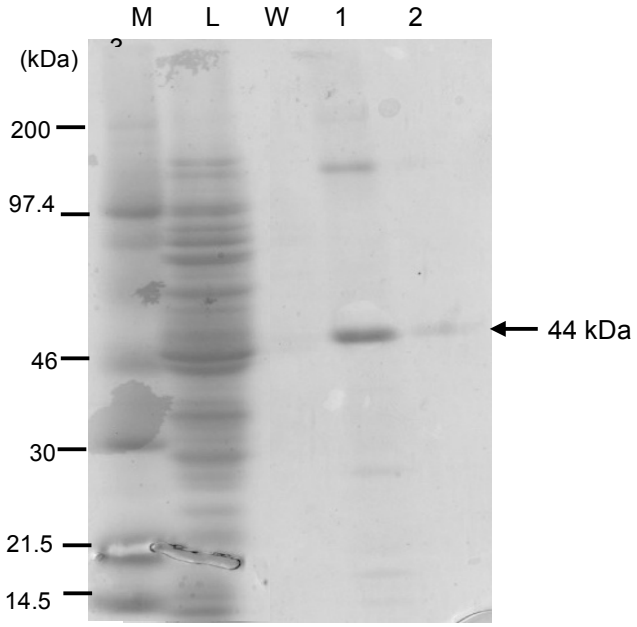


Figure 3.7 The 7.5 – 17.5% gradient SDS-PAGE of the K98R mutant PBGD protein of the purification done with the further reduced TALON™ resin volume of 1.0 ml together with the 10 mM imidazole in the sonication buffer. Lanes are labelled as follows: **(M)** Rainbow Marker **(L)** Load **(W)** Wash **(1-2)** eluted fractions.

The protein yield for the K98R was still very low, although we still had an adequate amount of protein to perform all assays. In Figure 3.7, only two protein fractions were loaded, as the samples were pooled prior to the running of the gel. In all of the optimization strategies, although the protein purity was increased, the protein stability remained problematic, even when purifications were performed at 4°C.

Kinetic Characterisation

Figure 3.8 below shows comparative illustrative kinetic substrate-velocity plots for the wild type, K98E and K98R generated using GraphPad Prism, Version 6.01.

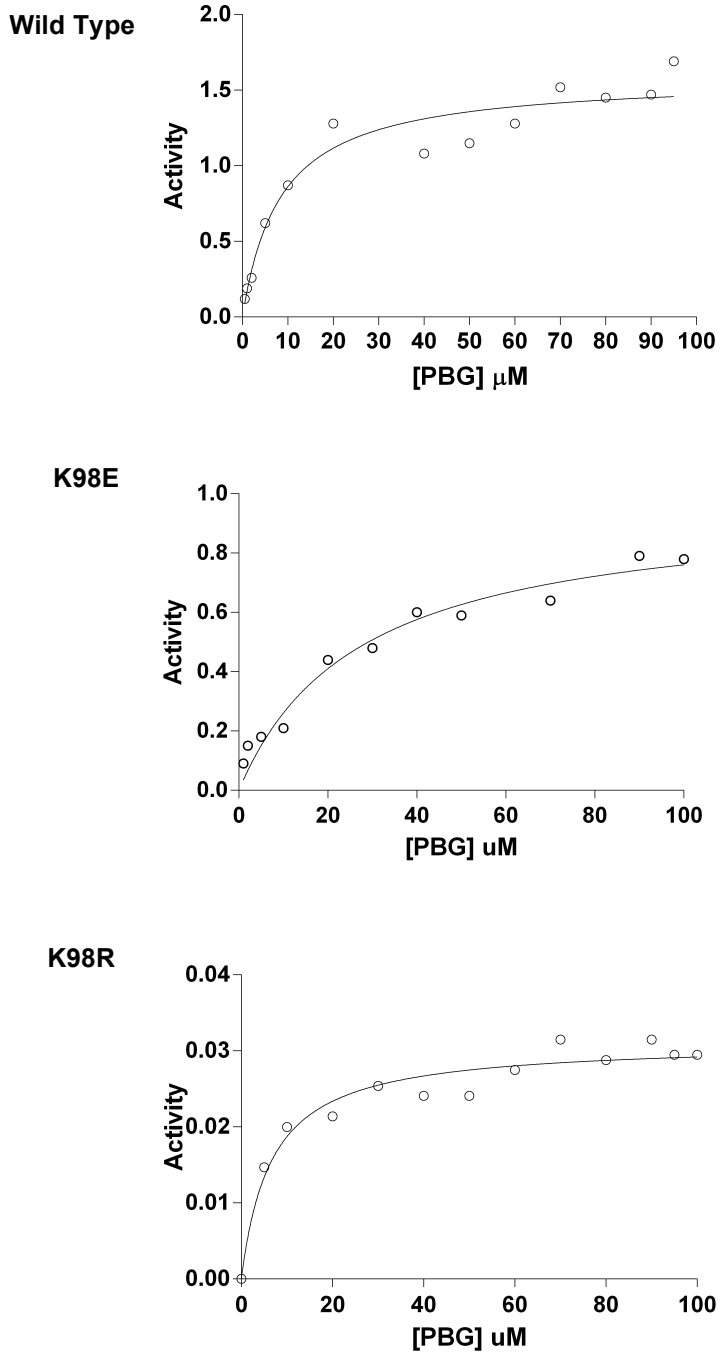


Figure 3.8 The Michaelis-Menten, substrate-velocity plots for the wild type, K98E and K98R mutants following the 20-minute kinetic assay. Activity is expressed as the number of picomoles of Uroporphyrin I produced per nanogram of PBGD added per hour [pmol/ng/h].

Table 3.5 illustrates the kinetic parameters calculated for the wild type, K98E, K98R and K98A from the substrate-velocity data. The results represent the mean from three independent assays using purified protein. The kinetic characterisation of the K98A mutant proved to be difficult. The extremely low/negligible activity of this mutant (Table 3.5; Figure 3.9) resulted in small variations in fluorescence readings correlating to severe scattering of kinetic plots and no stable reading was recorded, results ranging from 0.5 to 500 μM. Several alternatives were attempted including an increased incubation time, increases and decreases of enzyme concentration, to no avail. A graphical output of Table 3.5 is presented in Figure 3.9 below.

Table 3.5 The kinetic constants for the engineered K98A PBGD enzyme in relation to the previously analysed wild type and other mutant PBGD enzymes. Standard error is expressed as standard error of the mean (SEM).

	K_M (μM)	V_{max} (pmol/ng/hour)	k_{cat} (s^{-1})	k_{cat}/K_M ($\mu M^{-1}.s^{-1}$)	V_{max} % of WT
Wild Type	8.49 ± 0.136	1.71 ± 0.299	51.3 ± 8.95	6.04	100
K98E	30.7 ± 2.97	1.06 ± 0.072	31.7 ± 2.17	1.03	62.00
K98R	5.69 ± 0.812	0.029 ± 0.002	0.878 ± 0.06	0.15	1.70
K98A	-	0.026 ± 0.005	0.184 ± 0.03	-	1.52

The results for the kinetic constants for wild type and mutants, as shown in Table 3.5, are (where necessary) expressed to three significant figures. We did this as it was difficult to express the data to a two decimal notation with the low levels of activity for the K98R and K98A mutants. From Figure 3.9 it may be noted that the substrate affinity for the K98E mutant illustrated an approximate three-fold increase in comparison to wild type, suggesting a significant reduction in the substrate affinity of this mutant. The K98R had a comparable K_M to that of wild type (Figure 3.9).

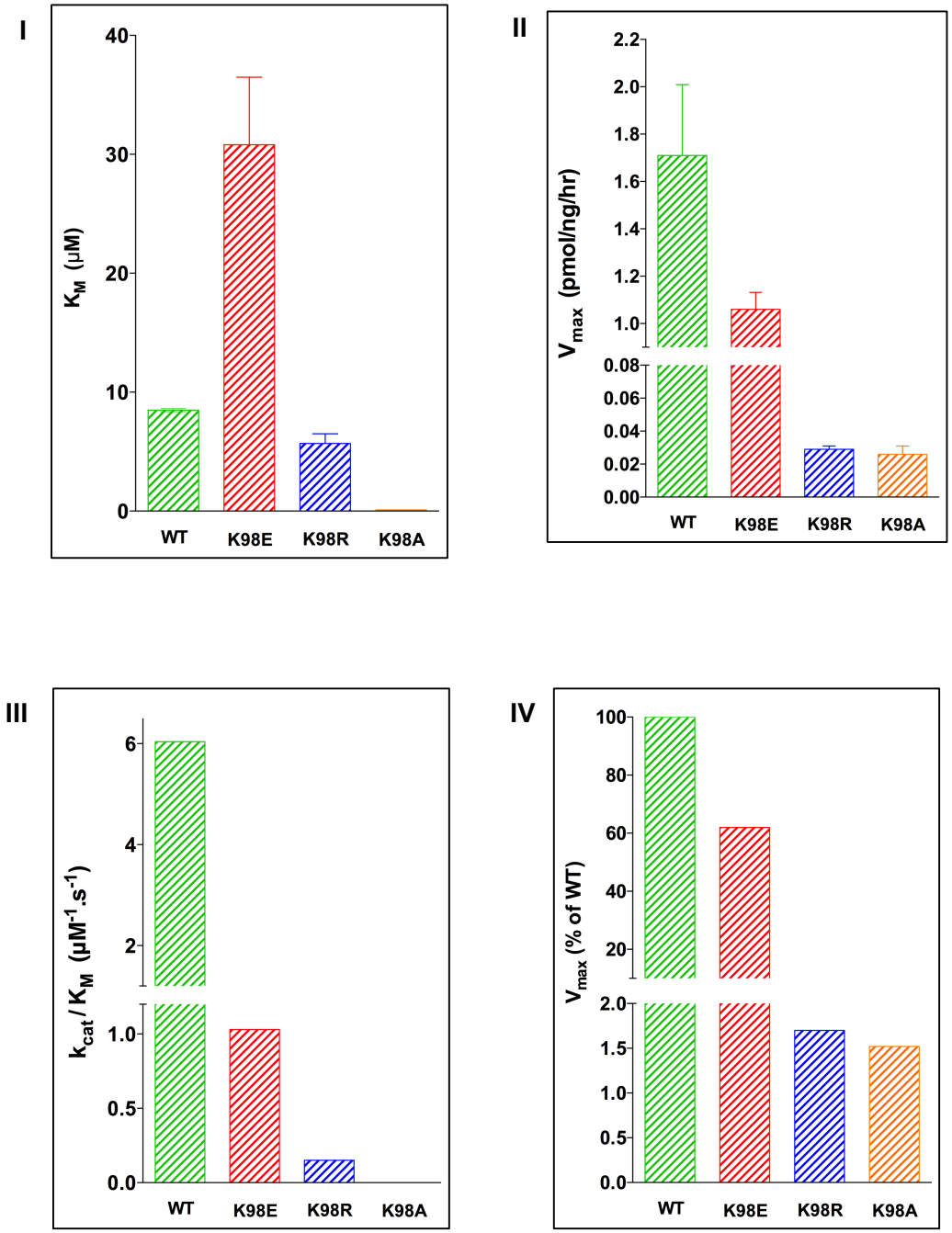


Figure 3.9 A graphical illustration, comparing the kinetic constants as determined in Table 3.5. **I)** Illustrates the comparative substrate affinity (K_M) of each sample. **II)** The maximum velocity achieved for each sample (V_{max}). **III)** The catalytic efficiency of each respective sample, as indicated by the (k_{cat}/K_M) ratio. **IV)** The V_{max} for each sample expressed as a percentage of wild type (%WT).

Native PAGE Analysis

Native PAGE gels were run in order to observe if the enzyme-substrate complex profiles for each mutant was different to that of wild-type. In addition, it was interesting to see whether any observed differences could be correlated to the mutants' kinetic results.

The different bands on the gel for wild type PBGD illustrate the different enzyme substrate complexes, as discussed in Chapter 1, Figure 1.11. However, following several optimization strategies and attempts, it appeared that none of the mutant proteins were stable enough to observe the enzyme complexes, as could be observed in the wild type, illustrated in Figure 3.10.

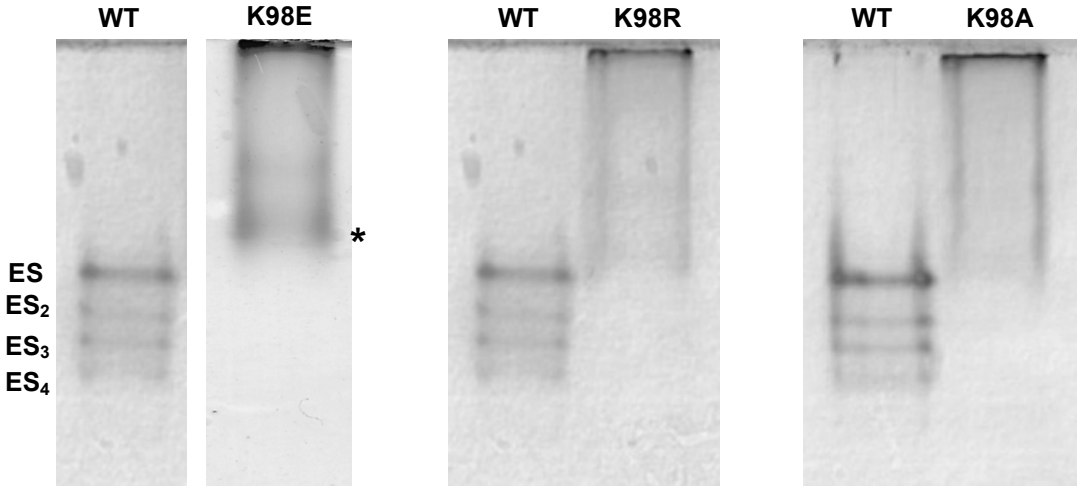


Figure 3.10 The 10% native PAGE for wild type PBGD and each of the mutants. The different bands indicated in the wild type enzyme are described as followed: ES – enzyme with one PBG; ES_2 – enzyme with 2 PBGs; ES_3 – enzyme with 3 PBGs; ES_4 – enzyme with 4 PBGs. The (*) indicated in the K98E represents a band on the gel observed migrating higher than the ES band.

The band that appears higher on the gel in the K98E mutant (indicated by the *), has been previously suggested in the literature to be the partially folded holoenzyme. This might suggest why K98E, although still able to bind the co-factor, still appeared unstable and shows structural defects, which is further discussed in Chapter 4.

DPM Co-Factor Analysis

The aim of this assay was to investigate the effect of the active site mutations on the enzyme's ability to bind its DPM co-factor. Figure 3.11 shows the spectral scan (450 – 650nm) for the wild type, K98E, K98R and K98A mutants. Both the wild type and K98E mutant PBGD illustrate the characteristic spectral shift from 564 nm to 495 nm over a 20-minute time period thus confirming the ability of the wild-type and K98E PBGD mutant to bind the co-factor. In contrast the absence of a spectral shift for the K98R and K98A mutants confirmed the loss of cofactor binding (which we had speculated due to the loss of characteristic 'pink' colour during purification).

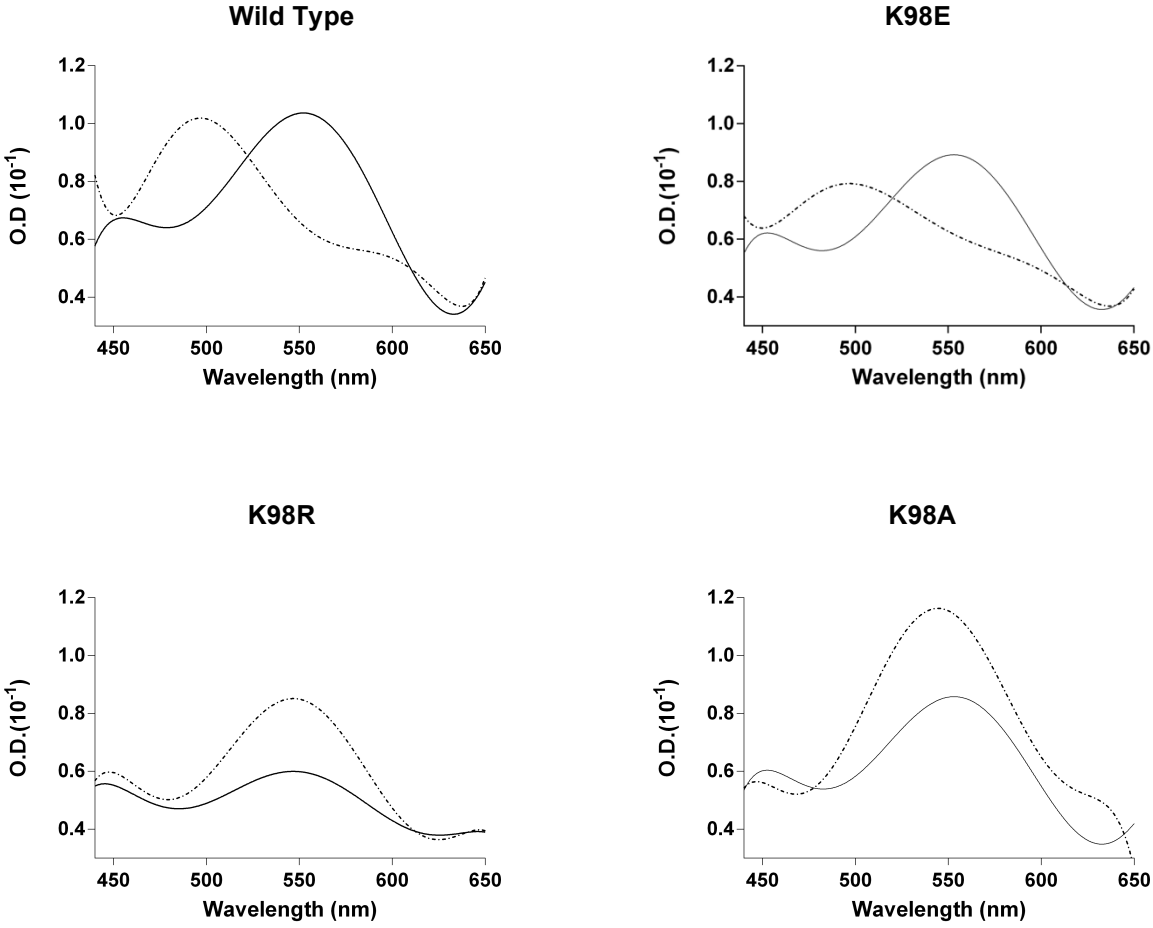


Figure 3.11 Spectral features of PBGD wild type and mutant proteins. Co-factor analysis showing the spectral shift from 564 nm to 495 nm following 20-minute incubation. The graphs are indicated as time zero T_0 indicated by the solid line, whereas time T_{20} is indicated by the (- - - - -) dotted line [80]. Graphs and images generated by GraphPad Prism, Version 6.01.

DPM Co-Factor Removal

In order to prepare the *'apo'* enzyme for the planned reconstitution assays to study DPM binding characteristics, wild type protein was hydrolysed with HCl in order to cleave the DPM from the enzyme (as discussed in Chapter 2). As shown in Figure 3.12, the DPM co-factor was successfully removed from the native wild type PBGD. This was proven by the negative reaction of the sample with the Ehrlich's reagent and no spectral shift was observed, thus behaving similarly to the K98R and K98A (as shown in Figure 3.11).

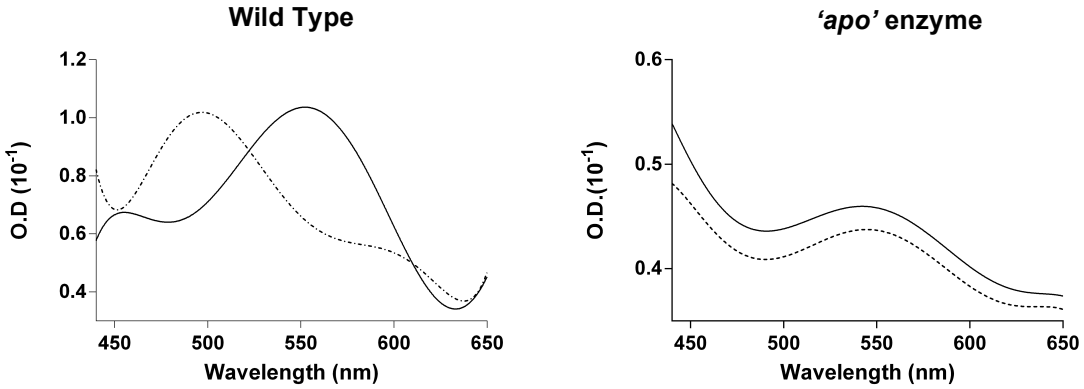


Figure 3.12 Spectral features of PBGD wild type and the *'apo'* enzyme. Co-factor analysis showing the spectral shift from 564 nm to 495 nm following 20-minute incubation. The graphs are indicated as time zero T_0 indicated by the solid line, whereas time T_{20} is indicated by the (- - - - -) dotted line [80]. Graphs and images generated by GraphPad Prism, Version 6.01.

Although we were able to successfully remove the co-factor, there were several difficulties with this protocol, including the stability of the protein for assays, which is discussed in detail in Chapter 4.

PBGD Reconstitution assay

DPM Solubility

In order to test the solubility of the DPM co-factor, 5 mM of DPM (in powder form, as purchased) was added to different ratios of methanol and imidazole. All samples were tested, and kept at 4°C overnight, to observe whether the samples would remain soluble, or precipitate. Imidazole was chosen as the solvent due to the similarity of the amine to that of the DPM co-factor (Chapter 1; Figure 1.8).

As illustrated in Table 3.6, the DPM was soluble in 50% v/v methanol and imidazole. Further reductions in methanol resulted in precipitation of the co-factor overnight. We therefore chose to use the 50% v/v solvent. With regards to the effect the methanol on protein, we tested 50% methanol with 0.1 mg/ml and 1 mg/ml concentrations of BSA, and found that it had no effect on the protein, and no precipitation of protein occurred.

Table 3.6 The range of solutions of imidazole and methanol in which the DPM co-factor solubility was tested.

100% methanol	70% methanol; 30% imidazole	50% methanol; 50% imidazole	30% methanol; 70% imidazole	100% imidazole
Soluble	Soluble	Soluble	Precipitates O/N	Not soluble

Assay

Table 3.7 illustrates the results for the reconstitution assay following the addition of either the PBG substrate or the combination of substrate with the DPM co-factor. This assay was based on the original assay we used for the PBGD activity (Chapter 2; Appendix 21). The maximum activity we were able to achieve following reconstitution was approximately 22% of the wild type.

