Biochemical and biophysical characterisation of *HMBS* variants associated with acute intermittent porphyria

Marthe Christie Sæter

This thesis is submitted in partial fulfilment of the requirements for the degree of
Master of Science in Biomedical Sciences

Department of Biomedicine
University of Bergen
1st June 2022

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Dr Helene Bustad Johannessen, Dr Juha Kallio and Prof. Aurora Martinez for allowing me to take on this project in their lab. Thank you so much for all your guidance and support, and for sharing your knowledge with me. I have learned a great deal from all of you during this past year. Helene, thank you for your enthusiasm and encouragement, for always being available to help, I truly appreciate it. Without you this would not have been possible.

I would also like to thank the whole Biorecognition group, especially Dayne, Ming, and my fellow master students Gro and Inga Elise. I've appreciated this opportunity to work with all of you, a big thank you to everyone in Biorecognition, for your kindness, helpfulness, and for creating such a good work environment. I appreciate your financial support, together with BioCat, which gave me the opportunity to attend the NBS conference where I got to present my work and attend a structural biology course in Oslo.

I am also very thankful to my family and friends, for all their love and support these past few years. Thank you for always making me laugh, for all the adventures and fun times we've had over the years. I would like to express my gratitude towards my fellow graduate students and especially Adrianna and Patrick, my partners in crime, these last two years would not have been the same without you, I am forever grateful.

Table of Contents

List of Abbreviations

AIP	Acute intermittent porphyria
ALA	δ-Aminolaevulinic acid
CD	Circular dichroism
DSF	Differential light scattering
DTT	Dithiothreitol
DPM	Dipyrromethane
EDTA	Ethylenediamine tetraacetic acid
FT-ICR MS	Fourier transform ion cyclotron mass spectrometry
HMB	1-hydroxymethylbilane
HMBS	Hydroxymethylbilane synthase
IEX	Ion-exchange chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani broth
MD	Molecular dynamics
PAGE	Polyacrylamide gel electrophoresis
PBG	Porphobilinogen
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
TCEP	Tris (2-carboxyethyl) phosphine
TEV	Tobacco Etch Virus
URO	Uroporphyrinogen
XRD	X-ray diffraction

Summary

Hydroxymethylbilane synthase (HMBS), the third enzyme in the haem biosynthesis, binds four consecutive porphobilinogen (PBG) substrate molecules creating the enzyme intermediate complexes ES, ES_2, ES_3 and ES_4 to form the final product, 1-hydroxymethylbilane (HMB), which is then cleaved from the enzyme via hydrolysis. Mutations in the *HMBS* gene are associated with the autosomal dominant disorder acute intermittent porphyria (AIP). Individuals carrying an *HMBS* mutation express the functional wild-type (wt) enzyme from only one allele, which