Several alternatives and troubleshooting mechanisms were adopted in order to attempt a better 'recovery' of the enzyme following re-addition of the native enzyme substrate (and or co-factor), all of which did not significantly alter or improve our final results. A detailed description of alternatives and the troubleshooting is described in Chapter 4.

Table 3.7 The reconstitution of *'apo'*-enzyme with either substrate (PBG) only, or in combination with the DPM co-factor. Assay was performed at 4°C as well as 37°C.

	Enzyme Sample		V_{max} (pmol/ng/hr)	V_{max} % of WT
37°C	Wild Type	Substrate	1.87	-
		Substrate & Co Factor	1.87	
37°C	*'apo'*-enzyme	Substrate (4 hours)	0.4254	22,4
		Substrate & DPM Co-factor	0.4030	21,3
4°C	*'apo'*-enzyme	Substrate (4 hours)	0.0683	3.45
		Substrate & DPM Co-factor	0.0598	3.02

Circular Dichroism Analysis

Far-UV

Figure 3.13 illustrates the spectrum obtained for each of our respective samples. Each sample is represented as the mean of a minimum of 3 spectral scans that were prepared from independently purified protein samples. Data was normalized and converted to mean residue ellipticity as described in Chapter 2.

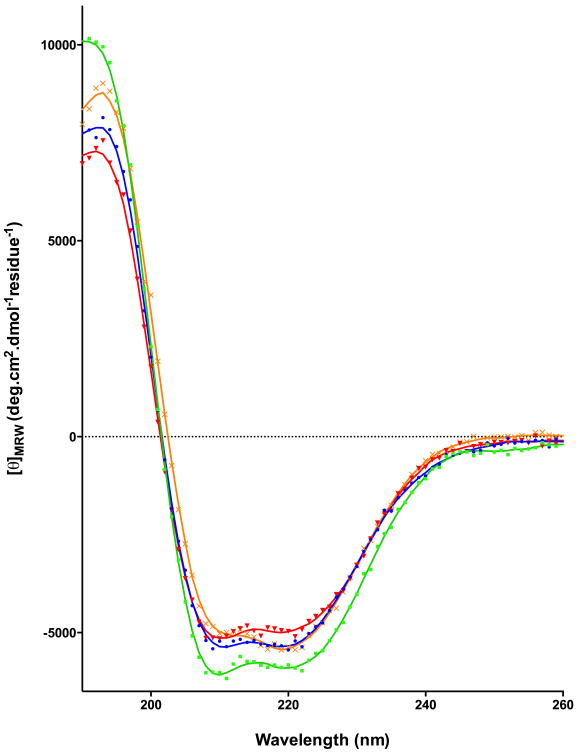


Figure 3.13 Illustration of the comparative far-UV analysis of wild type and mutants. Wild Type (green; ■), K98E (red; ▼), K98R (blue; ●) and K98A (orange; ×). Graphs were smoothed to a second order (polynomial) with a maximum of 4 neighbouring points and converted to mean residue ellipticity - $[\theta]_{MRW}$ ($deg.cm^2.dmol^{-1}residue^{-1}$). Graph images generated by GraphPad Prism, Version 6.01.

Data Analysis

The spectra analysis for protein secondary structure is within the range of 190 nm and 260 nm, with alpha helical proteins indicating negative peaks at 208 nm and 222 nm and a positive peak at 193 nm [100]. However, in order to ensure data collected was represented accurately in our graphic presentation, our spectra were collected between 190 nm to 300 nm. The raw data, in millidegrees (mdeg), were then minimised from the CD buffer blank, and normalized, by subtracting the average of the final readings (295 – 300 nm,) prior to data conversion to be expressed in mean residue ellipticity as shown in Figure 2.9. Important to note, is that this normalization does not in any way alter the results, but rather allows for the normalization of data points, resulting in a smoother curve, without the need to apply further smoothing techniques and exclusion of data points [108].

Each of the respective algorithms in the CDPro software has a maximum of 10 reference sets. For our samples we used reference sets SP37, SP43, SDP42, SDP48 and CLSTR. Our selection of protein reference was on grounds of consistent structure prediction [106,107]. Data displayed is the mean of the best five fits for each algorithm and the error expressed as the standard error of the mean (SEM) (Table 3.8). One-way ANOVA statistical tests revealed that the change in α-helical content of the K98E was significant ($p<0.01$) in comparison to the wild type. Although the β-sheet content for the K98R was higher for CONTINLL and SELCON3, these differences were insignificant in comparison to wild type.

Table 3.8 Secondary structure estimations performed with the CDPro software, using the CDSstr, CONTINLL and SELCON3 algorithms.

	α-helix (%)			**β-sheet (%)**		
	CDSstr	CONTINLL	SELCON3	CDSstr	CONTINLL	SELCON3
WT	20.8 ± 0.10	22.1 ± 0.20	21.6 ± 0.29	25.9 ± 0.47	25.8 ± 1.14	26.4 ± 0.48
K98E	16.7 ± 0.14	17.9 ± 0.35	17.8 ± 0.47	28.9 ± 0.47	29.9 ± 0.57	27.1 ± 1.7
K98R	19.7 ± 0.19	20.2 ± 0.23	21.0 ± 0.64	25.4 ± 1.06	38.1 ± 0.71	35.6 ± 0.66
K98A	20.3 ± 0.17	21.1 ± 0.19	21.6 ± 0.18	30.8 ± 0.37	28.3 ± 0.24	24.9 ± 0.09

In order to further correlate our findings with known wild type data, we calculated the mean of the wild type estimation data from CDPro and compared it to the two previously published crystal structures of human PBGD available in the PDB database (PDB ID(s): 3ECR and 3EQ1) [60,67]. We used two algorithms that are available via online servers, DSSP [114] and STRIDE [115,116]. These servers allow for the secondary structure prediction from crystal structure data coordinates. Data is outlined in Table 3.9 below. The 3EQ1 was not recognised in the STRIDE database, and therefore no estimations could be made.

Table 3.9 Secondary structure estimations performed from the two human crystal structures (3ECR and 3EQ1) with the online secondary structure algorithms, DSSP and STRIDE.

	WT	**3ECR**		**3EQ1**	
		DSSP	STRIDE	DSSP	STRIDE
α-helix (%)	21.5	28.7	30.0	32.0	-
β-sheet (%)	26.0	19.8	20.1	21.9	-

Although the data obtained for the wild type enzyme is comparable with those of the published crystal structures, the ratio of α-helices and β- sheets is different. However, this is most likely attributed to the fact that the wild type sample is a native soluble protein, whereas the crystal structures are 'static' representations of the enzyme. The final analysis for the far-UV data was to graphically illustrate all of the secondary structure predictions for the wild type and mutant proteins, as shown in Figure 3.14. The CDPro software does not allow for the prediction/estimation of the random coils and differences between parallel and antiparallel β-sheets [117]. The data is shown as percentage of the whole (100%), with the CDNN algorithm set to 'advanced spectra' in order to have the best and most accurate 'fit' of data estimation. The same data input that was used for each respective sample in the CDPro analysis was used for the CDNN estimation.

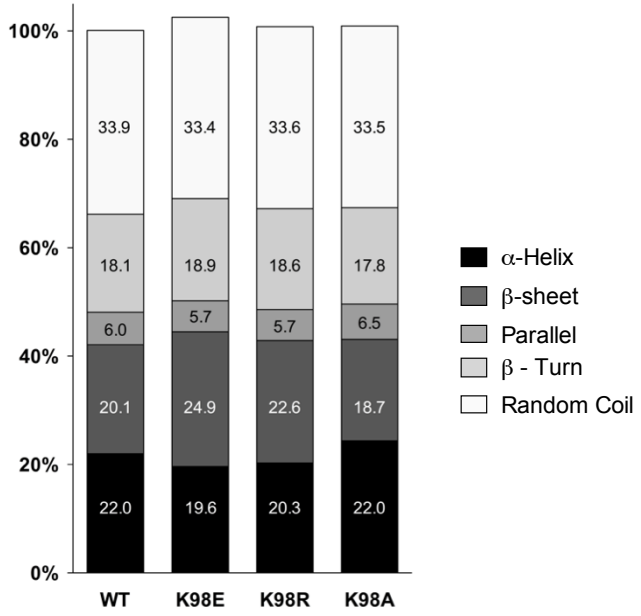
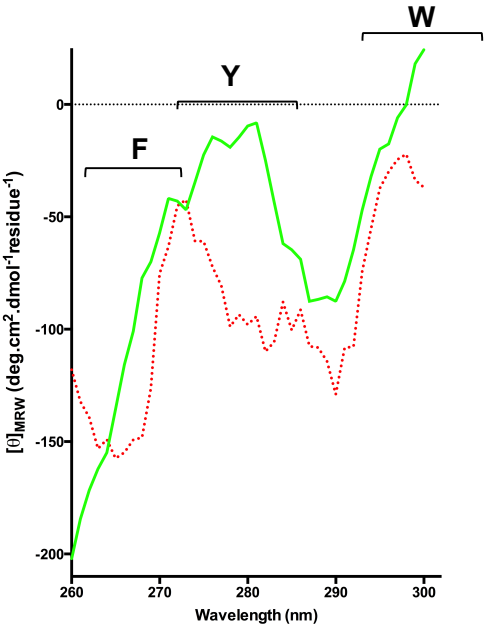


Figure 3.14 Illustration of the comparative far-UV analysis of wild type and mutants done by the CDNN algorithm in order to show the collective estimation of all structures in each respective sample [108]. Graph images generated by GraphPad Prism, Version 6.01.

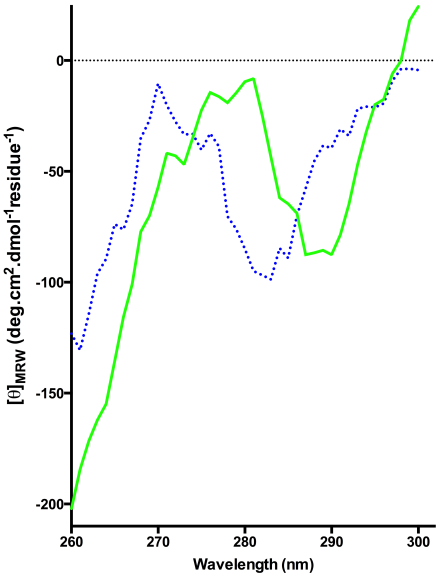
Near-UV

Figure 3.15 illustrates the near-UV spectra for each of the mutant samples (dotted lines) compared to wild type enzyme. Near-UV analysis is measured normally between 240 to 300 nm and indicates the presence of aromatic rings – specifically phenylalanine (F), tyrosine (Y) and tryptophan (W) [99]. It is said to be the 'tertiary footprint' of the protein structure, in that the presence of these peaks may indicate a tertiary structure, and hence correct folding of a protein [98]. The aromatic rings have characteristic peaks namely: between 255 and 275 nm for phenylalanine, tyrosine a peak between 275 and 282 nm, and tryptophan from 290 nm [116]. Graphs were smoothed to a second order (polynomial) with a maximum of three neighbouring points and expressed in mean residue ellipticity $[\theta]_{MRW}$ (deg.cm^2.dmol^{-1}residue^{-1}).

K98E



K98R



K98A

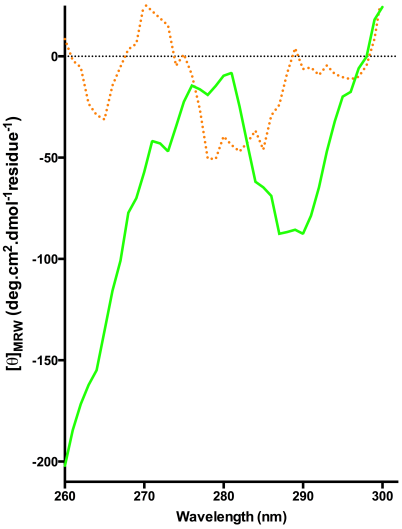


Figure 3.15 Comparative near-UV analysis of wild type and mutant PBGD. In the first graph of the K98E mutant, the peaks observed for phenylalanine (F) tyrosine (Y) and tryptophan (W) respectively are illustrated in the wild type spectra. Graph images generated by GraphPad Prism, Version 6.01.

CD Thermal Stability Analysis

Protein samples were prepared as for the far-UV and near-UV analysis as thermal assay is a thermodynamic application of the circular dichroism spectra for proteins. However, the graphical output for the thermal profiles may be analysed from the original spectra in millidegrees (mdeg). Figures 3.16 and 3.17 below illustrate the comparison of the thermal stability profiles of the wild type and mutant proteins. The change of ellipticity was observed at 208 nm **(I)** 222 nm **(II)** over a temperature range between 20 and 90 °C.

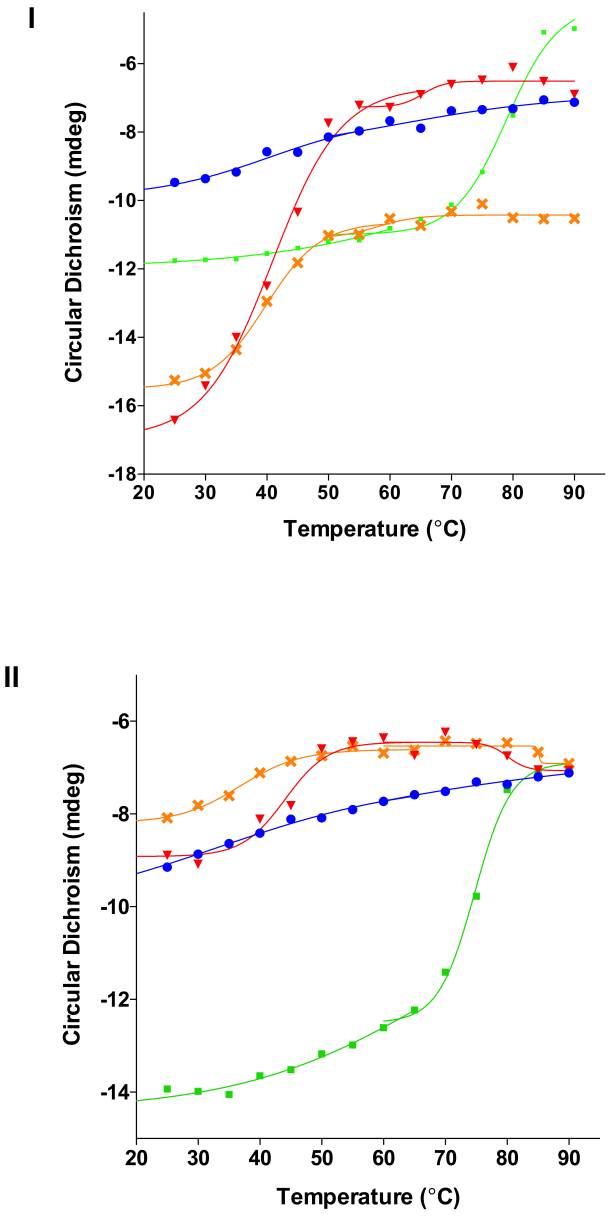


Figure 3.16 Comparative thermal analysis of wild type and mutants, measured at **(I)** 208 nm and **(II)** 222 nm. Wild Type (green; ■), K98E (red; ▼), K98R (blue; ●) and K98A (orange; ×). Graphs were fitted to a two-state unfolding model and data expressed as a non-linear regression (Boltzmann sigmoidal plot). Data is expressed in circular dichroism units, or millidegrees (mdeg). Graph images generated by GraphPad Prism, Version 6.01.

The melting Temperatures (T_m) for wild type and each mutant was determined from the fitted data, as described in Chapter 2, Figure 2.10.

Table 3.10 The melting temperatures calculated from the non-linear regression curves of the two-state unfolding modelling. Data was calculated from the non-linear regression of the thermal profile at 222nm (Figure 3.16 II).

	Wild Type	K98E	K98R	K98A
$T_{(m)}$	68.7 ± 0.41	44.1 ± 0.14	33.4 ± 0.87	36.8 ± 0.08

Data was fitted to a two-state unfolding model, as detailed in Chapter 2, and then converted to fraction of folded protein, as illustrated below in Figure 3.17.

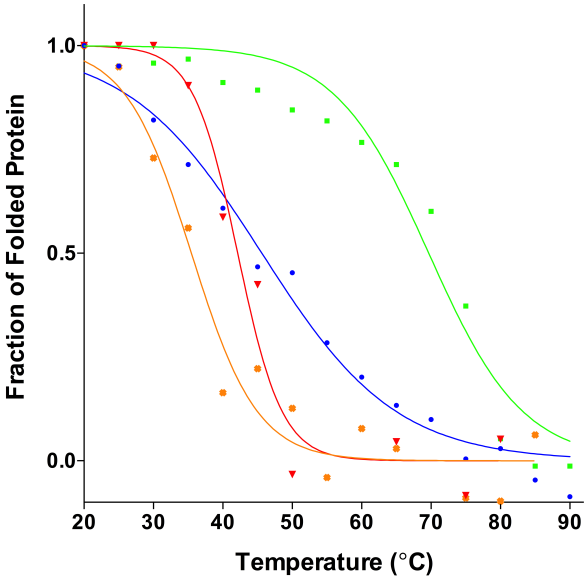


Figure 3.17 Fraction of folded protein as a function of temperature. Graph legends are as follows: Wild Type (green; ■), K98E (red; ▼), K98R (blue; ●) and K98A (orange; ×). Graphs were fitted to a two-state unfolding model. Graph images generated by GraphPad Prism, Version 6.01.

Structural Analysis

Following the structural analysis of the PBGD enzyme by UCSF Chimera, the interaction of the K98 residue with the DPM co-factor was visualized and a rationale established for the lack of co-factor binding in the case of the K98R and K98A mutants. Homology studies using the online bioinformatics server, Expasy was used, in order to observe the global effect on enzyme structure from each of the mutated sequences [109]. All Structures were minimised using Amber, and illustrations undertaken with USCF [61].

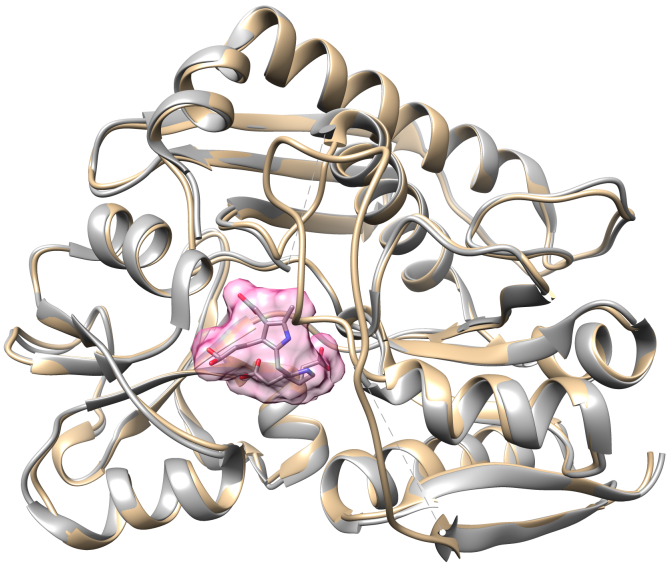


Figure 3.18 The superimposed structures of the modelled wild type enzyme from the CPH Model Server 3.0 (in tan colour) with the human crystal structure 3ECR (in grey).

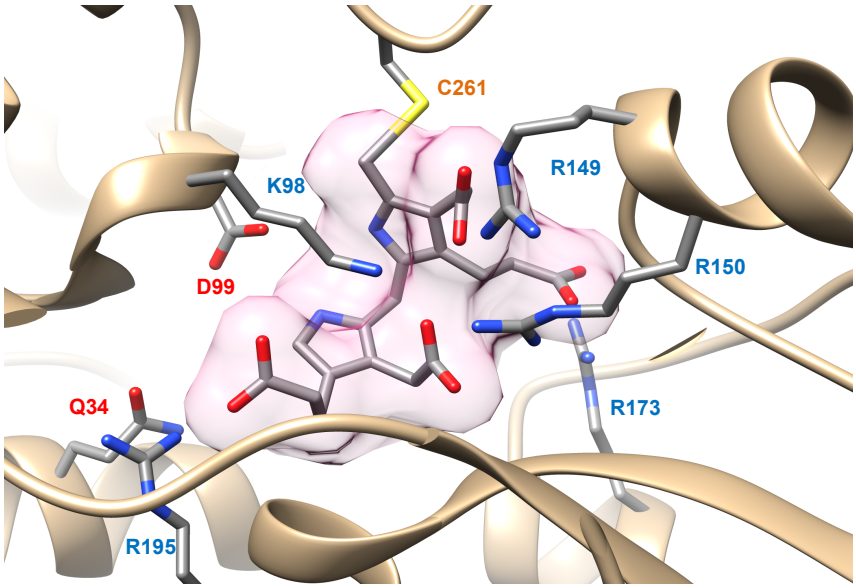


Figure 3.19 The active site of the PBGD enzyme.

Figure 3.20 illustrates the wild type enzyme with the interactions of the lysine (K98) residue with the carboxyl oxygen of the DPM co-factor. The close vicinity of the K98 residue suggests a stable ionic interaction with the co-factor. The modelled structures did not contain the DPM co-factor; hence it was modelled in from the minimised crystal structure (3ECR). This, however, does not allow for the bond formation of the co-factor to the cysteine 261 residue. The K98E mutant, as shown in Figure 3.21, although not able to form ionic interactions with the carboxyl groups of the co-factor, appears to maintain some interaction with the amine groups within the C1 and C2 of the co-factor.

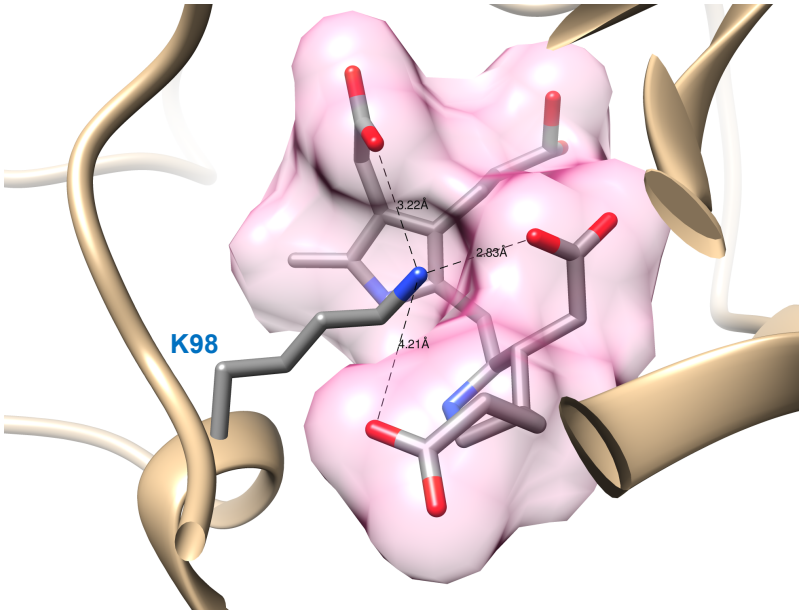


Figure 3.20 The illustration of the wild type, K98 residue and the interaction with the DPM co-factor.

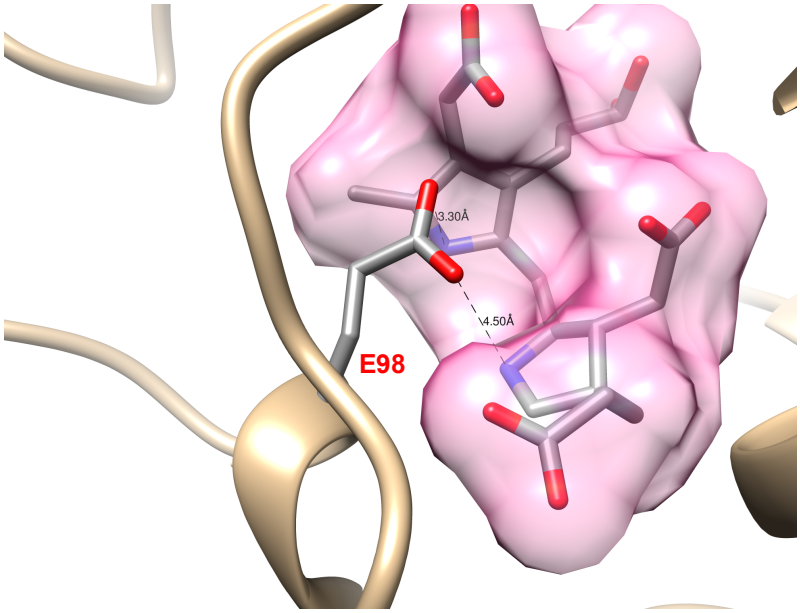


Figure 3.21 The mutated K98E residue and the suggested interactions with the DPM co-factor amine groups.

The K98R mutant, as shown in Figure 3.22, could possibly disrupt the architecture of the active site. The surface modelling shows how it interferes with the space occupancy of the DPM co-factor. The K98A, on the other hand, as shown in Figure 3.23 had no possible interaction with the co-factor, as it has no polar charge and therefore cannot form any stable ionic interactions. From Figure 3.23 it seems plausible that the K98A may, in fact, disrupt protein folding, as the lack of a side chain renders the active site with an open gap, which may result in increased solvent exposure, and hence increased protein instability.

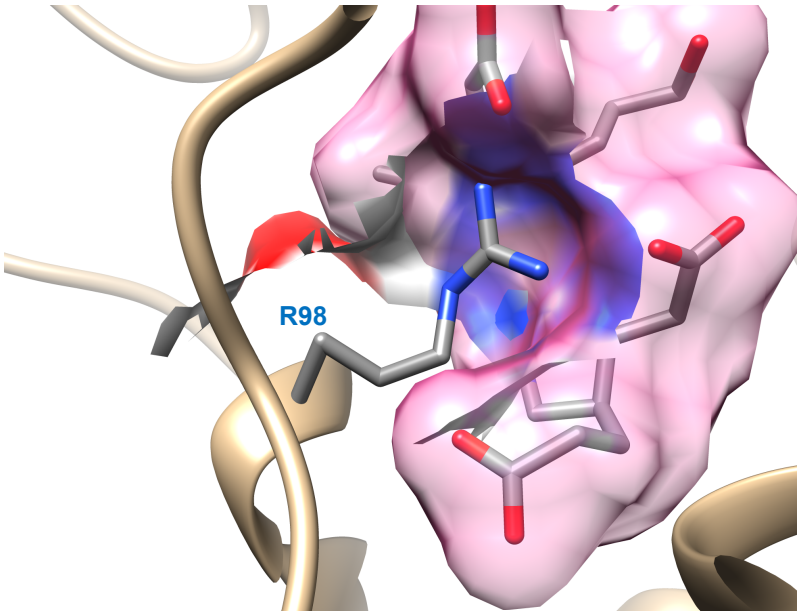


Figure 3.22 The mutated K98R residue with surface modelling, illustrating the possible disruption to the active site as it interferes with the DPM co-factor.

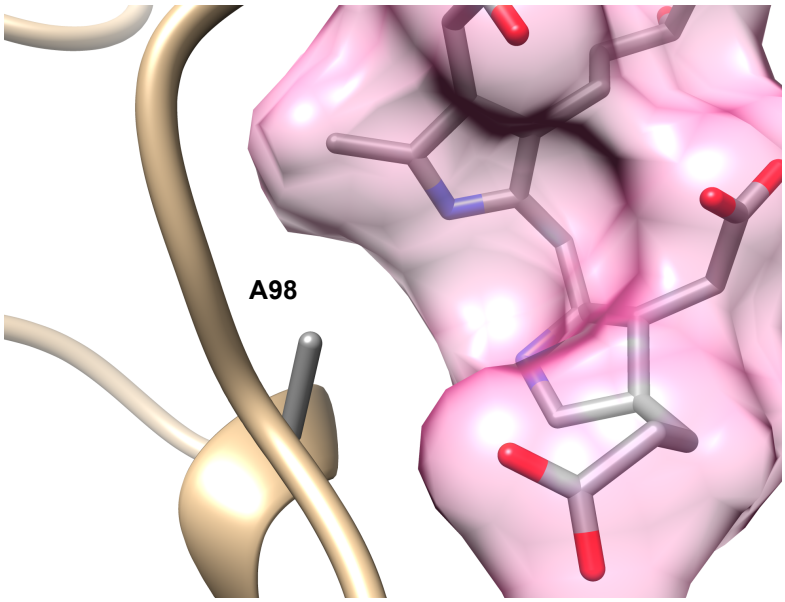


Figure 3.23 Mutated K98A residue with surface modelling, illustrating the lack of interaction of the alanine with the co-factor.

Homology modelling of the different mutants allowed for the speculation on observed changes in the overall structure of the enzyme. As shown in Figure 3.24 with the K98E mutant (indicated in red) superimposed on the wild type (in tan), the active site loop appeared to assume a different conformation to that of the wild type (indicated by green arrows). However, no other major differences were observed from the superimposed structures.

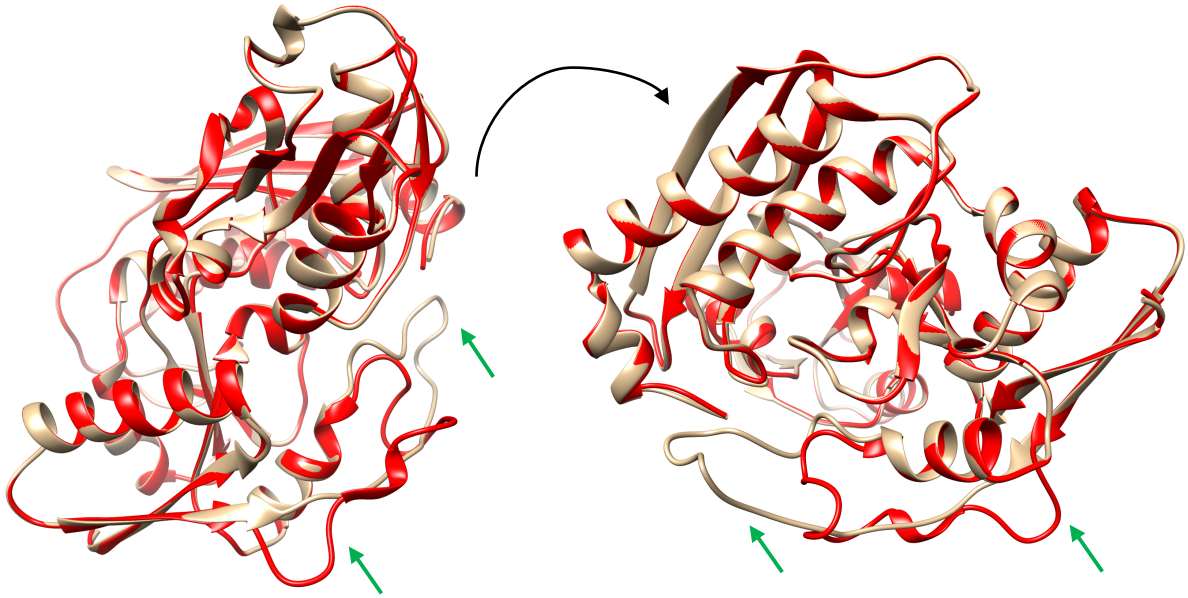


Figure 3.24 The superimposed structures of the wild type (tan) and K98E (red) mutant from the homology modelled structures using CPH Models Server 3.0 [110]. Green arrows highlight differences from the wild type and mutant.

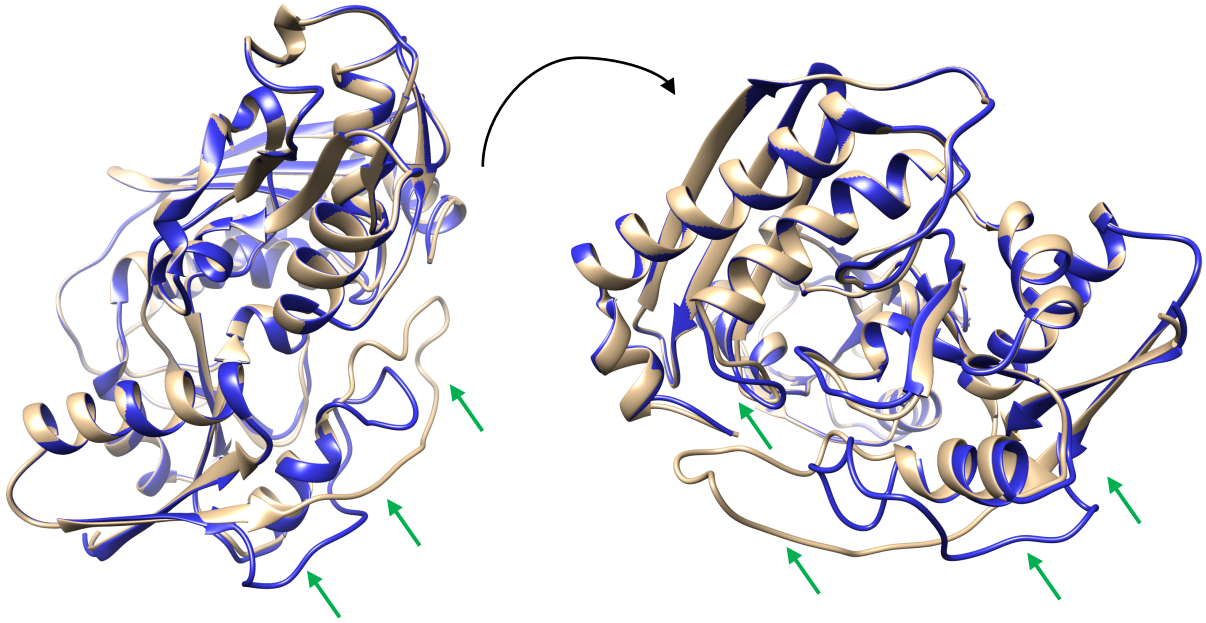


Figure 3.25 The superimposed structures of the wild type (tan) and K98R (blue) mutant from the homology modelled structures using CPH Models Server 3.0 [110]. Green arrows highlight differences from the wild type and mutant

In the K98R, as shown in Figure 3.25, the active site loop adopts a different conformation than the wild type enzyme, suggesting more of 'open' conformation, as discussed earlier in Chapter 1, regarding the movement of the domains. In addition, there were some movements noted within the inner loop regions of the enzyme, which may support the observations of this mutant's instability.

The K98A mutant, as shown in Figure 3.26, similar to the K98R mutant, showed some changes of the active site loop conformation. The active site loop appears to be further away from the other domains, again suggesting the 'open' conformation of the enzyme.

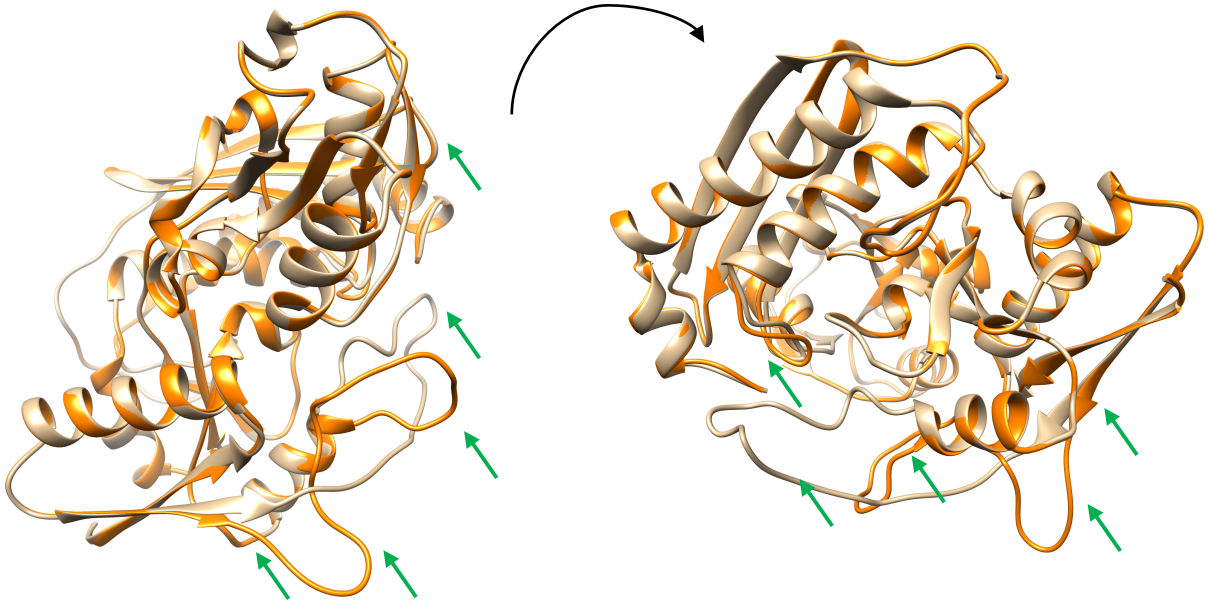


Figure 3.26 The superimposed structures of the wild type (tan) and K98A (orange) mutant from the homology modelled structures using CPH Models Server 3.0 [110]. Green arrows highlight differences from the wild type and mutant

Chapter 4

Discussion, Conclusion

Discussion

Overview

The purpose of this dissertation was to provide an overview of AIP focussing on the PBGD structure and function as well as experimentally demonstrating the use of kinetic, structural and thermodynamic techniques to shed light on the enzyme, with the key focus to investigate the effects the three K98 mutations have on conformation and hence enzyme activity. Two clinically relevant PBGD mutants, K98E and K98R, were expressed. Both of these mutants have previously been described in patients; the K98R was first identified in a patient in 1995 by Kauppinen et al., whereas the K98E mutation was identified by an MMed student in our laboratory in 2012 (reference to his MMed thesis https://open.uct.ac.za/handle/11427/6556) [1]. An additional mutant, K98A, was engineered and expressed in order to have a comparative view of the effect of charge at this residue. None of these mutations have been characterised in the literature prior to this study.

Site-Directed Mutagenesis

The adaptation of the QuickChange® SDM proved successful for the generation of the K98E and K98R recombinant mutant plasmids (Figure 3.2). The standard protocol was adapted for the engineering of the K98A plasmid, which required a two-base pair change from AAG for lysine to GCG for alanine. The PCR reaction for the site-directed mutagenesis was adjusted to increase the number of cycles in order to accommodate for this change.

Successful PCR products were obtained and expressed for the K98E and K98R mutants. There was initially some difficulty with the expression of the K98A mutant from overnight transfections. In order to address this, the antibiotic concentration was reduced as well as increased the volume of cells plated. This was previously done for PPOX mutant proteins in our laboratory and proved a successful alternative protocol [92,111]. Screening of selected colonies by direct sequencing, as shown in Figure 3.2 proved that all three of the desired mutations in the K98 residue had been introduced. As discussed in Chapter 2, further direct sequencing was also performed on the entire *HMBS* gene downstream of the K98 residue in order to ensure no other additional mutations had been introduced (data not shown).

Expression and Purification of PBGD

As this study included a kinetic and thermodynamic characterisation it was important that soluble, functionally assayable protein be expressed. Growth for 22 hours at 37ºC without IPTG induction followed by non-denaturing TALON™ metal affinity chromatography was satisfactory to yield sufficient amounts of wild type and K98E protein for purification and

subsequent characterisation (see Table 3.3). For the K98A and K98R mutants the incubation time was increased to 24 hours as the expression of these two mutant proteins was low. This prolonged the growth period and we were able to obtain a satisfactory amount of protein.

Imidazole based purification was used and this enabled the His-tagged PBGD to be competitively eluted off the column. The presence of high salt (NaCl) and TWEEN®-20 detergent in the column load reduced the levels of non-specific protein binding. However, this protocol proved successful only for the wild type and the K98E proteins.

Our visualization on the gradient SDS-PAGE electrophoresis showed that the eluted fractions for both the K98R and K98A contained several other protein fragments (Figure 3.4). This led us to speculate that 1) the expressed mutant protein was unstable and thus led to denaturation in cells or 2) due to the introduced mutation, there were several non-specific binding fractions that were interacting with the TALON™ resin. This non-specific binding is a common feature in recombinant expressed proteins [90,91].

Owing to the non-specific binding and impure fractions within the K98R and K98A mutant proteins, we altered the protocol following a review of previous work in our lab with mutant enzymes [92]. The TALON™ resin volume was first reduced from 1.6 ml to 1.2 ml, whilst adding 5mM imidazole to the sonication buffer. Reducing the resin volume allows for decreased amount of sample to bind (and saturate) the resin beads. In turn, it also then limits the amount of non-specific binding on the resin. As shown in Figure 3.5 we were able to significantly reduce the large non-specific fragments that were eluting with our desired K98A PBGD protein. However, there were still several small contaminating fragments in our samples. Thus we further altered the protocol as follows: The TALON™ resin was further reduced from 1.2 ml to 1.0 ml, and the imidazole concentration was further increased to 10mM in the sonication buffer. As seen in Figures 3.6 and 3.7, this alteration reduced the non-specific binding whilst not affecting the yield of our desired PBGD protein. The final adjustment in addition to those already mentioned, was to perform the entire purification at 4°C. This was done in order to promote the stability of the protein, and improved the purity of all our samples and was thus set as the new standard protocol for all purifications.

There were several other issues with the purification of the K98A mutant that delayed further characterisation. The K98A mutant was highly unstable, and following elution of the purified protein fractions, the pooled fractions precipitated after only eight hours which was insufficient time to perform the required assays. In order to address this, we kept eluted fractions separate and did not pool all eluted protein following purification. The smaller the volume of sample, as well as the lower the concentration, the less likely a protein will precipitate. At the same time,

we pre-incubated the glass collecting tubes on ice, ensuring the protein remains as close to 4°C as possible, at all times. This delayed precipitation to between 12 and 16 hours for the highest concentration protein fractions. Upon further observation it was noted that the protein fractions with concentrations of 1 mg/ml and less, did not precipitate. Thus we diluted the high protein fractions to 1 mg/ml with the normal elution buffer (containing the correct amount of PMSF) directly after purification and protein quantification. By doing so, we were able to prevent the precipitation of protein fractions for more than 72 hours, which was a sufficient time window to perform the required assays and analyses.

When analysing the quality of each purification with SDS PAGE electrophoresis, we utilised the Coomasssie®–R250 stain. Although Vacutec Aqua Stain product is advertised as being able to detect nanomolar quantities of protein, it provided satisfactorily results for the wild type and was a faster and more convenient approach, the non-specific binding fractions could not be observed with this method in the mutants. False assumptions regarding purity were thus a potential problem using the latter stain.

Specific Activity

Wild type PBGD protein and each of the generated mutants were tested for specific activity with the use of the modified assay by Anderson and Desnick [118]. The assay is based on the production of HMB from the PBG substrate, which is non-enzymatically converted to uroporphyrinogen, and oxidised to fluorescent uroporphyrin I (as detailed in Chapter 2, Figure 2.5). We initially had some issues pertaining to the solubility of the uroporhyrin I standard, which was not soluble in the Tris/HCl assay buffer, but following some additional research we were able to determine that it would be soluble in a stronger molarity of hydrochloric acid (HCl) [119]. The initial stock solution of 1mg/ml uroporphyrin I was made up in a 5N HCl solution in order to have the uroporphyrin I soluble, whereas the aliquoted standard solutions were in a 0.5N solution of HCl (as further detailed in Appendix 17).

Optimization of the assay had to be performed for each of the mutant PBGD proteins, and the protein concentration measured and taken into account during the conversion of fluorescent units to specific activity (Chapter 2). For the low expressing mutations, K98A and K98R, significantly higher protein concentrations were required in order to have accurate and meaningful fluorescent readings.

K98R is known to be an AIP-causing mutation associated with the phenotype of this autosomal dominant disease. This fits with our measurements which showed it had negligible activity and would result in approximately 50% activity in AIP subjects as such patients could be assumed

to have inherited the dysfunctional K98R allele plus a normal allele from the unaffected parent [26,57]. The K98E mutant showed 54% of the wild type PBGD specific activity. K98E has also been shown to be an AIP-causing mutation (reference to MMed thesis https://open.uct.ac.za/handle/11427/6556) although PBGD activity was not measured in patient blood samples. We could predict that the PBGD activity measured in such a patient's blood should reveal approximately 75% of normal activity. We were not able to follow up with such an investigation as fresh blood samples from this patient were unobtainable (reference to MMed thesis https://open.uct.ac.za/handle/11427/6556). With regards to the K98R and K98A mutant proteins, the specific activity was negligible for both, with results below 1% in relation to wild type (Table 3.4). It therefore appears that these mutations either disrupt the active site architecture or alternatively affect the protein folding, both of which render the enzyme inactive.

Kinetic Characterisation

For determining the kinetic constants of each of the respective samples, the specific activity assay was modified in order to ensure achieving substrate excess and the determination of the kinetic constants from the Michaelis-Menten fitted curves, as illustrated in Figure 3.8. All assays were performed a minimum of three times from independently purified samples in order to ensure reproducibility.

The Michaelis constant (K_M) determined for wild-type PBGD in this study agrees with that of Meissner (1990) (K_M of 8.7 µM), and the more recent work of Roberts et al. (2013) [68,120]. It also falls within the region of previously published PBGD K_M (4-20 µM), as reviewed for HMBS (EC 2.5.1.61) from the online enzyme database BRENDA (http://www.brenda-enzymes.info) [120,121]. With regards to the engineered mutants, the K98R had a similar K_M to the wild-type, suggesting that the affinity of the enzyme for its PBG substrate is unaffected by the mutation (Table 3.5; Figure 3.9). In contrast, the K98E mutant had an approximately three-fold increase of K_M in comparison to wild type showing a significant loss of substrate affinity (Table 3.5; Figure 3.9). A K_M for K98A could not be determined. We initially speculated that the K98A mutant would mimic the K98E by having a significantly larger K_M (in comparison to the wild type), indicative of a loss of substrate affinity. However, during the activity and kinetic assays, the total enzyme activity was almost negligible, and even after several attempts and alterations to the protocol, the K_M had an unrealistic variation in values, either very low numbers, (close to zero) or values as high as 100 µM. In other repeated attempts, data was scattered and did not converge, hence no kinetic constants could be determined. One possible reason for this might be due to the effect of small variations on almost negligible activity. In other words, because the fluorescent readings are within the lowest reading frame (at highest sensitivity

settings on the fluorimeter), a small amount of variation results in drastic scattering of activity reflected in the final calculation [122,123]. Although it may well be that the K98A mutant has a weak affinity towards the PBG substrate, the instability of the protein together with its negligible activity makes it impossible to accurately determine the K_M.

In terms of the enzyme's catalytic efficiency, in other words the rate of substrate converted to product, taking into account both enzyme and substrate concentrations, expressed as the ratio of k_{cat}/K_M, it may be noted that the K98E as well as the K98R have a four-fold and eight-fold (respectively) lower enzyme efficiency compared to the wild type (Table 5.5). This implies that although the specific activity of the K98E is only half that of the wild-type PBGD enzyme, the loss of the substrate affinity (K_M) consequently results in the enzyme turnover rate to be practically negligible. Similarly, in the case of the K98R, even though the substrate is still able to bind, the negligible activity results in an equally poor enzyme catalytic function. These findings support the fact that the K98E and K98R are known disease causing mutations (https://open.uct.ac.za/handle/11427/6556) [117]. Since we were not able to determine a stable K_M for the K98A, we were consequently unable to determine the catalytic efficiency for this mutation. However, based on the close resemblance of this mutant's activity to that of the K98R, together with its significant instability, it may be suggested that that this mutant has almost no enzymatic function.

Native PAGE Analysis

The native PAGE is often used as a technique to view and profile the different enzyme complexes of the PBGD enzyme [124]. As highlighted previously in chapter 2, the stability of the enzyme during the polymerization steps, allows for the viewing of each respective enzyme-substrate complex. We anticipated that the analysis of the accumulation of different enzyme complexes may allow for some insight into how a specific mutation may affect the enzyme function in either the inability of the enzyme to polymerize its substrate (accumulation of the first two complexes) or difficulties pertaining to the release of the product (accumulation of the final enzyme-substrate complex ES_4) (Figure 1.11).

The initial investigation of the native PAGE gels showed mostly smearing of bands, or alternatively no product was seen. We attempted several different protocols without success. Finally, following contact and discussions with Ms Helene J. Bustad, from the University of Bergen in Norway, who published an article with relation to similar work (in Norwegian mutations) as we have done for the K98, including native PAGE gels, we achieved some results by using her protocol [96]. We also adapted the system in order to perform these native PAGE gels in the smaller and Bio-Rad® electrophoresis systems which allowed for reduced

running time of the gel. Finally, following the successful visualization of the wild type ES complexes (Figure 3.10), we adapted to run our mutant samples at 4°C, as the stability was a continuous issue. However, as illustrated in Figure 3.10, all of our mutants proved too unstable for this native PAGE analysis and showed mostly band 'smearing'. The K98E showed a faint band, higher up on the gel in comparison to the ES complex of the wild type, which has been previously suggested to be a partially formed holoenzyme [96,124]. In other words, the K98E mutant maintains interaction of the co-factor (holoenzyme), but there are some issues pertaining to the protein folding. This is particularly interesting as this might explain why the K98E mutant is able to retain 54% of its activity in comparison to the wild type, but at the same time show a significant difference in secondary structure and a lack of thermal/conformational stability.

It would have have been interesting to expand on these native PAGE gels with the addition of the PBG substrate, in order to see how each mutation would respond, and thereby promote the understanding of how each mutation affects the enzyme mechanism. However, considering that no enzyme complex bands could be observed for any of our mutants, it was not possible to do so. In addition, it has been previously reported that the PBGD enzyme is destabilised upon substrate binding, thereby further promoting destabilisation to already highly unstable mutants would not yield positive results [75,96]. We did, however, attempt a native PAGE with the addition of PBG to each respective mutant prior to the running of the native PAGE gel, but no changes were observed.

DPM Co-Factor Analysis

During the first attempts to purify the K98R mutant, it was noted that the eluted mutant protein lacked the characteristic pink colour (due to presence of a pyrrolic chromophore of the DPM). This led us to speculate (in subsequent repeated purifications with the same outcome) that there could possibly be a loss of the DPM co-factor in this mutant [71]. Following subsequent testing of the wild type PBGD as well as our mutants using the original assay performed by Hart et al. (1988) in their work on the DPM co-factor, we were able to show that the K98R had lost its ability to bind the DPM co-factor, since there was no spectral shift in the absorbance, which is a key feature of PBGD with bound vs. unbound DPM (Figure 3.11) [78]. On the other hand K98E was able to retain co-factor binding ability (Figure 3.11) as demonstrated by retention of the spectral shift. Based on these results, we speculated that the K98A mutant would have the combined effects noted in each of these mutants, namely a loss of substrate affinity (K_M) as well as a loss of co-factor interaction. Indeed the K98A mutant showed no interaction with the enzyme's native DPM co-factor but we could not ascertain substrate binding affinity as outlined in the previous section.

As detailed in Chapter 1, the DPM co-factor is a striking and significant feature of functional PBGD enzyme. Among other residues within the vicinity of the enzyme active site, the K98 has been suggested to be important for the interaction and maintenance of the co-factor orientation. However, no evidence has been put forward to prove this. Thus with our K98 mutants, we agree with this speculation, and have provided results to support that the K98 is indeed a key residue for co-factor interaction as mutation(s) of the residue results in failure of the enzyme to bind its co-factor. In contrast, the PBGD mutation in the glutamate residue (E) maintains the co-factor binding. A possible reason for this is that the negative glutamate, although no longer able to form ionic interactions with the carboxyl groups of the co-factor, may be able to form ionic interactions with the amine rings of the co-factor, thereby stabilising the enzyme and maintaining a stable co-factor interaction. In the case of K98R it could be speculated that the bulky arginine (R) side chain disrupts the orientation and architecture within the active site cleft, and consequently disrupts the niche of specific interactions of the co-factor. Alternatively, perhaps the arginine interferes with the orientation of the D99 residue, as previously discussed to be a key residue for enzyme catalysis, thereby destabilizing the ionic interactions of the enzyme with the DPM co-factor on more than one residue simultaneously. On the other hand, the K98A which lacks a side chain as well as polar charge, results in an opening or gap within the active site, thus increasing the solvent exposure of the active site. This could be one possible reason why this mutant is particularly unstable.

DPM Co-Factor Removal

One of our aims was to investigate first the removal of the DPM co-factor, followed by re-addition of the DPM co-factor and or PBG (substrate). The method of Hart et al. (1988) has been cited several times in the literature as a means of the most reproducible and successful method for the removal of the DPM- co-factor from native wild type enzyme [78]. An alternative methodology for 'removal' of the co-factor was to not generate the PBG required for co-factor formation and insertion in the first place. Such involves the generation of a vector that mutates the *hemA* gene, and consequently ALA synthase, which produces the PBG substrate, is not produced. However, this methodology was beyond the scope of this work, and assumed the first-mentioned method was viable [58,80,125].

Initially there was a significant number of problems with the methodology from Hart et al. (1988), and we attempted several optimization strategies, including the use of alternate buffers to the sodium phosphate buffers used by them. Tris/HCl buffers were chosen in order to accommodate the required pH of 8.0 versus 7.2 as used by Hart et al. (1988) [78]. Further, following increased dialysis buffer exchanges, allowed for successful removal of the co-factor

from the enzyme, as indicated from the results, with the lack of a spectral shift noted in the treated *'apo'* enzyme sample (Figure 3.11).

The presumed *'apo'* enzyme sample was very unstable, and had a specific activity below 2%. Our results were in agreement with the published results from Hart et al. (1988) as we were able to recover a similar amount of protein following dialysis; they stated a recovery of 36% of the original protein concentration and in our results we were able to achieve 34% from the original protein [78]. This, of course, made it increasingly difficult to obtain sufficient quantities of protein for our analysis, so we overcame this by hydrolysing multiple samples simultaneously. Because of the instability of the protein, we kept all fractions separate (similar to that of K98R and K98A) rather than pooling the samples, in order to prevent protein precipitation. The timing and planning around the acid hydrolysis and dialysis of the so called *'apo'* enzyme was crucial, as the protein precipitated completely overnight.

PBGD Reconstitution Assay

As highlighted earlier, our aim was to test whether we would be able to 'rescue' the *'apo'* enzyme from which co-factor had been removed, by adding back DPM co-factor. The rationale was as follows: As suggested in Chapter 1, it has been shown that there are two ways in which the PBGD enzyme is able to obtain its DPM co-factor, the one involving reaction with the use of the product, HMB (or preuroporhyrinogen) or the alternative slower reaction with the use of the monopyrrolic PBG substrate. Thus we evaluated what effect it would have if, instead of the preuorporphhyrinogen product, the DPM co-factor *per se* was pre-incubated with the enzyme. The further downstream experiments, if this first experiment had proved successful (i.e. in which the wild type enzyme had the co-factor successfully removed and reconstituted), would have been to perform similar assays with the K98R and K98A, in order to evaluate if indeed the enzyme would be able to take up and bind the DPM co-factor and thus promote the stability of the enzyme as a whole. In addition, it would have been interesting to evaluate and produce the kinetic parameters involved in the binding of the co-factor to the enzyme.

As detailed in Chapter 3, we were only able to achieve 22.4% activity of the '*apo'* enzyme in comparison to wild type after addition of/pre-incubation with the DPM co-factor, PBG substrate or both (Table 3.7). There were no differences between the addition of the substrate only versus the pre-incubation of the enzyme with the DPM co-factor (Table 3.7). Thus there is no evidence to suggest that the enzyme shows any preference for the DPM co-factor over the PBG substrate under these experimental conditions. Hence the use/binding of the dipyrrolic form of the co-factor appears no more efficient/preferred than the monopyrrolic PBG. In addition, there is no additive/synergistic effect seen by the co-factor and the substrate.

We performed several optimization strategies and alterations to the protocol in order to test whether our reconstituted activities could possibly be improved. For example, the solubility of the DPM co-factor was also tested and adapted from the prescribed solvent of pure methanol, which we felt could have been problematic to the refolding/reconstitution of the protein. We were able to dissolve the DPM in a 50% v/v of methanol and imidazole solution. This solution did not have any effect on the native wild type enzyme (Table 3.7). We also used a concentrated DPM solution (0.5 mM), in order to maintain the volume of the methanol-imidazole solution added in each assay sample, to a minimum (Appendix 21). Further, the assay time was increased to 4 hours to ensure the co-factor reconstitution reaction could run to completion, keeping in mind the slow 'lag' phase suggested for the *'apo'* enzyme when utilising either PBG and/or possibly the DPM co-factor. The additional assay time was also allowed to ensure that our fluorescence readings were within an accurate range. We also ran an additional assay at 4°C, in parallel to the normal assay at 37°C, in order to test whether the higher temperature results in spontaneous oxidation of the co-factor which would be problematic. As detailed in Chapter 1, the DPM co-factor is quite prone to oxidation, which renders the enzyme inactive [70]. Indeed, we noted that in some of our samples the solution turned bright yellow following completion of the assay. However, the assays performed at 4°C showed negligible activity restoration after reconstitution with DPM (and/or PBG), even following the prolonged incubation time of the assay. We performed additional assays for a 6-hour period, but no improvement of the results was observed.

It could be argued that either the enzyme is not completely functional at 4°C, as most assays for PBGD are performed at physiological temperature (37°C) [40,126]. Alternatively, it could be that the higher temperature does promote some oxidation of the DPM co-factor, thereby making it non-viable for the enzyme to utilise. As noted in Chapter 1, the oxidised co-factor is observed to have a different conformation in comparison to the reduced form. This change in conformation may lead to a reduced reactivity of the DPM co-factor, which might be a possible reason for the enzyme not to favour the DPM co-factor versus the PBG substrate. As quoted from Awan et al. (1997) in their work with the preference of the PBGD enzyme to preuroporphyrinogen:

"The answer undoubtedly lies in the fact that the 1-hydroxymethyl group of preuroporphyrinogen is more reactive than the aminomethyl group of porphobilinogen and more readily forms a azafulvene that can rapidly alkylate cysteine-361 to form the thioether link..." [73].

An alternative scenario that could possibly explain the low recovered activity is that the PBGD enzyme has previously been shown to be destabilised upon PBG substrate binding. This correlates with the suggestion that the enzyme is able to 'open' and 'close' in order to accommodate the growing pyrrole product (as discussed in Chapter 1) [75,96]. Thus, considering the unstable nature of the *'apo'* enzyme, the binding of the PBG causing additional destabilisation may lead to an enzyme that is not entirely functional. In contrast to all of these scenarios it is also possible to speculate that the enzyme was not correctly refolded during the dialysis and was consequently dysfunctional to start off with. Further work to establish this was beyond the scope of this study within funded and required time period – although an answer could be sought by a host of different experiments to investigate protein folding and aggregation, including intrinsic and extrinsic fluorescence, CD and NMR [127].

CD Far – UV Analysis

As noted in Chapter 1, the K98 residue has been suggested to be important for enzyme catalysis. Based on our findings of the K98R and K98A whereby the interaction with the DPM co-factor is lost, we wanted to investigate whether this loss possibly correlates with alterations in the enzyme secondary structure (i.e. the general three dimensional form of local segments of the PBGD protein). Our rationale was that although the K98 has been suggested to be important in enzyme catalysis, we may be able to further define its role by establishing whether or not it was important for conformational stability of the protein.

The methodology for our CD far-UV analysis was based on the recent work of Bustad et al. (2013) in other PBGD mutants [96]. The only modification to their methods was the use of ammonium sulphate buffer instead of potassium fluoride (KF). Ammonium sulphate has the same lower wavelength limit as potassium fluoride (185 nm) and was within our range of spectra measurements [98]. In addition all of our eluted protein samples required buffer exchanges, as the high content of imidazole in the elution buffer is optically active and would therefore adversely affect the spectral data [98]. All of our samples were surprisingly stable in the exchanged CD buffer (data not shown). The buffer exchange results in approximately a one and half times dilution, thus we had to perform a protein quantification after the samples were exchanged into the CD buffer. This is particularly important when performing repeat experiments, and also since we are comparing the ellipticity amongst all of our mutants. Changes in concentration might result in changes of data recorded and thus lead to false results and/or artefacts and incorrect conclusions.

Once CD spectral data was collected we analysed the data for our different mutants. We utilised CDPro software which has three different algorithms namely, the CDSstr, CONTINLL

and SELCON 3 [100–104]. CDPro is a robust program since it allows for the analysis of three different algorithms which can be referenced against a dataset of ten proteins. Importantly, although the data was normalized for the graphical display of the far-UV analysis, we used only the raw data which had been minimised against the blank. The CDPro algorithms only read spectra within the range of 180 nm to 240 nm and therefore data was only collected between 190 nm and 240 nm. We did not include spectra below 190 nm as the buffer absorbance increases exponentially from this point and it is generally advised to keep the absorbance below 2.0 absorbance units.

As shown in Table 3.8, our results indicate that there are modest differences in secondary structure estimates between the WT and mutant PBGD. It is important to note that these structure estimations are done based on a reference set of proteins for each algorithm. The data may be considered as valid estimations if there is convergence amongst the estimations of each algorithm. From the data we can observe that there is a decreased α-helical content predicted for the K98E and K98R in comparison to wild type. However, the difference is only significant for the α-helical prediction of K98E, as noted in Chapter 3 (Table 3.8). This is particularly interesting as it suggests that the 'non-conservative' change from the positive lysine to the negative glutamate, does hinder the secondary α-helical structure. In terms of the β-sheet estimation, although the K98R and K98A appeared to have increased estimations (as noted in Table 3.8) the changes were not significant. This finding is interesting as the loss of the co-factor we found for both these mutants, there was an expectation that some secondary changes would be observed.

In addition to the analysis with the CDPro software, we also performed two other analyses of predictions for the PBGD enzyme secondary structure. These included the predictions based from the crystal structures PBGD enzyme, as well as another software package CDNN [106,112,114]. The comparison of the wild type to the known human crystal structures was done in order to determine if our results are within the same range as for the crystal structure. As shown in Table 3.9, we compared the mean secondary structure prediction of the wild type to the two human PBGD crystal structures (3ECR and 3EQ1) [60,67]. The two algorithms, DSSP and STRIDE are able to integrate the coordination points of the crystal structure and provide a prediction of the enzyme secondary structure. Although our results are within the same range as the predictions of DSSP and STRIDE, it appears that the ratio of α-helices and β-sheets differ slightly. However, our results are in agreement with Bustad et al. (2013), who also found a higher β-sheet and lower α-helix prediction [96]. It could be argued that crystal structures, although providing some insight into the possible structure, will naturally vary from a native functional enzyme.

To further investigate the secondary structure predications, we analysed the data using CDNN. Although this program does not have multiple algorithms, it screens the data against protein reference sets and provides a global picture of all of the possible structure predictions. Indeed as noted in Figure 3.14, the CDNN software allows for the differentiation of β-sheets and parallel sheets. Interestingly, here we are able to observe that the K98E shows a variation of its secondary structure, with a decrease in predicted α-helical content, and an increase in β-sheets. The K98A, in contrast seems to have a similar α-helical content compared to the wild type, whereas its β-sheet content is decreased. This might provide an explanation for the instability of the K98A PBGD [128].

CD Near – UV analysis

Whereas changes in the far-UV CD will usually reflect major backbone changes in the protein, small conformational changes are more likely to be detected in a near-UV analysis since the CD contributions of the aromatic side chains is usually very sensitive to their environment [129,130]. Thus, the near-UV analysis (260 nm – 300 nm) was carried out in order to investigate whether the mutation of K98 affected the spectra of the aromatic amino acids in PBGD (phenylalanine, tyrosine and tryptophan) [115,131]. Although this does not provide an in-depth structural insight of the enzyme as a whole, the 'tertiary footprint' of these amino acids does allow for a comparison of the wild type to each of the respective mutants [97].

For the near-UV analysis, the data was treated the same as for the far-UV. The final graphic output was then adjusted in order to only include the spectra for the wavelength range from 260 nm to 300 nm. From Figure 3.15 illustrating the peaks observed for phenylalanine, tyrosine and tryptophan there is an indication that the aromatic rings, specifically tyrosine, is disrupted in each of the mutants. As noted in Chapter 3, all of the mutants illustrate a 'blue-shift' towards the lower wavelength of the tyrosine movement and possible increased exposure to solvent. This suggests that the alteration of the K98 residue does indeed disrupt some of the active site architecture and may alter the conformation of the enzyme. The K98E however seems to retain the intensity of the ellipticity signal, whereas the K98R and especially the K98A illustrates an intensity shift towards zero, indicating the 'quenching' of the signal. This may further support the idea of the loss of co-factor in these two mutants as the quenching is as a result of exposure to either surrounding water, or alternative polar environments [130,132].

Thermal Stability Analysis

The PBGD enzyme is known to be thermally stable and therefore does not readily unfold with increases in temperature [68,86]. However considering the effect of the K98R and K98A

mutation on the enzyme's stability, we wanted to further characterise the effect of these mutations by investigating their effects on the enzyme's thermal stability.

The PBGD mutations caused alterations in the spectral profiles over a temperature range of 20 – 90 °C (Figure 3.16). The wild type illustrates a high conformational stability, with a melting temperature ($T_{(m)}$) of 68.7°C, which correlates with previous studies on PBGD [68,72]. The K98E, which is the most stable of all the mutants in terms of protein activity, however shows a similar sensitivity to thermal denaturation as the K98R and K98A (Table 3.10). The signal at 222 nm for all of the mutants is significantly reduced in comparison to wild type, which again highlights the instability of the mutants. The K98R, unlike the K98E and K98A, seems to have a more linear (or flatter) regression. This is also seen in Figure 3.17, where the graphs were fitted in order to express data as a fraction of folded protein. The K98A unfolds very rapidly, followed by the K98E. The K98R, again illustrates a slower transition from folded to unfolded protein. Attempts to fit multiple sets of data from different, independent purifications illustrated the same trend. This may highlight the arginine residue (R) itself, indicating that although the K98A also has negligible activity, the arginine may be disrupting additional residue networks and interactions.

Structural Analysis

A structural analysis was undertaken to allow for further visualization of the possible effect of each mutation in terms of the enzyme structure. We chose the human crystal structure (PDB ID: 3ECR) published from Song et al. (2009) [67]. The rationale being that the alternative human model from Gill et al (2009) was crystallised with the R167Q mutation [60]. Bung et al. (2014) used the *E.coli* protein in their work using molecular dynamics – as it is the best characterised [75]. However, we chose to maintain the analysis in the human, un-mutated enzyme. An online server, hosted by the global database and bioinformatics tool, Expasy was used for the homology modelling of the PBGD mutants [109]. CPHModels 3.0 Server allowed for the simple uploading of the sequences (in FASTA) format, followed by the prediction model based on all known crystal structures of the sequence.

Figure 3.18 illustrates the superimposed structure of our wild type modelled query to the known human crystal 3ECR. As detailed in Chapter 3, there is a clear and direct overlay of the enzyme, and the missing active site loop region was successfully modelled. This gave us a good indication that the homology modelling proved successful and we could continue with further analysis. Importantly, each of the structures were independently minimised following the modelling, as well as an energy minimisation of 3ECR performed prior to superimposing of structures.

With regards to the global structural analysis of the enzyme, from Figures 3.24 – 3.26 in Chapter 3, it is evident that there are some changes of the active site loop conformation. What is very interesting, and somewhat surprising, is that each of the respective K98 mutations seems to affect the active site loop conformation in a different way. In Figure 3.24 it can be seen that the K98E mutation appears to change the conformation of the active site loop. It assumes a more 'closed' conformation as it moves towards the other domains in comparison to wild type. This may be an effect of the negative glutamate charge. The active site cleft as highlighted in Chapter 1, has predominant positively charged residues. So it could be argued that the negative glutamate, although it does allow for some means of stabilisation from the interaction with the DPM co-factor, as is suggested in Figure 3.21, may cause destabilisation of surrounding residues within the positive active site. In addition this might also explain the significant reduction in substrate affinity, as the active site loop adapts a more closed conformation which may reduce the availability of the enzyme to its PBG substrate. Furthermore, the altered conformation of the loop may also correlate to the significantly reduced α-helical content noted in the far-UV analysis for the K98E (Chapter 3, Table 3.8).

With regards to the K98R (Figure 3.25) and the K98A (Figure 3.26) mutants, the opposite effect is seen. The active site loop resembles more of an 'open' conformation, which is normally adapted for the enzyme with the release of the product, as shown by Bung et al. (2013) (and previously discussed in Chapter 1) [75]. The surface modelling of the bulky arginine, as shown in Figure 3.22, indicates that there is a possible steric hindrance with the DPM co-factor that might result in a change in conformation or positioning of the co-factor. In doing so it may however promote the oxidation of the co-factor, as suggested by Azim et al. (2014) (and discussed in Chapter 1, Figure 1.10). [70]. A similar picture may be seen from the K98A mutation. The active site loop is seen to move away from the other domains (in comparison to the wild type), again suggesting the 'open' conformation of the enzyme. In contrast to the K98R, the alanine has no charge, as seen in Figure 3.23, which in turn may further destabilize the enzyme, as the lack of a side chain may create an additional open area within the active site. This opening may render the enzyme open to increased solvent exposure, which may disrupt the conformation of the enzyme as a whole.

Conclusions

The purpose of this study was to investigate how a range of mutations within the active site (at the K98 residue) of PBGD affects the enzyme mechanism and conformation. In general terms, our results support the idea that expansion of our knowledge of a single active site residue may provide important insight into the structure and function of the PBGD enzyme as a whole.

Site directed mutagenesis and expression of the K98E, K98R and K98A PBGD variants were successfully carried out. Furthermore, purification using affinity chromatography was optimised in order to obtain satisfactory pure and soluble protein for kinetic and spectroscopic characterisation. Compared to wild-type PBGD, K98R and K98A had negligible specific activity and the K98E had approximately 50% of wild-type specific activity. In K98R and K98E mutants the catalytic efficiency was significantly reduced. K98R had a similar K_M to that of wild-type, but K98E showed reduced substrate binding affinity. These kinetic parameters could not be determined in K98A due to protein stability issues. The significance of the positive charge of the K98 residue on substrate affinity was illustrated by a significant loss of substrate affinity (higher K_M) in the negatively charged K98E mutant, whereas the conservative mutation, K98R, had a similar K_M to that of the wild type enzyme. However, the fact that the K98E mutant was partially functional suggested that the positive charge in the lysine residue is not a critical requirement for substrate conversion.

We also studied DPM co-factor binding using a spectral shift assay. We conclude that the lack of activity seen in the K98R and K98A mutants were as a result of failure to bind the DPM co-factor. The K98E, in contrast, retains binding of the DPM co-factor which could be explained through structural modelling, illustrating potential interaction between the negatively-charged carbonyl oxygens on the glutamate (E) residue and the positive amine groups in the DPM co-factor.

In addition, we were able to establish a protocol for native PAGE analysis in order to observe the enzyme-substrate complexes for the wild type enzyme. However, we could not achieve such with any of the mutant PBGDs, as they were too unstable, even with attempted optimization strategies. We therefore could not determine the effect of each of the respective mutations on the enzyme-substrate complexes required for conversion of PBG substrate to HMB product, using this methodology.

Further investigation of the DPM co-factor binding interaction proved successful in part, as we were able to remove the co-factor. However, attempts to reconstitute the wild type enzyme from the *'apo'* enzyme by re-addition of the DPM co-factor were unsuccessful. We found no

evidence that the *'apo'* enzyme shows a preference for the DPM co-factor versus the PBG substrate; however, it is likely that further methodological optimization is needed. The generated *'apo'* enzyme is unstable, and the methodology from Hart et al. (1988) for the refolding of the protein is not ideal, as the protein recovery is too low to perform further analysis [78]. Due to time constraints, we were unable to adapt a different protocol for the protein refolding.

We successfully established a protocol for the analysis of the secondary structure for each of our mutants using far-UV analysis, as well as spectral observations of the 'tertiary footprint' from near-UV analysis. The K98E mutant PBGD illustrated significant changes in its α-helical content, thereby suggesting that the negative glutamate may interfere with other residues within the positive active site cleft. The K98R and K98A showed marginal differences, however none of these changes were significant. The near–UV analysis, on the other hand, allowed for speculation regarding the stability of each of the mutants, from the observations of spectral shifts of aromatic rings. The blue shift of the tyrosine seen in all of the mutants may indicate increased exposure to solvent, and hence a decrease in protein stability. The 'quenching' of the ellipticity signal observed in the K98R and K98A may correlate with changes in enzyme conformation.

In addition, the thermal stability assays we performed further illustrated the instability of the mutants. The wild type demonstrated the characteristic high thermal stability, whereas, each of the mutants had melting temperatures slightly above normal biological temperature (37°C).

The structural analysis and modelling of each of the mutants allowed for the correlation of the kinetic and thermodynamic results, as well as the co-factor assay. Based on molecular modelling, K98E may retain co-factor interaction by forming ionic interactions with the amine groups of the co-factor (as mentioned). In contrast, the bulky arginine of K98R may sterically hinder the binding of the DPM co-factor. The K98A has no ability to form an ionic network, which allowed us to speculate that it promotes increased solvent exposure, thus supporting the instability of this mutant. Homology modelling of each of the mutants to the wild type illustrated that there might be possible changes in conformation of the active site loop. This finding may explain the compromised kinetic function of the mutants as well as the conformational results.

Future Recommendations

It would be interesting to test the kinetics of the DPM co-factor binding with the use of isothermal calorimetry. In brief, this technique analyses changes in heat from binding and dissociation, which would enable the calculation of the dissociation constant k_d. Future work should aim to generate the *'apo'* enzyme via the silencing of the *hemA* gene. Previous work has shown a great success of this method, which enables the growth and culturing of the *'apo'* PBGD. In brief, the silencing of the *hemA* gene results in a lack of transcription of the ALAS enzyme, and therefore no PBG can be produced.

A crystal structure of the human K98E would also be advantageous. This might provide an insight into the effect of the negative charge within the active site and how the enzyme is able to retain co-factor binding. In addition, it would be good to be able to visualize the effect of the mutation on the active site loop, provided the human enzyme allows for the crystallisation of this area (which could not be determined in previously published crystals of the human enzyme, due to distortion of this area).

Finally, there should be a follow-up on the work of Bung et al. (2014) in terms of molecular dynamics [75]. Not only for the K98, but for highlighted residues within the active site as well. This may provide further insight into movement of the domains, the movement of the active site loop, and the interaction network within the active site as a whole.

Appendix 1 – Human nucleotide *HMBS* Sequence (non-erythroid form)

cDNA Sequence (GenBank ID: NM_000190.3)

```
CCGGAAGTGACGCGAGGCTCTGCGGAGACCAGGAGTCAGACTGTAGGACGACCTCGGGTCCCACGTGTCCCCGGT
ACTCGCCGGCCGGAGCCCCCGGCTTCCCGGGGCCGGGGGACCTTAGCGGCACCCACACACAGCCTACTTTCCAAG
CGGAGCCATGTCTGGTAACGGCAATGCGGCTGCAACGGCGGAAGAAAACAGCCCAAAGATGAGAGTGATTCGCGT
GGGTACCCGCAAGAGCCAGCTTGCTCGCATACAGACGGACAGTGTGGTGGCAACATTGAAAGCCTCGTACCCTGG
CCTGCAGTTTGAAATCATTGCTATGTCCACCACAGGGGACAAGATTCTTGATACTGCACTCTCTAAGATTGGAGA
GAAAAGCCTGTTTACCAAGGAGCTTGAACATGCCCTGGAGAAGAATGAAGTGGACCTGGTTGTTCACTCCTTGAA
GGACCTGCCCACTGTGCTTCCTCCTGGCTTCACCATCGGAGCCATCTGCAAGCGGGAAAACCCTCATGATGCTGT
TGTCTTTCACCCAAAATTTGTTGGGAAACCCTAGAAACCCTGCCAGAGAAGAGTGTGGTGGGAACCAGCTCCCTG
CGAAGAGCAGCCCAGCTGCAGGAAAGTTCCCGCATCTGGAGTTCAGGAGTATTCGGGGAAACCTCAACACCCGGC
TTCGGAAGCTGGACGAGCAGCAGGAGTTCAGTGCCATCATCCTGGCAACAGCTGGCCTGCAGCGCATGGGCTGGC
ACAACCGGGTGGGGCAGATCCTGCACCCTGAGGAATGCATGTATGCTGTGGGCCAGGGGGCCTTGGGCGTGGAAG
TGCGAGCCAAGGACCAGGACATCTTGGATCTGGTGGGTGTGCTGCACGATCCCGAGACTCTGCTTCGCTGCATCG
CTGAAAGGGCCTTCCTGAGGCACCTGGAAGGAGGCTGCAGTGTGCCAGTAGCCGTGCATACAGCTATGAAGGATG
GGCAACTGTACCTGACTGGAGGAGTCTGGAGTCTAGACGGCTCAGATAGCATACAAGAGACCATGCAGGCTACCA
TCCATGTCCCTGCCCAGCATGAAGATGGCCCTGAGGATGACCCACAGTTGGTAGGCATCACTGCTCGTAACATTC
CACGAGGGCCCCAGTTGGCTGCCCAGAACTTGGGCATCAGCCTGGCCAACTTGTTGCTGAGCAAAGGAGCCAAAA
ACATCCTGGATGTTGCACGGCAGCTTAACGATGCCCATTAACTGGTTTGTGGGGCACAGATGCCTGGGTTGCTGC
TGTCCAGTGCCTACATCCCGGGCCTCAGTGCCCCATTCTCACTGCTATCTGGGGAGTGATTACCCCGGGAGACTG
AACTGCAGGGTTCAAGCCTTCCAGGGATTTGCCTCACCTTGGGGCCTTGATGACTGCCTTGCCTCCTCAGTATGT
GGGGGCTTCATCTCTTTAGAGAAGTCCAAGCAACAGCCTTTGAATGTAACCAATCCTACTAATAAACCAGTTCTG
AAGGTGTAAAAAAAAAAAAAAAAA
```

Amino Acid Sequence (GenBank ID: P08397.2)

```
MSGNGNAAATAEENSPKMRVIRVGTRKSQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSKIGEKS
LFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPHDAVVFHPKFVGKTLETLPEKSVVGTSSLRR
AAQLQRKFPHLEFRSIRGNLNTRLRKLDEQQEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVR
AKDQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIH
VPAQHEDGPEDDPQLVGITARNIPRGPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH
```

Appendix 2 – Primers for Human *HMBS* Sequence

Fragment 1

GGCACTCGACCGGAATTATCGATTAACTTTATTATTAAAAATTAAAGAGGTATATATTAATGTATCGATTAAATA
VecHMBS1F

[+1]

AGGAGGAATAAACCATGGGGGGTTCT*CATCATCATCATCATCATCAT*GGTATGG**ATG**TCTGGTAACGGCAATGCGGCT

GCAACGGCGGAAGAAAACAGCCCAAAGATGAGAGTGATTCGCGTGGGTACCCGCAAGAGCCAGCTTGCTCGCATA

CAGACGGACAGTGTGGTGGCAACATTGAAAGCCTCGTACCCTGGCCTGCAGTTTGAAATCATTGCTATGTCCACC

ACAGGGGACAAGATTCTTGATACTGCACTCTCTAAGATTGGAGAGAAAAGCCTGTTTACCAAGGAGCTTGAACAT

K98

GCCCTGGAGAAGAATGAAGTGGACCTGGTTGTTCACTCCTTG**AAG**GACCTGCCCACTGTGCTTCCTCCTGGCTTC

ACCATCGGAGCCATCTGCAAGCGGGAAAACCCTCATGATGCTGTTGTCTTTCACCCAAAATTTGTTGGGAAGACC

CTAGAAACCCTGCCAGAGAAGAGTGTGGTGGGAACCAGCTCCCTGCGAAGAGCAGCCCAGCTGCAGAGAAAGTTC

CCGCATCTGGAGTTCAGGAGTATTCGGGGAAACCTCAACACCCGGCTTCGGAAGCTG**GACGAGCAGCAGGAGTTC**
HMBS1R

AG

Fragment 2

CGCATCTGGAGTTCAGGAGTATTCGGGGAAACCTCAACACCCGGCTTCGGAAGCTGGACGAGCAGCAGGAGTTCA
HMBS2F

GTGCCATCATCCTGGCAACAGCTGGCCTGCAGCGCATGGGCTGGCACAACCGGGTGGGGCAGATCCTGCACCCTG

AGGAATGCATGTATGCTGTGGGCCAGGGGGCCTTGGGCGTGGAAGTGCGAGCCAAGGACCAGGACATCTTGGATC

TGGTGGGTGTGCTGCACGATCCCGAGACTCTGCTTCGCTGCATCGCTGAAAGGGCCTTCCTGAGGCACCTGGAAG

GAGGCTGCAGTGTGCCAGTAGCCGTGCATACAGCTATGAAGGATGGGCAACTGTACCTGACTGGAGGAGTCTGGA

GTCTAGACGGCTCAGATAGCATACAAGAGACCATGCAGGCTACCATCCATGTCCCTGCCCAGCATGAAGATGGCC

CTGAGGATGACCCACAGTTGGTAGGCATCACTGCTCGTAACATTCCACGAGGGCCCCAGTTGGCTGCCCAGAACT

TGGGCATCAGCCTGGCCAACTTGTTGCTGAGCAAAGGAGCCAAAAACATCCTGGATGTTGCACGGCAGCTTAACG
Stop

ATGCCCAT**TAA**CTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGATCGATG

GGGATCCGAGCTCGAGATCTGCAGCTGGTACCATATGGGAATTCGAAGCTTGGCTGTTTTGGCGGATGAGAGAAG

ATTTTCAGCCTGATACAGATTAAAT**CAGAACGCAGAAGCGGTCT**
VecHMBS2R

Appendix 3 – Media

Equipment

Autoclave Huxley, Speedy (Laboratory and Scientific Co. (Pty) Ltd., Cape Town, South Africa)

Reagents

- Agar Noble (Biolab Diagnostics, Gauteng, South Africa)
- Ampicillin 100mg/ml
- Bacto® - Yeast Extract
- Bacto® - Tryptone
- NaCl

Methodology

Luria-Bertani (LB) Medium 1L pH 7.5

- Weigh out the following reagents:

NaCl	5.0 g
Yeast Extract	5.0 g
Tryptone	10.0 g

- Add 1L distilled water to dissolve
- Sterilise by autoclaving

Store at 4ºC

LB Agar Plates 200mL pH 7.5

- Weigh out the following reagents:

Agar	3.0 g
NaCl	1.0 g
Tryptone	2.0 g
Yeast Extract	1.0 g

- Add 200 ml distilled water to dissolve
- Sterilise by autoclaving

- Allow to cool to approximately 35 – 45 ºC (warm to touch)
- Add 0.25 ml ampicillin (100 mg/ml stock)
- Pipette 20 ml into Petri dishes (pipette to the side to prevent bubble formation)
- Allow 20 minutes to cool/set with lid slightly open to prevent condensation
- Close lid and seal dish with parafilm

Store inverted at 4ºC in sealed plastic bag

Appendix 4 – Generation and Storage *HMBS* stocks

Equipment

- Orbital shaker incubator (Yidher LM-510, Taiwan)
- Sterile Falcon™ Round-Bottom Polypropylene Tubes (15 ml) (Fischer Scientific, Pittsburgh)
- *Cryo*-storage tubes (2 ml) (Bio-Rad Laboratories, Munich, Germany)

Reagents

- 100 mg/ml Ampicillin
- LB medium
- Sterile 100% Glycerol

Methodology

- Inoculate 6 ml LB medium containing ampicillin (0.1 mg/ml) with 6 µl *E.coli* JM109 *HMBS* plasmid (from overnight culture)
- Incubate at 37ºC on shaking incubator (225 rpm) overnight (14 - 18 hours)

- Pipette 300 µl of glycerol and 700 µl of culture into 2 ml *cryo*-storage tubes
- Mix by inverting (± 6 times)
- Label appropriately

Store at -80ºC

Appendix 5 – Plasmid DNA Isolation (from *E.coli*)

Protocol from Wizard® Plus SV Minipreps DNA Purification kit (Promega Corporation, Madison, USA)

Equipment

- Hermle Z400 centrifuge (Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)
- Labnet Spectrafuge 24D (Labnet International, Inc., Edison, NJ, USA)
- 1.5 ml Sterile microcentrifuge tubes (Axygen Scientific Inc., Union City, CA, USA)

Methodology

- Centrifuge 6 ml overnight culture (Appendix 4) at 4000xg for 10 minutes
- Discard supernatant (LB Waste); resuspend pellet with 250 µl *Cell Resuspension solution*
- Transfer resuspended sample to a 1.5 ml microcentrifuge tube

- Add 250 µl *Cell Lysis solution* to sample; mix by inverting 4 - 6 times
- Incubate at room temperature until solution clears (4 - 5 minutes)
- Add 10 µl Alkaline Protease Solution; mix by inverting 4 - 6 times
- Incubate at room temperature for **exactly** 5 minutes

- Add 350 µl Neutralization Solution; mix by inverting 4 - 6 times
- Centrifuge at 14000xg for 10 minutes at room temperature
- Place column into collection tube. Label both column and collection tube clearly
- Pipette approximately 850 µl (maximum) of supernatant into column; centrifuge at 14000xg for 1 - 2 minutes
- Discard flow through and re-insert column into collection tube

- *Importantly: leave small amount of supernatant behind above precipitate to ensure no precipitate is pipetted out with supernatant. If, however, traces of precipitate are observed -pipette into sterile microcentrifuge tube and centrifuge at 14000xg for 5 minutes. Transfer supernatant into column*

- Add 750 µl of Wash Solution to column; centrifuge at 14000xg for 1 - 2 minutes
- Discard flow through and re-insert column into collection tube

- Repeat wash procedure with 250 µl of Wash solution; Centrifuge at 14000xg for 2 minutes
- Transfer column into a sterile 1.5 ml microcentrifuge tube; add 50 µl of Nuclease-free water directly on to membrane in column.
- Incubate at room temperature for 2 minutes
- Elute DNA by centrifuging at 14000xg for 2 minutes
- Quantify DNA and determine purity (Appendix 6)

Store at -20^{o}C

Appendix 6 – DNA Quantification

Equipment

- Gene Quant Spectrophotometer (Pharmacia Biotech, Cambridge, UK)
- Glass capillaries (Separation Scientific, Honeydew, RSA)

Reagents

- Reference/ Blank (nuclease-free water used to elute purified DNA – Appendix 5)
- Extracted DNA (Appendix 5)

Methodology

- Set GeneQuant for double stranded DNA
- Path Length 0.5 mm
- Factor dsDNA
- Tare GeneQuant:

- Aliquot 5 μl of reference into sterile 0.5 microcentrifuge tube
- Dip capillary into reference sample until approximately 3/4 of the capillary is filled with solution
- Prevent solution from leaking out by sealing top end of capillary with a small amount of Presstick
- Remove excess solution on the outside of capillary by wiping with clean paper towel
- "Blank" GeneQuant using reference sample

Quantifying DNA

- Insert capillary into sample as described above
- Read the O.D of the DNA and note the percentage purity, DNA concentration and the 260:280 ratio

Appendix 7 – Primer (Oligonucleotide) Design & Preparation for use

Equipment

- Vortex - 2 Genie (Scientific Industries Inc., Bohemia, NY, USA)
- Primer Designer

Internet Sources

- Oligo analyser (Integrated DNA Technologies Inc. Coralville, Iowa 52241, USA)

Table A7.1 Table of Primers designed for screening of SDM mutants.

Primer Name	Region	5'--- 3'	base pairs
VecHMBS1F	Vector	5'-GGC ACT CGA CCG GAA TTA TC-3'	20
HMBS1R	HMBS	5'-CTG AAC TCC TGC TGC TCG TC-3'	20
HMBS2F	HMBS	5'-CGC ATC TGG AGT TCA GGA G-3'	19
VecHMBS2R	Vector	5'-AGA CCG CTT CTG CGT TCT G-3'	19

F - Forward; R - Reverse; Vec - Vector; 1&2 - Fragment numbers

Preparation for use:

- Briefly spin tube containing primer (oligo)
- Resuspend oligo in 1 ml sterile water
- Incubate at room temperature for a minimum of 20 minutes
- Vortex to ensure oligo is completely dissolved, and spin down briefly
- Prepare working stocks of 25 µM (20 µl aliquots)

Store at -20° C

Appendix 8 – Agarose Gel Electrophoresis

Equipment:

- BioRad® Sub cell GT System for agarose gels (Bio-Rad Laboratories, Munich, Germany)
- BioRad® GT UVTP gel tray 15x7 cm (Bio-Rad Laboratories, Munich, Germany)
- PS-1500 DC power supply (Hoefer Scientific Instruments, San Francisco, USA)
- UVItech gel documentation system (UVItech, Limited, Cambridge, UK)

Reagents:

- Seakem LE agarose (Whitehead Scientific, Cape Town, RSA)
- 10 x TAE Buffer *(pH 8.0)*
- 1 mg/ml ethidium bromide (Sigma-Aldrich, St. Louis, MO, USA)
- Sucrose sample solution
    - 30g sucrose
    - 10 mg Bromophenol blue
    - 5 ml 0.5M Na_2EDTA (pH 8.0)

Methodology:

- Weigh out agarose (for 1% gel = 0.5 g in 50 ml 1 x TAE) in 250 ml flask and add TAE
- Microwave agarose and cool briefly
- Add 10 µl of 1 mg/ml ethidium bromide, mix and pour into gel tray
- Insert spacer and allow to set
- Pour 1 x TAE buffer to cover gel in gel tank
- Load samples (2:1 ratio with sucrose sample solution)
- Run gel at 100 V for 1.5 hours
- Visualize gel using UVItech documentation system

Appendix 9 – QuickChange® Site-Directed Mutagenesis

Equipment and Reagents

- 0.6 ml sterile PCR microcentrifuge tubes (Axygen Scientific Inc., Union City, USA)
- GeneQuant spectrophotometer (Pharmacia Biotech, Cambridge, UK)
- Hybaid Omnigene thermal cycler (The Scientific Group, Cape Town, RSA)
- Labnet Spectrafuge 24D (Labnet International, Inc., Edison, NJ, USA)
- Orbital Shaker Incubator (Yidher LM-510, Taiwan)
- SI-18 incubator (Stuart Scientific, Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)
- Sterile 17 x 100 mm polypropylene tubes (Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)
- Water Bath (Memmert, Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)

Reagents

- Agar Plates (Appendix 3)
- *DpnI* restriction endonuclease (Fermentas, Life Sciences, Inqaba Biotech, RSA)
- Mutagenic oligonucleotides (Integrated DNA Technologies Inc., Coralville, IA, USA)
- *Pfu Turbo*® Buffer (Promega Corporation, Madison, W1, USA)
- *Pfu Turbo*® DNA polymerase (Promega Corporation, Madison, W1, USA)
- SABAX sterile, d.H_2O (Adcock Ingram Critical Care, Johannesburg, RSA)

Methodology:

9.1 Mutagenic Oligonucleotide design

1. Design two complementary mutagenic oligonucleotides containing the desired base change. The desired mutation should be centered in the primer with 10 - 15 bases of wild type sequence on either side. It is highly recommended that the designed oligo ends with a G or C nucleotide in order to stabilize the oligo during the annealing between oligo and the template.

2. Using Oligonucleotide parameter analyser determines the GC content, melting temperature (Tm); self-dimer and heterodimer as well as ΔG.

Parameter	Minimum	Maximum	Optimum
%GC	40	60	55
length	25	45	25

*The Tm of the hairpin as well as oligo hetero-dimer should be lowers than the annealing temperature of the reaction.

3. Select a restriction enzyme from a relevant website (for example Webcutter 2.0) to enable restriction analysis to be performed that will allow for identification between the mutant and the wild type.

Table A9.1 HMBS mutant primers (oligos)

Mutant	Sequence	Length (bp)
K98E.F	5'-GTTCACTCCTTGGAGGACCTGCCCAC-3'	26
K98E.R	5'-GTGGGCAGGTCCTCCAAGGAGTGAAC-3'	
K98R.F	5'-GTTCACTCCTTGAGGGACCTGCCCAC-3'	26
K98R.R	5'-GTGGGCAGGTCCCTCAAGGAGTGAAC-3'	
K98A.F	5'-GTTGTTCACTCCTTGGCGGACCTGCCC-3'	27
K98A.R	5'- CAGTGGGCAGGTCCGCCAAGGAGTGA-3'	

9.2 DNA Extraction

See Appendix 5

9.3 Mutant Strand Synthesis - Polymerase Chain Reaction

Reagents

- 10 x Reaction Buffer
- 10 µM mutagenic oligos (Table A9.1)
- d.H_2O
- dNTPs (25 mM)
- dsDNA (10 ng/µl)
- *Pfu Turbo*® polymerase (100U)

Methodology

Table A9.2 PCR Reaction Set up

Reagent	Final Concentration	Volume (µl)
Reaction Buffer (10 X)	1 X	5
dNTPs (25 mM)	10 mM	2
Forward and Reverse Mutagenic Primers (10 µM)	7.4 µM	2 (each)
Pfu Turbo® polymerase (3 U/µl)	1 U	0.5
d.H_2O		25.5
Purified DNA (50 ng/µl)	10 ng/µl	5

For each sample add 37 µl of master mix. Hereafter determine amount of magnesium to be added;

50 µl reaction			
Tube Samples	**DNA (µl)**	**H_2O (µl)**	**Mg^{2+} (µl)**
No DNA	-	13	0
No Mg^{2+}		8	0
2 mM Mg^{2+}	5	4	4
4 mM Mg^{2+}		0	8

Take care not to vortex to prevent bubble formation; flick samples to mix and spin down briefly

Table A9.3 Hybaid Omnigene Program Cycle - For one base change

	Temperature (ºC)	**Time (minutes)**	**Number of Cycles**
Initiation Denaturation	94	5:00	1
Denaturation	94	0:30	
Annealing	55	0:30	12# (16**)
Elongation	72	12:00*	

* Protocol stipulates 2 min/kb; *HMBS* together with insert = ± 6kb

Number of cycles equals 12 for single base pair change. Wild type – AAG; K98E – GAG; K98R – AGG

** For a 2 base pair change mutation 16 cycles are required. K98A – GCG

- Place samples on ice following PCR
- Confirmation of PCR product
  - Run samples on a 0.8% agarose gel (Appendix 8)
  - Add 8 µl ethidium bromide (1mg/ml) to gel prior to pouring
  - Mix & load 5 µl of PCR product + 5 µl d.H_2O + 1.5 µl loading dye
  - Run gel at 80 V

- Select strongest and purest PCR product
- *DpnI* RE digest of PCR product
  - 45 µl PCR product
  - 2 µl *DpnI*
  - 4.7 µl *DpnI* buffer added directly to sample
  - Centrifuge samples briefly
  - Incubate at 37ºC for one hour

9.4 Transformation

- Chill 17 x 100 mm falcon tubes on ice
- Allow stored JM109 cells (stored at - 70ºC) to thaw on ice

- Immediately aliquot 100 µl JM109 cells into chilled falcon tubes
- Add 5 - 10 µl of digested product (ensure no mineral oil is included); flick 4 - 6 times
- Incubate on ice or 30 minutes
- Heat shock reaction exactly - 42ºC water bath for 45 seconds
- Immediately place on ice for 2 minutes

- Add 900 µl LB media, incubate with shaking incubator (225 rpm) at 37ºC for 1 hour
- Transfer to a 1.5 ml microfuge tube
- Spin down cells at 4000xg for 2 minutes
- Discard 900 µl of supernatant and resuspend cells with remaining supernatant
- Spread previously prepared agar plate (Appendix 3) with 50 µl of transformed JM109 cells
- Leave to air dry for ± 20 minutes
- Invert plates and incubate at 37ºC overnight
- Visible single clones may be isolated and inoculated for expression (Appendix 4)

Appendix 10 – Amplification of *HMBS* using Polymerase Chain Reaction (PCR)

Equipment and Reagents

- 0.6 ml sterile PCR microcentrifuge tubes (Axygen Scientific Inc., Union City, USA)
- GeneQuant spectrophotometer (Pharmacia Biotech, Cambridge, UK)
- Labnet MultiGene OptiMax Thermal Cycler. (The Scientific Group, Cape Town, RSA)
- Labnet Spectrafuge 24D (Labnet International, Inc., Edison, NJ, USA)
- Vortex - 2 Genie (Scientific Industries Inc., Bohemia, NY, USA)
- Microcentrifuge tubes 1.5 ml and 0.5 ml (Axygen Scientific Inc., Union City, USA)

Reagents:

- *GoTaq Flexi* (Promega Corporation, Madison, W1, USA)
- Magnesium-free DNA polymerase 5 x Buffer (Promega Corporation, Madison, W1, USA)
- 25 mM $MgCl_2$ (Promega Corporation, Madison, W1, USA)
- Deoxynucleotide triphosphate (dNTPs) (Promega Corporation, Madison, W1, USA)
    - dATP, dCTP, dGTP, dTTP
- Designed primers (see Appendix 9)
- SABAX sterile, d.H_2O (Adcock Ingram Critical Care, Johannesburg, RSA)

Methodology:

- Set up PCR reaction (master mix) in a 1.5 ml microcentrifuge tube on ice for the number of desired reactions.

Table A10.1 The standard PCR reaction mixture

Reagent	Final Concentration	Volume (µl)
Buffer (5 X)	1 X	10
Deoxynucleotide triphosphates; dNTPs (2.5 mM)	100 µM	1
Forward and Reverse Primers (25 µM)	0.5 µM	$1_{(each)}$
GoTaq flexi DNA Polymerase (5 U/ µl)	1 U	0.2
$MgCl_2$ (25 mM)	1.5 mM	3
Purified DNA (100-200 ng/µl) For Negative Control used d.HO	2 – 4 ng/ µl	1

- Aliquot 49 µl of master mix into a 0.5 ml microcentrifuge tube
- Add 1 µl DNA (100-200 ng)
- Add 1µl of H_2O to the blank
- Mix and spin down briefly

Note: The PCR cycling conditions used for the *HMBS* primers were all optimised, and it was found that the annealing temperature (T_a) for all primer pairs was 58ºC.

Table A10.2 The PCR Cycling Conditions for the *HMBS* Primers.

	Temperature (ºC)	Time (minutes)	Number of Cycles
Initiation	95	1:00	1
DNA Denaturation	95	0:30	
Annealing	58	0:30	35
Elongation	72	0:30	
Final Extension	72	8:00	1

Appendix 11 - 6% Non-Denaturing Polyacrylamide Gel Electrophoresis (PAGE)

Equipment:

- 50 ml glass measuring cylinder
- 100 ml glass beaker
- Hoefer vertical slab gel electrophoresis unit, SE600 series (Hoefer Scientific Instruments, San Francisco, USA)
    - Glass gel plates
    - Spacer comb (1.5 mm)
    - Clamps
    - Combs and spacers (1.5 mm)
    - Casting stand
    - Upper & lower buffer chambers

- Hamilton blunt end syringe (The Hamilton Company, Nevada, USA)
- PS1500 DC power supply (Hoefer Scientific Instruments, San Francisco, USA)
- UVItec gel documentation system (UVItec Limited, Cambridge, UK)

Reagents

- 10 X TBE Buffer pH 8.0
- Tris/HCL 890 mM
- Boric acid 890 mM
- EDTA 20 mM

- 30% acrylamide, 0.8% bis-acrylamide (A-Bis-A)
- 10% ammonium persulphate (Promega Corporation, Madison, W1, USA)
- TEMED (N,N,N',N', - Tetramethyl-ethelenediamine)
  Store at 4 ºC

- Sucrose sample solution (Loading Dye)
    - Sucrose 30 g
    - Bromophenol Blue 10 mg
    - Na_2EDTA (0.5M ; pH 8) 5 mL

- Make up to a total volume of 50 mL with $d.H_2O$
- Store in aliquots at - 20 ºC

- Ethidium Bromide
    - 100 µl of 1 mg/ml stock solution in 200 ml $d.H_2O$ (Sigma-Aldrich, St. Louis, MO, USA)

Methodology

- Assemble glass plates (using 1.5 mm spacers) with clamps; mount plates on casting stand
- 6% Gel Solution (50 mL Glass Cylinder)
    - 10 x TBE Buffer 4 ml
    - A-Bis-A 8 ml
    - Add d.H_2O to a final volume of 40 ml; then add,
    - TEMED 40 µl
    - Ammonium persulphate 400 µl

- Mix Gel Solution by pouring into 50 ml Glass Beaker
- Pour solution into the space between the clamped plates in the gel casting stand
- Insert 20 bay sample comb
- Allow gel to polymerize at room temperature for minimum of 30 minutes
- Assemble upper chamber buffer on top of gel plates with clamps and fill with 500 ml 1 x TBE
- Fill lower chamber with 1L 1 x TBE
- Prepare 10 µl of samples and molecular marker (1:1 with loading dye solution)
- Load samples into gel wells using Hamilton syringe
- Run gel at 250 V for 1.5 hours
- Remove gel from between glass plates and allow to stain in a solution of ethidium bromide for 10 minutes
- Rinse in d.H_2O
- Visualize gel on UVltec documentation system and photograph

Appendix 12 – DNA Purification from PCR Product

Equipment and Reagents:

- 1.5 ml microfuge tubes (Greiner, Lasec, Cape Town, South Africa)
- SABAX sterile d.H_2O (Adcock Ingram Critical Care, Johannesburg, South Africa)
- Illustra^{TM} GFX^{TM} PCR DNA and Gel Band Purification Kit (Amersham Biosciences Ltd.)
- Labnet Spectrafuge 24D Centrifuge (Labnet International Inc.)
- Vortex – 2 Genie G560-E Mixer (Scientific Industries Inc

Methodology:

- Transfer PCR product into a sterile 1.5 ml microfuge tube
- Add 500 µl Capture buffer type 3 (blue cap)
- Mix solution thoroughly
- Spin down briefly

- Place GFX column into a sterile collection tube
- Transfer the capture buffer 3 and sample mixture into the GFX column
- Centrifuge at 14000xg for 40 seconds
- Discard flow through in the collection tube. Reinsert GFX column back into collection tube

- Add 500 µl of wash buffer type 1 (yellow cap)
- Centrifuge at 14000xg for 40 seconds; discard flow through in the collection tube
- Repeat with 250 µl of wash buffer type 1

- Place GFX column into a sterile 1.5 ml microfuge tube
- Add 30 µl elution buffer type 6 (pink cap) directly onto membrane in the GFX column
- Incubate at room temperature for 2 minutes
- Centrifuge at 14000xg for 1 minute

- Determine DNA concentration and purity using the GeneQuant spectrophotometer (Appendix 6); use the elution buffer type 6 as the reference

  Store at - 20ºC

Appendix 13 – TALON™ Affinity Protein Purification

Equipment and Reagents

- Orbital Shaker Incubator (Yidher LM-510, Taiwan)
- Refrigerated Centrifuge (Centrikon T-324, Kontron Instruments, Italy)
- Misonix Sonicator 300 (Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)
- Source 15Q glass column (6 mm x 300 mm) (Amicon Ltd, Stonehouse, UK)
- Gilson Minipuls 3 Pump (Laboratory and Scientific Equipment Co. (Pty) Ltd., Cape Town, RSA)
- Glass pipettes
- Small sterile glass test tubes (12 x 75 mm) (B&M Scientific, Cape Town, South Africa)

Reagents:

- TALON® metal affinity Resin (Clonetch Laboratories, Palo Alto, CA, USA)
- Phenyl-methyl-sulfonyl fluoride (PMSF) (10mg/ml) Sigma-Aldrich, St. Louis, MO, USA)
    - PMSF 100 mg
    - Isopropanol 10 ml

- ***Sonication Buffer*** ***pH 8.0***
    - Tris-HCl 20 mM
    - NaCl 100 mM
    - Tween-20 1%

- ***Equilibration Buffer*** ***pH 8.0***
    - Tris-HCl 20 mM
    - NaCl 100 mM
    - Tween-20 0.5%

- ***Wash Buffer*** ***pH 8.0***
    - Tris-HCl 20 mM
    - NaCl 100 mM
    - Tween-20 0.5%
    - Imidazole 25 mM

- ***Elution Buffer*** ***pH 8.0***
    - Tris-HCl 20 mM
    - NaCl 100 mM
    - Tween-20 0.5%
    - Imidazole 150 mM

Methodology:

- Inoculate 1 ml of stock cells and 1 ml ampicillin (100 mg/ml) into 1L LB Media (Appendix 3)
- Incubate on shaker (225 rpm) at 37ºC for 22 hours
- Aliquot culture into centrifuge tubes of 125 ml
- Centrifuge for 30 minutes at 4ºC and 3000 g
- Discard supernatant; resuspend pellets in 30 ml sonication buffer + 30 µl PMSF; use the club shaped glass rod
- Avoid 'frothing' and bubble formation during re-suspension

Sonication:

- In a plastic beaker, sonicate solution using the macro-probe of the Sonicator
- Sonicate sample at 60 Watts for 30 seconds; keep probe at the bottom of beaker to prevent 'frothing'
- Place sample on ice for 2 minutes – prevent overheating and consequent protein denaturation
- Repeat 30 second sonication another three times; keep on ice following every sonication
- Ultracentrifuge sonicate at 4ºC (105000 g) for 30 minutes
- Collect supernatant for column and label (LOAD); note volume

Column Preparation:

- Set up the column on a retort stand, so that column is approximately 10 cm from bottom
- Connect the column to the pump by connecting plastic tubing
- Add 1.0 ml of resin into the column and allow to settle
- Equilibrate the column with 10 ml of equilibration buffer (containing 10 µl of PMSF); take care not to disrupt resin bed
- PBGD is light sensitive; cover resin with aluminium foil

Protein binding and elution

- Load the supernatant of sonicate into the column (rate 15 ml/hr)
- Collect flow-through in fractions; pool samples and label VOID; note volume
- Add 10 ml equilibration buffer (with 10 µl PMSF) to ensure binding of final load to column**;** after 20 minutes increase rate to 16-18 ml/hr
- Add 10 ml wash buffer (with 10 µl PMSF) and collect in fractions; label WASH; note volume

- To elute protein: Add 10 ml elution buffer (with 10 µl PMSF) to column
- Clamp and remove tubing from pump;
- Collect flow through in fractions of 10 drops per tube (approximately 500 µl)
- Collect 6-8 fraction samples
- Pool all samples (PBGD) and note final volume

Appendix 14 – 7.5 – 15.5% Gradient SDS Polyacrylamide Gel Electrophoresis (PAGE)

Equipment and Reagents

- Hoefer vertical slab gel electrophoresis unit (SE600 Series) (Amersham Pharmacia Biotechnology Inc.)
- UVITech Gel Documentation apparatus, (UVItech, Limited, Cambridge, UK)
- K65HM High-density synthetic paper (Whitehead Scientific)
- SG Gradient Maker (Hoefer Scientific Instruments, San Francisco, CA, USA)
- 25 ml glass measuring cylinders
- 10 ml volumetric Flask

Additional Reagents

- 17.5% Resolving solution (in 25 ml measuring cylinder)
  - 7 ml "High" Buffer (1M Tris/HCl, pH 8.8, 30% glycerol)
  - 12 ml A-Bis-A Solution (30% Acrylamide, 0.8% bis-acrylamide)
  - 0.3 ml 10% SDS
  - Make up to 20 ml with d.H_2O and mix well

- 7.5% Resolving solution (in 25 ml measuring cylinder)
  - 7 ml "Low" Buffer (1M Tris/HCl, pH 8.8, 7.5% glycerol)
  - 5 ml A-Bis-A Solution (30% Acrylamide, 0.8% bisacrylamide)
  - 0.3 ml 10% SDS
  - Make up to 20 ml with d.H_2O and mix well

- Spacer Solution (in 10 ml volumetric Flask)
  - 1.2 ml A-Bis-A Solution
  - 8.6 ml spacer buffer (0.125 M Tris/HCl, pH 6.8)
  - 0.1 ml 10% SDS

Methodology

- Boil samples for 5 minutes with an equal volume of 2% SDS, 0.2% mercaptoethanol, 20% glycerol and 0.002% bromophenol blue in 0.125 M Tris/HCl pH 6.8
- To each resolving solution add 100µl of freshly prepared 5% ammonium persulphate and 10 µl TEMED immediately before use. Mix Well.
- Set up glass plates with 1.5 mm spacers and mount on casting stand
- Pour 16 ml of both solutions into the gradient pourer. The low solution in the left chamber and the high solution in the right chamber – containing a stirrer and placed on a magnetic plate
- Introduce water at the top
- Once resolving gel has set add spacer solution containing freshly prepared 100µl 15% ammonium persulphate and 10µl TEMED
- Insert 1.5 mm comb
- Load samples into wells using Hamilton syringe
- Run gel at 80 V for 16 hours
- Track bromophenol blue to asses gel front; should be 1.5 cm from the bottom
- Remove the gel between plates
- Stain gel (Appendix 15)
- Visualize gel using UVItech documentation system

Appendix 15 – Coomassie® R-250 PAGE Gel Staining

Equipment

- UVITech Gel Documentation apparatus, (UVItech, Limited, Cambridge, UK)
- K65HM High-density synthetic paper (Whitehead Scientific)
- Plastic container (old microtip box)

Reagents

- Staining Solution
  - 0.1% Coomassie® R250 Brilliant Blue Dye (Life Technologies, South Africa)
  - 50% MeOH
  - 10% glacial acetic acid

- De-staining solution
  - 40% MeOH
  - 10% glacial acetic acid

- Storage Solution
  - 5% glacial acetic acid

Methodology:

- Stain for 20-30 minutes on shaker
- Decant stain (Can be re-used 2-3 times).
- De-stain gel on shaker
  - Use small amount and exchange every 15-20 minutes*
    *Place sponge in box to aid the de-staining
- Store in final solution
- Visualize gel using UVItech documentation system

Appendix 16 – BioRad® Bovine Serum Albumin (BSA) micro-assay for Protein Quantification

Equipment

- Small sterile glass test tubes (12 x 75 mm) (B&M Scientific, Cape Town, South Africa)
- Hitachi U-1100 UV/VIS spectrophotometer (Koki CO, Ltd., Tokyo, Japan)
- Cuvette (plastic)

Reagents

- 1 mg/ml bovine serum albumin (BSA) (Sigma-Aldrich, St. Louis, MO, USA)
- Bio-Rad® Protein Assay Dye Reagent Concentrate (Bio-Rad® Laboratories, Munich, Germany)

Methodology

- Make a 1/10 dilution of BSA (i.e. 0.1 mg/ml)
- Prepare 6 linear dilutions of the 0.1 mg/ml BSA (20-70 µl in final volume of 800 µl)
- Prepare relevant dilutions of samples and include a blank (negative control)
- Add 200 µl Bio-Rad® dye reagent to each tube and vortex; avoid bubbles
- Incubate at room temperature for 5 minutes (1 hr max)
- Set spectrophotometer (595 nm) and read samples against relevant blank

- Generate and plot standard curve from BSA concentrations and absorbance
- Plot linear trend-line
- Extrapolate concentration of samples from the standard curve

Appendix 17 PBGD Activity Assay

Equipment

- 0.2 µm PTFE filters (B&M Scientific, Cape Town)
- Aluminium tube rack
- Eppendorf repeater pipette (Merck Millipore, RSA)
- Small sterile glass test tubes (12 x 75 mm) (B&M Scientific, Cape Town, South Africa)
- Hitachi 650/10S Fluorescence Spectrophotometer (Koki, Co (Ltd), Japan)
- Hitachi U-1100 UV/VIS spectrophotometer (Koki CO, Ltd., Tokyo, Japan)
- Memmert shaking water incubator (Lab and Scientific, Cape Town)
- Plastic cuvettes (4.4 ml)

Reagents

- 5N HCl solution
- 0.5N HCl solution
- 1 mg/ml Uroporphyin I (in 5N HCl) (Frontier Scientific, USA)
- 5.5 mM Dithiothreitol (DTT) (Sigma, St Louis, MA USA)
- 5 mM PBG (substrate) (Frontier Scientific, USA)
- 39 nM PBGD
  Store at 4ºC

- Assay Buffer: 0.1M Tris-HCl
- Cuvette Buffer: 0.01M Tris-HCl

Methodology

Standard

- Pipette 50 µl from 1 mg/ml Uroporphyrin I stock solution and add 4.95 ml of 5N HCl. This solution is the working solution and is stable for ± 1 week. Importantly, filter and measure absorbance prior to every assay
- From the working solution make a 1/100 dilution in 0.5N HCl
- Measure the absorbance at 405 nm in the spectrophotometer. Determine the concentration of the working solution by the Beer-Lambert Law ($A=\varepsilon Cl$) where $\varepsilon_{(UroI)}$ = 541 mM
- From the working solution pipette 250 µl into a glass tube and add 6 ml of 5N HCl. This is a 1000X standard
- Make a serial dilution range from 1000X in 5N HCl. 10, 25, 50, 100, 150, 200, 250
- Use the 250 dilution and set the fluorimeter to 1000 (excitation wavelength, 405 nm and emission, 595 nm), with slit widths of 6 and 6.5 cm, respectively at a. Sensitivity of x1
- Plot a graph (linear) using the standards readings to determine activity

Assay

- Set shaking water incubator to 37ºC
- Place aluminium tube rack on ice
- Make up a 1 mg/ml Uroporphyrin I solution in 5N HCl to use as standard; taking care to keep sample on ice and in the dark (cover with foil). Work in dim light.
- Make up the 5.5 mM DTT and 5 mM PBG in Assay Buffer; keep on ice
- Dilute PBGD (sample) in Assay Buffer
- Prepare samples of BSA of equivalent concentration as the PBGD to act as a blank
- Add the following to each cuvette; where possible using an Eppendorf repeater pipette
    - 920 µl Assay Buffer
    - 10 µl DTT
    - 20 µl PBG
    - 50 µl PBGD
- Immediately after adding the PBGD vortex, and incubate in shaking water incubator at 37ºC in the dark for 20 minutes.
- Terminate the reaction by adding 100 µl of 5N HCl and vortex briefly
- Place samples in area of bright light for 20 minutes.
- Add 1 ml Cuvette Buffer to each sample.

- Measure fluorescence at excitation wavelength of 405 nm and emission wavelength at 595 nm

Note: Ensure (if necessary, by optimization) that all concentrations of PBDG measured fall within the linear part of the curve

Appendix 18 Kinetic Assay

All equipment and reagents are as detailed in Appendix 17, apart from these mentioned below:

Reagents

- 0.5 mM PBG (substrate)
- 19.5 pM PBGD

Methodology

Standard

As detailed in Appendix 17

Assay

- Make up 0.5 mM PBG in Assay Buffer; keep on ice
- Add increasing concentrations of PBG to each successive tube over an appropriate range at the same concentration of PBDG (19.5 pM)
- For each concentration of PBG – have a BSA blank
- Volumes for the highest PBG concentration as follows: (as an example)
    - 760 µl Assay Buffer
    - 10 µl DTT
    - 180 µl PBG
    - 50 µl PBGD

- Perform Assay as detailed in Appendix 17

Note: Kinetic Assay requires - optimization with regards to the PBGD concentration for both wild type and mutants in order to ensure that maximal velocity is reached and that the concentration of PBGD used is not too great.

Equipment:

- SpectraMaxPlus 384 Spectrophotometer, Molecular Devices (Labotec, RSA)
- Geiner bio-one 96-well flat bottom microplate (Laboratory & Scientific (Pty) Ltd, CT, RSA)

Software

- SoftMax Pro Microplate Data Acquisition & Analysis Software, Molecular Devices (Labotec, RSA)
- GraphPad Prism, Version 6.03 for Windows. (Used Spaghetti graphs – multiple y output for single x input)

Reagents

- 2M HCl freshly prepared
- 2% 4 - (dimethylamino)benzaldehyde (DAB)
- 98% formic acid
- 0.5 mg/ml PBGD enzyme solution

Methodology

- Prepare the modified Ehrlich's reagent as follows: Dissolve 2% DAB in a 98% formic acid/2M HCl solution in the ratio of (7:3 v/v)
- Dilute the pure eluted protein (Appendix 13 & 16) to a final concentration of 0.5 mg/ml
- Set up the SoftMax Pro software to scan absorbance from 440 nm to 650 nm at 5 nm intervals. Apply the auto-calibration settings for every reading.
- Add 75 μl of both enzyme and Ehrlich's reagent into the 96-well microplate. Ensure no bubble formation & that the bottom of the well is completely covered to avoid false readings.
- Include a well containing only Ehrlich's reagent as a negative blank.
- Immediately after adding the Ehrlich's reagent to samples (*note in WT and K98E bright pink colour)*, insert microplate into spectrophotometer to scan. Note these readings as time 0 (T_0). Set a timer for 20 minutes
- Following 20-minute incubation repeats the scan. Note readings as time 20 (T_{20})
- Export readings as text files. Analyse and generate graphs using desired software

Equipment:

- Amicon® Ultra 4 ml centrifugal filters (30 kDa cut-off) (Merck Millipore, South Africa)
- Standard (RC) Membrane Dialysis tubing 10mm (25 kDa cut-off) (GIC Scientific, Roodepoort, South Africa)
- Small sterile glass test tubes (12 x 75 mm) (B&M Scientific, Cape Town, South Africa)

Software

- SoftMax Pro Microplate Data Acquisition & Analysis Software, Molecular Devices (Labotec, RSA)
- GraphPad Prism, Version 6.03 for Windows. (Used Spaghetti graphs – multiple y output for single x input)

Reagents

- 2 mg/ml DTT
- 1.75 mg/ml HMBS purified enzyme
- Conc. HCl
- 6M Urea solution

- **"Buffer 1" (pH 7.2)**
  - 50 mM $NaPO_4$
  - 0.6 mM EDTA
  - 1 mM DTT
  - 0.1M NaCl
  - 0.1M Urea

- **"Buffer 2" (pH 7.2)**
  - 50 mM $NaPO_4$
  - 0.6 mM EDTA
  - 1 mM DTT
  - 0.1M NaCl

- **"Buffer 3" (pH 7.4)**
  - 50 mM $NaPO_4$
  - 0.6 mM EDTA
  - 0.2 mM DTT

Methodology

- DTT was added at a concentration of 2 mg/ml to HMBS (1.75 mg/ml) followed by concentrated HCl to a final concentration of 1M
- Protein will immediately precipitate; cover with foil and leave in the dark for 25 hours at RTP. (Solution will turn bright pink as pyrromethane has been cleaved)
- Collect protein by centrifugation (20 000 g for 10 minutes)
- Resuspended protein in 45 ml of Buffer 1 and added to dialysis tubing & dialysed against same buffer for 90 minutes at RTP. (Stirring required)
- Dialysis tube transferred to 5L bottle of Buffer 2 and continues dialysis at 5°C for 24 hours. Exchange buffer minimum of 3 times (no stirring)
- Dialysis exchanged to Buffer 3 for overnight dialysis (with stirring) at 5°C. Exchange buffer minimum of 3 times
- Collect sample from dialysis tubing and keep on ice.
- Concentrate protein using Amicon® filter tubes
    - Important – exchange samples in order to prevent protein precipitation. Then concentrate final sample volume
- Determine protein concentration (Appendix 16)

Appendix 21 – '*Apo*'- Enzyme Reconstitution Assay

All equipment and reagents are as detailed in Appendix 17, apart from these mentioned below:

Reagents

- 0.5 mM PBG (substrate) (1.22 mg/ml)
- 0.5 mM DPM cofactor (1.69 mg/2ml)
- 50% MeOH
- 1M imidazole solution
- 0.5 mg/ml PBGD enzyme solution

Methodology

Standard

As detailed in Appendix 17.

Assay

- Make up 0.5 mM PBG in Assay Buffer; keep on ice
- Make up 0.5 mM DPM cofactor in 50% Meth-OH and 50% imidazole solution; keep on ice
- Volumes for the assay as follows: (as an example)
    - 920 µl Assay Buffer
    - 10 µl DTT
    - 20 µl PBG (or DPM)
    - 50 µl PBGD
- Perform Assay as detailed in Appendix 17

Appendix 22 – PD10- Column Buffer Exchange

Equipment:

- PD-10 Desalting Columns (GE South Africa Pty Ltd, Midrand, South Africa)
- Small sterile glass test tubes (12 x 75 mm) (B&M Scientific, Cape Town, South Africa)

Reagents:

- 0.2% sodium azide solution
- CD Buffer (Appendix 23)
- 2.5 ml Purified PBGD enzyme

Methodology:

- Decant storage solution in PD 10 column
- Equilibrate column with minimum of 25 ml of desired final buffer (i.e. CD Buffer)

- Aliquot 2.5 ml of purified protein.
    - In case of mutants with smaller amount yield, aliquot as much of protein as possible and make-up to a final volume of 2.5 ml with equilibration buffer (Appendix 13)

- Add the 2.5 ml of protein and allow to flow through
    - Wait until there is no further flow through

- Measure 3 ml of CD buffer & add to column in order to elute protein
- Store at 4°C
- Determine protein concentration (Appendix 16)

Appendix 23 – Circular Dichroism (Far UV & Near UV) *from Bustad et al. (2013)*[96]

Equipment:

- 106 QS, 0.5 mm P1X Cuvette (200 µl) (CAF, Stellenbosch University)
- Chirascan Plus CD Spectrophotometer (Applied Photophysics Limited, United Kingdom)
- PCS.3 Single Cell Peltier Temperature Controller (Applied Photophysics Limited, United Kingdom)
- Soft Kleenex Tissues

Software:

- Chirascan Windows™ Software (Applied Photophysics Limited, United Kingdom)
- Pro-Data Viewer Version 4.0 (Applied Photophysics Limited, United Kingdom)

Reagents:

- **CD Buffer (pH 8.2)**
  - 10 mM KPO_4
  - 100 mM $(NH_4)_2\,SO_4$
- Distilled H_2O
- Acetone
- Conc. HCl

Methodology

- Start-up CD machine, and allow 15 seconds before switching on the lamp
- Start-up and set the temperature controller to 20°C
- Open the nitrogen gas valve and ensure flow-meter gas inlet as follows:
  - sealed lamp housing: requiring 1 l/min
  - sealed monochromator: requiring 3 l/min (or 1 l/min when working > 200 nm)
  - sealed light path within the sample chamber: requiring 1 l/min (or zero purge > 200 nm)

- Start up the lamp and allow minimum of **15 minutes for lamp to equilibrate**
- Dilute protein samples to required concentration using CD Buffer (7 µM) and keep on ice
- Wash cuvette with acetone to remove and residual sample or contaminants.
- Wash cuvette with distilled H_2O a minimum of 3 times
- Rinse cuvette with CD buffer multiple times
- Start the Chirascan software program
  - Set the wavelength range to start at 300 nm and end at 190 nm.
  - Set machine to read 'millidegrees' and not absorbance
  - Select 'adaptive sampling'
  - Select 'spectrum'
- Start Pro Data Viewer
  - Preferences > 'Go online'
  - Set the folder of choice as the working directory

- Aliquot CD buffer into cuvette and read a scan for the wavelength. (Blank)

*Note: if the blank reading shows peaks, specifically around 222 and 190 nm – there is possible *residual sample left in the cuvette. Soak cuvette with acid for 2-5 minutes and repeat cuvette washing.*

**Should also be visible if y-axis is changes to absorbance – value should not exceed 2.0 A.*

- Aliquot sample and scan for far-UV range at 25°C.
- Rinse cuvette & repeat. (ensure concentration is uniform – check absorbance levels)

Note: more information on the specifications of the Chirascan plus, accessories and manuals can be found on the Applied Photophysics website [https://www.photophysics.com] as well as on the Central Analytical Facility website [http://academic.sun.ac.za/saf/]

Similar set up as Appendix 23

Methodology:

- Set the wavelength to measure at 222 nm

- Set the temperature (under Advanced setting)
- Allow the temperature range from 20 - 90°C
    - Set the interval in 5°C
- Following temperature set machine to return to starting temperature (20°C)

- Enable 'stepped ramping'
    - Rationale – the temperature monitor of the sample holder heats up faster than the solution (i.e. true temperature measured by thermometer) hence allowing for stepped ramping ensures the true sample temperature is as close as possible to the sample holder
    - Allow setting time of 300 seconds maximum.
- Before attempting the temperature scan – perform a quick far UV scan to ensure protein is correct concentration as well as checking for contaminants (strange signal etc.)

Equipment:

- Bio-Rad® Gel Casting Kit (Courtesy of Department of Surgery)
- PS1500 DC Power Supply (Hoefer Scientific Instruments, San Francisco, USA)
- UVItec gel documentation system (UVItec Limited, Cambridge, UK)
- K65HM high-density synthetic paper (Whitehead Scientific, South Africa)

Reagents:

- Running gel buffer: 1.5M Tris-HCl, pH 8.8
- Stacking gel buffer: 0.5M Tris-HCl, pH 6.8

- 10 % Running gel
  - 10 ml A-Bis-A (30% acrylamide, 0,8% bis-acrylamide)
  - 12 ml H_2O
  - 7.5 ml Running Buffer
  - 20 µl TEMED
  - 150 µl 10% ammonium persulphate (fresh)

- 4 % Stacking gel
  - 1.33 ml A-Bis-A
  - 6.1 ml H_2O
  - 2.5 ml Stacking buffer
  - 10 µl TEMED
  - 50 µl 10% ammonium persulphate (fresh)

- Running the gel
  - Tank buffer: 0.025M Tris-HCl, 0.192M Glycine, pH 8.3
  - Sample buffer (5 mL, w/o denaturing agents) (store in aliquots at −20 °C):
    - 125 mM Tris/HCl (0.1514 g)
    - 2.5 ml Stacking gel buffer
    - 10 % v/v glycerol (1.15 mL of 87 % glycerol)
    - Very little (0.002 %) Bromphenol Blue A few grains
    - H2O to 5 mL
- Do not boil samples (will cause denaturation)
- Load between 5 – 10 µg of protein
- Run gel at 150 V and 20 mA (approximately 45 minutes)
- Stain (Appendix 15)
- Visualize using UVItec gel documentation system

References

1 Nilsson, R., Schultz, I. J., Pierce, E. L., Soltis, K. a, Naranuntarat, A., Ward, D. M., Baughman, J. M., Paradkar, P. N., Kingsley, P. D., Culotta, V. C., et al. (2009) Discovery of Genes Essential for Heme Biosynthesis through Large-Scale Gene Expression Analysis. Cell Metab., Elsevier Ltd **10**, 119–130.

2 Karim, Z., Lyoumi, S., Nicolas, G., Deybach, J.-C., Gouya, L. and Puy, H. (2015) Porphyrias: A 2015 update. Clin. Res. Hepatol. Gastroenterol., Elsevier Masson SAS **39**, 412–425.

3 Ajioka, R. S., Phillips, J. D. and Kushner, J. P. (2006) Biosynthesis of heme in mammals. Biochim. Biophys. Acta **1763**, 723–736.

4 Furuyama, K., Kaneko, K. and Vargas, P. D. (2007) Heme as a Magnificent Molecule with Multiple Missions: Heme Determines Its Own Fate and Governs Cellular Homeostasis. Tohoku J. Exp. Med. **213**, 1–16.

5 ChemSketch Version 2.0. (2015) Advanced Chemistry Development, Inc., Toronto, On, Canada. www.acdlabs.com.

6 Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., Muthukrishnan, V., Owen, G., Turner, S., Williams, M., et al. (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acid Res. **41**, D456–D463.

7 Zhang, J., Kang, Z., Chen, J. and Du, G. (2015) Optimization of the heme biosynthesis pathway for the production of 5-aminolevulinic acid in Escherichia coli. Nat. Sci. Reports **5**, 1–7.

8 Franken, A. C. W., Lokman, B. C., Ram, A. F. J., Punt, P. J., van den Hondel, C. a M. J. J. and de Weert, S. (2011) Heme biosynthesis and its regulation: towards understanding and improvement of heme biosynthesis in filamentous fungi. Appl. Microbiol. Biotechnol. **91**, 447–460.

9 Hoffman, M., Góra, M. and Rytka, J. (2003) Identification of rate-limiting steps in yeast heme biosynthesis. Biochem. Biophys. Res. Commun. **310**, 1247–1253.

10 Tsiftsoglou, A. S., Tsamadou, A. I. and Papadopoulou, L. C. (2006) Heme as key regulator of major mammalian cellular functions: Molecular, cellular, and pharmacological aspects. Pharmacol. Ther. **111**, 327–345.

11 Dailey, H. a. and Meissner, P. N. (2013) Erythroid Heme Biosynthesis and Its Disorders. In Cold Spring Harbor perspectives in medicine (Weatherall, D., Schecter, A. N., and Nathan, David, G., eds.), pp 1–18, Cold Spring Harbor Laboratory Press.

12 Layer, G., Reichelt, J., Jahn, D. and Heinz, D. W. (2010) Structure and function of enzymes in heme biosynthesis. Protein Sci. **19**, 1137–1161.

13 Deybach, J. C. and Puy, H. (2003) Acute Intermittent Porphyria: From Clinical to Molecular Aspects. In Porphyrin Handbook (Kadish, K., Smith, K., and Guilard, R., eds.), pp 23–41, Elsevier Science, San Diego.

14 Hamza, I. and Dailey, H. A. (2012) One ring to rule them all: Trafficking of heme and heme synthesis intermediates in the metazoans. Biochim. Biophys. Acta, Elsevier B.V. **1823**, 1617–1632.

15 Koch, M., Breithaupt, C., Kiefersauer, R., Freigang, J., Huber, R. and Messerschmidt, A. (2004) Crystal structure of protoporphyrinogen IX oxidase: a key enzyme in haem and chlorophyll biosynthesis. EMBO J. **23**, 1720–1728.

16 Handschin, C., Lin, J., Rhee, J., Peyer, A.-K., Chin, S., Wu, P.-H., Meyer, U. A. and Spiegelman, B. M. (2005) Nutritional Regulation of Hepatic Heme Biosynthesis and Porphyria through PGC-1alpha. Cell **122**, 505–515.

17 Thunell, S. (2006) Genomic Approach to Acute Porphyria. Physiol. Res. **55**, S43–S66.

18 Homedan, C., Laafi, J., Schmitt, C., Gueguen, N., Lefebvre, T., Karim, Z., Desquiret-Dumas, V., Wetterwald, C., Deybach, J.-C., Gouya, L., et al. (2014) Acute intermittent porphyria causes hepatic mitochondrial energetic failure in a mouse model. Int. J.

Biochem. Cell Biol., Elsevier Ltd **51**, 93–101.

19 Roberts, A. G. and Elder, G. H. (2001) Alternative splicing and tissue-specific transcription of human and rodent ubiquitous 5-aminolevulinate synthase (ALAS1) genes. Biochim. Biophys. Acta **1518**, 95–105.

20 Hofer, T., Wenger, R. H., Kramer, M. F., Ferreira, G. C. and Gassmann, M. (2003) Hypoxic up-regulation of erythroid 5-aminolevulinate synthase. Blood **101**, 348–350.

21 Elder, G., Harper, P., Badminton, M., Sandberg, S. and Deybach, J. C. (2013) The incidence of inherited porphyrias in Europe. J. Inherit. Metab. Dis. **36**, 849–857.

22 Meissner, P. N., Corrigall, A. V. and Hift, R. J. (2012) Fifty years of porphyria at the University of Cape Town. South African Med. J. **102**, 422–426.

23 Puy, H., Gouya, L. and Deybach, J.-C. (2010) Porphyrias. Lancet **375**, 924–937.

24 Chemmanur, A. T. and Bonkovsky, H. L. (2004) Hepatic porphyrias: diagnosis and management. Clin. Liver Dis. **8**, 807–838.

25 Besur, S., Schmeltzer, P. and Bonkovsky, H. L. (2015) Acute Porphyrias. J. Emerg. Med., Elsevier Ltd **49**, 305–312.

26 Nordmann, Y. and Puy, H. (2002) Human hereditary hepatic porphyrias. Clin. Chim. Acta. **325**, 17–37.

27 Puy, H., Deybach, J. C., Lamoril, J., Robreau, A. M., Da Silva, V., Gouya, L., Grandchamp, B. and Nordmann, Y. (1997) Molecular Epidemiology and Diagnosis of PBG Deaminase Gene Defects in Acute Intermittent Porphyria. Am. J. Hum. Genet. **60**, 1373–1383.

28 Herrick, A. L. and McColl, K. E. L. (2005) Acute intermittent porphyria. Best Pract. Res. Clin. Gastroenterol. **19**, 235–249.

29 Daniell, W. E., Stockbridge, H. L., Labbe, R. F., Woods, J. S., Anderson, K. E., Bissell, D. M., Bloomer, J. R., Ellefson, R. D., Moore, M. R., Pierach, C. A., et al. (1997) Environmental chemical exposures and disturbances of heme synthesis. Environ. Health Perspect. **105 Suppl**, 37–53.

30 Singal, A. K., Parker, C., Bowden, C., Thapar, M., Liu, L. and McGuire, B. M. (2014) Liver Transplantation in the Management of Porphyria. Hepatology **60**, 1082–1089.

31 Unzu, C., Sampedro, A., Mauleón, I., Vanrell, L., Dubrot, J., de Salamanca, R. E., González-Aseguinolaza, G., Melero, I., Prieto, J. and Fontanellas, A. (2010) Porphobilinogen deaminase over-expression in hepatocytes, but not in erythrocytes, prevents accumulation of toxic porphyrin precursors in a mouse model of acute intermittent porphyria. J. Hepatol., European Association for the Study of the Liver **52**, 417–424.

32 Balwani, M. and Desnick, R. J. (2012) The porphyrias: advances in diagnosis and treatment. Blood **120**, 4496–4504.

33 Siegesmund, M., van Tuyll van Serooskerken, A.-M., Poblete-Gutiérrez, P. and Frank, J. (2010) The acute hepatic porphyrias : Current status and future challenges. Best Pract. Res. Clin. Gastroenterol., Elsevier Ltd **24**, 593–605.

34 Elder, G. H. (1998) Update on enzyme and molecular defects in porphyria. Photodermatol. Photoimmunol. Photomed. **14**, 66–69.

35 Hindmarsh, J. T., Oliveras, L. and Greenway, D. C. (1999) Biochemical Differentiation of the Porphyrias. Clin. Biochem. **32**, 609–619.

36 Anderson, K. E., Sassa, S. S., Bishop, D. F. and Desnick, R. J. (2001) Disorders of Heme Biosynthesis: X-Linked Sideroblastic Anemia and the Porphyrias. In The metabolic and molecular bases of inherited disease. (Scriver, C., Beaudet, A., Sly, W., and Valle, D., eds.) 8th ed., pp 2991–3062, McGraw-Hill, New York.

37 Mathews, M., Paish, L., Kumar, A., Khan, N. and Maroules, M. (2015) Hepatocellular Carcinoma in Acute Intermittent Porphyria: Incidental or Expected? Am. J. Med. Case Reports **3**, 141–143.

38 Trier, H., Krishnasamy, V. P. and Kasi, P. M. (2013) Case Report Clinical Manifestations

and Diagnostic Challenges in Acute Porphyrias. Case Rep. Hematol., Hindawi Publishing Coorportaion **2013**, 1–6.

39 Pozo, O. J., Marcos, J., Fabregat, A., Ventura, R., Casals, G., Aguilera, P., Segura, J. and To-Figueras, J. (2014) Adrenal hormonal imbalance in acute intermittent porphyria patients: results of a case control study. Orphanet J. Rare Dis. **9**, 1–11.

40 Ulbrichova, D., Flachsova, E., Hrdinka, M., Saligova, J., Bazar, J., Raman, C. S. and Martasek, P. (2006) De Novo mutation found in the porphobilinogen deaminase gene in Slovak acute intermittent porphyria patient: molecular biochemical study. Physiol. Res. **55**, S145–S154.

41 Kotze, M. J., De Villiers, J. N., Groenewald, J. Z., Rooney, R. N., Loubser, O., Thiart, R., Oosthuizen, C. J., van Niekerk, M. M., Groenewald, I. M., Retief, a E., et al. (1998) Molecular analysis reveals a high mutation frequency in the first untranslated exon of the PPOX gene and largely excludes variegate porphyria in a subset of clinically affected Afrikaner families. Mol. Cell. Probes **12**, 293–300.

42 Sasaki, H., Kaneko, K., Tsuneyama, H., Daimon, M., Yamatani, K. and Manaka, H. (1996) Family study of acute intermittent porphyria and hereditary coproporphyria in Niigata and Akita Prefectures, Japan. J. Clin. Epidemiol. **49**, 1117–1123.

43 Ulbrichova, D., Hrdinka, M., Saudek, V. and Martasek, P. (2009) Acute intermittent porphyria - impact of mutations found in the hydroxymethylbilane synthase gene on biochemical and enzymatic protein properties. FEBS J. **276**, 2106–2115.

44 Kauppinen, R. and von und zu Fraunberg, M. (2002) Molecular and biochemical studies of acute intermittent porphyria in 196 patients and their families. Clin. Chem. **48**, 1891–1900.

45 von Brasch, L., Zang, C., Haverkamp, T., Schlechte, H., Heckers, H. and Petrides, P. E. (2004) Molecular analysis of acute intermittent porphyria: mutation screening in 20 patients in Germany reveals 11 novel mutations. Blood Cells. Mol. Dis. **32**, 309–314.

46 Floderus, Y., Shoolingin-Jordan, P. M. and Harper, P. (2002) Acute intermittent porphyria in Sweden. Molecular, functional and clinical consequences of some new mutations found in the porphobilinogen deaminase gene. Clin. Genet. **62**, 288–297.

47 Ong, P. M., Lanyon, W. G., Hift, R. J., Halkett, J., Moore, M. R., Mgone, C. S. and Connor, J. M. (1996) Detection of four mutations in six unrelated South African patients with acute intermittent porphyria. Mol. Cell. Probes **10**, 57–61.

48 Li, N., Chu, X., Wu, L., Liu, X. and Li, D. (2008) Functional studies of rat hydroxymethylbilane synthase. Bioorg. Chem. **36**, 241–251.

49 Gregor, A., Schneider-Yin, X., Szlendak, U., Wettstein, A., Lipniacka, A., Rüfenacht, U. B. and Minder, E. I. (2002) Molecular Study of the Hydroxymethlybilane Synthase Gene (HMBS) Among Polish Patients With Acute Intermittent Porphyria. Hum. Mutat. **491**, 1–5.

50 Lee, J. S. and Anvret, M. (1991) Identification of the most common mutation within the porphobilinogen deaminase gene in Swedish patients with acute intermittent porphyria. Proc. Natl. Acad. Sci. U. S. A. **88**, 10912–10915.

51 Stenson, P. D., Mort, M., Ball, E. V, Howells, K., Phillips, A. D., Thomas, N. S. and Cooper, D. N. (2009) The Human Gene Mutation Database: 2008 update. Genome Med. **1**, 13.

52 Raich, N., Romeo, P. H., Dubart, A., Beaupain, A., Cohen-Solal, M. and Goossens, M. (1986) Molecular cloning and complete primary sequence of human erythrocyte porphobilinogen deaminase. Nucleic Acids Res. **14**, 7453–7472.

53 Grandchamp, B., Picat, C., Mignotte, V., Wilson, J. H., Te Velde, K., Sandkuyl, L., Roméo, P. H., Goossens, M. and Nordmann, Y. (1989) Tissue-specific splicing mutation in acute intermittent porphyria. Proc. Natl. Acad. Sci. U. S. A. **86**, 661–664.

54 Chretien, S., Dubart, a, Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M. and Romeo, P. H. (1988) Alternative transcription and splicing of the human

porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. Proc. Natl. Acad. Sci. U. S. A. **85**, 6–10.

55 Namba, H., Narahara, K., Tsuji, K., Yokoyama, Y. and Seino, Y. (1991) Assignment of human porphobilinogen deaminase to 11q24.1→q24.2 by in situ hybridization and gene dosage studies. Cytogenet. Genome Res. **57**, 105–108.

56 Gubin, A. N. and Miller, J. L. (2001) Human erythroid porphobilinogen deaminase exists in 2 splice variants. Blood **97**, 815–817.

57 Elder, G. H. (1998) Genetic Defects in the Porphyrias: Types and Significance. Clin. Dermatol. **16**, 225–233.

58 Warren, M. J. and Jordan, P. M. (1988) Investigation into the Nature of Substrate Binding to the Dipyrromethane Cofactor of Escherichia coli Porphobilinogen Deaminase. Biochemistry **25**, 9020–9030.

59 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) The three-dimensional structures of mutants of porphobilinogen deaminase: toward an understanding of the structural basis of acute intermittent porphyria. Protein Sci. **3**, 1644–1650.

60 Gill, R., Kolstoe, S. E., Mohammed, F., Al D-Bass, A., Mosely, J. E., Sarwar, M., Cooper, J. B., Wood, S. P. and Shoolingin-Jordan, P. M. (2009) Structure of human porphobilinogen deaminase at 2.8Å: the molecular basis of acute intermittent porphyria. Biochem. J. **420**, 17–25.

61 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605–1612.

62 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) Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. Nature **359**, 33–39.

63 Helliwell, J. R. J., Nieh, Y. Y.-P., Raftery, J., Cassetta, A., Habash, J., Carr, P. D., Ursby, T., Wulff, M., Thompson, A. W., Niemann, A. C., et al. (1998) Time-resolved structures of hydroxymethylbilane synthase (Lys59Gln mutant) as it is loaded with substrate in the crystal determined by Laue diffraction. J. Chem. Soc. Faraday Trans. **94**, 2615–2622.

64 Nieh, Y. P., Raftery, J., Weisgerber, S., Habash, J., Schotte, F., Ursby, T., Wulff, M., Hädener, A., Campbell, J. W., Hao, Q., et al. (1999) Accurate and highly complete synchrotron protein crystal Laue diffraction data using the ESRF CCD and the Daresbury Laue software. J. Synchrotron Radiat. **6**, 995–1006.

65 Hädener, A., Matzinger, P. K., Battersby, A. R., McSweeney, S., Thompson, A. W., Hammersley, A. P., Harrop, S. J., Cassetta, A., Deacon, A., Hunter, W. N., et al. (1999) Determination of the structure of selenomethionine-labelled hydroxymethylbilane synthase in its active form by multi-wavelength anomalous dispersion. Acta Crystallogr. Sect. D Biol. Crystallogr. **55**, 631–643.

66 Helliwell, J. R., Nieh, Y. P., Habash, J., Faulder, P. F., Raftery, J., Cianci, M., Wulff, M. and Hädener, A. (2003) Time-resolved and static-ensemble structural chemistry of hydroxymethylbilane synthase. Faraday Discuss. **122**, 131–144.

67 Song, G., Li, Y., Cheng, C., Zhao, Y., Gao, A., Zhang, R., Joachimiak, A., Shaw, N. and Liu, Z.-J. (2009) Structural insight into acute intermittent porphyria. FASEB J. **23**, 396–404.

68 Roberts, A., Gill, R., Hussey, R. J., Mikolajek, H., Erskine, P. T., Cooper, J. B., Wood, S. P., Chrystal, E. J. T. and Shoolingin-Jordan, P. M. (2013) Insights into the mechanism of pyrrole polymerization catalysed by porphobilinogen deaminase: high-resolution X-ray studies of the Arabidopsis thaliana enzyme. Acta Crystallogr. D. Biol. Crystallogr., International Union of Crystallography **D69**, 471–485.

69 Roberts, A., Gill, R., Hussey, R. J., Mikolajek, H., Erskine, P. T., Cooper, J. B., Wood, S. P., Chrystal, E. J. T. and Shoolingin-Jordan, P. M. (2012) Crystallization and

preliminary X-ray characterization of the tetrapyrrole-biosynthetic enzyme porphobilinogen deaminase from Arabidopsis thaliana. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun., International Union of Crystallography **F68**, 1491–1493.

70 Azim, N., Deery, E., Wolfenden, B. A. A., Erskine, P., Cooper, J. B., Coker, A., Wood, S. P. and Akhtar, M. (2014) Structural evidence for the partially oxidized dipyrromethene and dipyrromethanone forms of the cofactor of porphobilinogen deaminase: structures of the Bacillus megaterium enzyme at near-atomic resolution. Acta Crystallogr. D. Biol. Crystallogr., International Union of Crystallography **D70**, 744–751.

71 Azim, N., Deery, E., Warren, M. J., Erskine, P., Cooper, J. B., Wood, S. P. and Akhtar, M. (2013) Crystallization and preliminary X-ray characterization of the tetrapyrrole-biosynthetic enzyme porphobilinogen deaminase from Bacillus megaterium. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun., International Union of Crystallography **F69**, 906–908.

72 Louie, G. V, Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Malashkevich, V. N., Hädener, A., Warren, M. J. and Shoolingin-Jordan, P. M. (1996) The Three-Dimensional Structure of Escherichia coli Porphobilinogen Deaminase at 1.76-A Resolution. PROTEINS Struct. Funct. Genet. **25**, 48–78.

73 Awan, S. J., Siligardi, G., Shoolingin-Jordan, P. M. and Warren, M. J. (1997) Reconstitution of the holoenzyme form of Escherichia coli porphobilinogen deaminase from apoenzyme with porphobilinogen and preuroporphyrinogen: A study using circular dichroism spectroscopy. Biochemistry **36**, 9273–9282.

74 Jordan, P. M. and Woodcock, S. C. (1991) Mutagenesis of arginine residues in the catalytic cleft of Escherichia coli porphobilinogen deaminase that affects dipyrromethane cofactor assembly and tetrapyrrole chain initiation and elongation. Biochem. J. **280**, 445–449.

75 Bung, N., Pradhan, M., Srinivasan, H. and Bulusu, G. (2014) Structural insights into E. coli porphobilinogen deaminase during synthesis and exit of 1-hydroxymethylbilane. PLoS Comput. Biol. **10**, 1–16.

76 Jordan, P. M. and Warren, M. J. (1987) Evidence for a dipyrromethane cofactor at the catalytic site of E. coli porphobilinogen deaminase. FEBS Lett. **225**, 87–92.

77 Hart, G. J., Miller, A. D., Leeper, F. J. and Battersby, A. R. (1987) Biosynthesis of the Natural Porphyrins: Proof that Hydroxymethylbilane Synthase (Porphobilinogen Deaminase) uses a Novel Binding Group in its Catalytic Action. J. Chem. Soc. Chem. Commun. 1762–1765.

78 Hart, G. J., Miller, A. D. and Battersby, A. R. (1988) Evidence that the pyrromethane cofactor of hydroxymethylbilane synthase (porphobilinogen deaminase) is bound through the sulphur atom of a cysteine residue. Biochem. J. **252**, 909–912.

79 Jordan, P. M., Warren, M. J., Williams, H. J., Stolowich, N. J., Roessner, C. A., Grant, S. K. and Scott, A. I. (1988) Identification of a cysteine residue as the binding site for the dipyrromethane cofactor at the active site of Escherichia coli porphobilinogen deaminase. FEBS Lett. **235**, 189–193.

80 Scott, A. I., Rossnerr, C. A., Stolowich, N. J., Karuso, H. J., Williams, H. J., Grant, S. K., Gonzalez, M. D. and Hoshino, T. (1988) Site-Directed Mutagenesis and High-Resolution NMR Spectroscopy of the Active Site of Porphobilinogen Deaminase. Biochemistry **27**, 7985–7990.

81 Scott, A. I., Clemens, K. R., Stolowich, N. J., Santander, P. J., Gonzalez, M. D. and Roessner, C. A. (1989) Reconstitution of apo-porphobilinogen deaminase: structural changes induced by cofactor binding. FEBS Lett. **242**, 319–324.

82 Shoolingin-Jordan, P. M., Warren, M. J. and Awan, S. J. (1996) Discovery that the assembly of the dipyrromethane cofactor of porphobilinogen deaminase holoenzyme proceeds initially by the reaction of preuroporphyrinogen with the apoenzyme. Biochem. J. **376**, 373–376.

83 Shoolingin-Jordan, P. M., Al-Dbass, A., McNeill, L. A., Sarwar, M. and Butler, D. (2003)

Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. Biochem. Soc. Trans. **31**, 731–735.

84 Cole, D. J., Tirado-Rives, J. and Jorgensen, W. L. (2015) Molecular dynamics and Monte Carlo simulations for protein-ligand binding and inhibitor design. Biochim. Biophys. Acta **1850**, 966–971.

85 Dayan, F. E., Daga, P. R., Duke, S. O., Lee, R. M., Tranel, P. J. and Doerksen, R. J. (2010) Biochemical and structural consequences of a glycine deletion in the alpha-8 helix of protoporphyrinogen oxidase. Biochim. Biophys. Acta, Elsevier B.V. **1804**, 1548–1556.

86 Jones, R. M. and Jordan, P. M. (1994) Purification and properties of porphobilinogen deaminase from Arabidopsis thaliana. Biochem. J. **299**, 895–902.

87 Tsaniras, S. C., Fakourelis, P., Filintisi, A., Polychronidou, E. and Papageorgiou, L. (2015) Insights into 3D molecular modelling methodologies. J. Struct. Bioinforma. **1**, 1–7.

88 Keskin, O., Jernigan, R. L. and Bahar, I. (2000) Proteins with similar architecture exhibit similar large-scale dynamic behavior. Biophys. J., Elsevier **78**, 2093–106.

89 Cappellini, M. D., di Montemuros, F. M., Di Pierro, E. and Fiorelli, G. (2002) Hematologically Important Mutations: Acute Intermittent Porphyria. Blood Cells. Mol. Dis. **28**, 5–12.

90 Chi, R., Cheung, F., Wong, J. H. and Ng, T. B. (2012) Immobilized metal ion affinity chromatography : a review on its applications. Appl. Microbiol. Biotechnol. **96**, 1411–1420.

91 Chaga, G., Hopp, J. and Nelson, P. (1999) Immobilized metal ion affinity chromatography on Cobalt - carboxymethylaspartate – agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. Biotechnol. Appl. Biochem. **24**, 19–24.

92 Maneli, M. H., Corrigall, A. V., Klump, H. H., Davids, L. M., Kirsch, R. E. and Meissner, P. N. (2003) Kinetic and physical characterisation of recombinant wild-type and mutant human protoporphyrinogen oxidases. Biochim. Biophys. Acta - Proteins Proteomics **1650**, 10–21.

93 Bondos, S. E. and Bicknell, A. (2003) Detection and prevention of protein aggregation before, during, and after purification. Anal. Biochem. **316**, 223–231.

94 Wang, Y., Scott, C. R., Gelb, M. H. and Turecek, F. (2008) Direct assay of enzymes in heme biosynthesis for the detection of porphyrias by tandem mass spectrometry. Porphobilinogen deaminase. Anal. Chem. **80**, 2606–2611.

95 Akagi, R., Kato, N., Inoue, R., Anderson, K. E., Jave, E. K. and Sassa, S. (2006) δ-Aminolevulinate dehydratase (ALAD) porphyria: The first case in North America with two novel ALAD mutations. Mol. Genet. Metab. **87**, 329–336.

96 Bustad, H. J., Vorland, M., Rønneseth, E., Sandberg, S., Martinez, A. and Toska, K. (2013) Conformational stability and activity analysis of two hydroxymethylbilane synthase mutants, K132N and V215E, with different phenotypic association with acute intermittent porphyria. Biosci. Rep. **33**, 617–627.

97 Kelly, S. M., Jess, T. J. and Price, N. C. (2005) How to study proteins by circular dichroism. Biochim. Biophys. Acta - Proteins Proteomics **1751**, 119–139.

98 Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. Nat. Protoc. **1**, 2876–2890.

99 Parbhoo, N. (2013) Topologically conserved hydrophobic residues of the thioredoxin C-subdomain stabilise GSTs, University of the Witwatersrand.

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

101 van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R. and Groen, F.

C. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra. Anal. Biochem. **191**, 110–118.

102 Sreerama, N. and Woody, R. W. (1993) A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular Dichroism. Anal. Biochem. **209**, 32–44.

103 Sreerama, N., Venyaminov, S. Y. and Woody, R. W. (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. Protein Sci. **8**, 370–380.

104 Sreerama, N. and Woody, R. W. (2000) Estimation of Protein Secondary Structure from Circular Dichroism Spectra : Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. Anal. Biochem. **260**, 252–260.

105 Naidoo, V. B. and Rautenbach, M. (2013) Self-assembling organo-peptide bolaphiles with KLK tripeptide head groups display selective antibacterial activity. J. Pept. Sci. **19**, 784–791.

106 Böhm, G., Muhr, R. and Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. Protein Eng. Des. Sel. **5**, 191–195.

107 Agashe, V. R. and Udgaonkar, J. B. (1995) Thermodynamics of Denaturation of Barstar: Evidence for Cold Denaturation and Evaluation of the Interaction with Guanidine Hydrochloride. Biochemistry **34**, 3286–3299.

108 Greenfield, N. J. (2006) Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. Nat. Protoc. **1**, 2733–2741.

109 Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., et al. (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. Web Serv. Issue **40**, W597–W603.

110 Nielsen, M., Lundegaard, C., Lund, O. and Petersen, T. N. (2010) CPHmodels-3.0-remote homology modeling using structure-guided sequence profiles. Nucleic Acids Res. Web Serv. Issue **38**, W576–W581.

111 Boateng, M. O., Corrigall, A. V., Sturrock, E. and Meissner, P. N. (2015) Characterisation of the flavin adenine dinucleotide binding region of Myxococcus xanthus protoporphyrinogen oxidase. Biochem. Biophys. Reports, Elsevier **4**, 306–311.

112 Kabsch, W. and Sander, C. (1983) Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. Biopolymers **22**, 2577–2637.

113 Frishman, D. and Argos, P. (1995) Knowledge-based protein secondary structure assignment. Proteins Struct. Funct. Genet. **23**, 566–579.

114 Heinig, M. and Frishman, D. (2004) STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. Web Serv. Issue **32**, W500–W502.

115 Sreerama, N., Venyaminov, S. Y. and Woody, R. W. (2001) Analysis of Protein Circular Dichroism Spectra Based on the Tertiary Structure Classification. Anal. Biochem. **299**, 271–274.

116 Arakawa, T., Philo, J. S. and Kita, Y. (2001) Kinetic and thermodynamic analysis of thermal unfolding of recombinant erythropoietin. Biosci Biotechnol Biochem **65**, 1321–1327.

117 Hrdinka, M., Puy, H. and Martasek, P. (2006) May 2006 Update in Porphobilinogen Deaminase Gene Polymorphisms and Mutations Causing Acute Intermittent Porphyria. Comparison with the Situation in Slavic Population. Physiol. Res. **55**, S119–S136.

118 Anderson, P. and Desnick, R. (1982) Porphobilinogen deaminase: methods and principles of the enzymatic assay. Enzyme **23**, 146–157.

119 Chou, J.-H., Kosal, M. E., Nalwal, H. S., Rakow, N. A. and Suslick, K. S. (2000) Applications of Porphyrins and Metalloporphyrins to Materials Chemistry. In The Porphryin Handbook (Kadish, K. M., Smith, K. M., and Guilard, R., eds.) Volume 6., pp 44–128, Academic Press, Illinois, USA.

120 Meissner, P. N. (1990) Enzyme Studies in Variegate Porphyria, University of Cape Town.

121 Schomburg, I., Chang, A., Placzek, S., Söhngen, C., Rother, M., Lang, M., Munaretto, C., Ulas, S., Stelzer, M., Grote, A., et al. (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic acids Res. Database Issue **41**, D764–D772.

122 Piepkorn, M. W., Hamernyik, P. and Labbé, R. F. (1978) Modified erythrocyte uroporphyrinogen I synthase assay, and its clinical interpretation. Clin. Chem. **24**, 1751–4.

123 Meissner, P., Adams, P. and Kirsch, R. (1993) Allosteric inhibition of human lymphoblast and purified porphobilinogen deaminase by protoporphyrinogen and coproporphyrinogen. A possible mechanism for the acute attack of variegate porphyria. J. Clin. Invest. **91**, 1436–44.

124 Inoue, R. and Akagi, R. (2008) Co-synthesis of Human delta-Aminolevulinate Dehydratase (ALAD) Mutants with the Wild-type Enzyme in Cell-free System -Critical Importance of Conformation on Enzyme Activity. J. Clin. Biochem. Nutr. **43**, 143–153.

125 Lee, M. J., Kim, H.-J., Lee, J.-Y., Kwon, A. S., Jun, S. Y., Kang, S. H. and Kim, P. (2013) Effect of gene amplifications in porphyrin pathway on heme biosynthesis in a recombinant Escherichia coli. J. Microbiol. Biotechnol. **23**, 668–673.

126 Jover, R., Hoffmann, F., Scheffler-Koch, V. and Lindberg, R. L. P. (2000) Limited heme synthesis in porphobilinogen deaminase-deficient mice impairs transcriptional activation of specific cytochrome P450 genes by phenobarbital. Eur. J. Biochem. **267**, 7128–7137.

127 Dobson, C. M. (2004) Experimental investigation of protein folding and misfolding. Methods **34**, 4–14.

128 Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat. Protoc. **1**, 2527–2535.

129 Mach, H., Middaugh, C. R. and Lewis, R. V. (1992) Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. Anal. Biochem. **200**, 74–80.

130 Lim, B. T. and Kimura, T. (1980) Conformation-associated Anomalous Tyrosine Fluorescence of Adrenodoxin. J. Biol. Chem. **255**, 2440–2444.

131 Greenfield, N. J. (2006) Analysis of the kinetics of folding of proteins and peptides using circular dichroism. Nat. Protoc. **1**, 2891–2899.

132 Osysko, A. P. and Muíño, P. L. (2011) Fluorescence quenching of tryptophan and tryptophanyl dipeptides in solution. J. Biophys. Chem. **2**, 316–321.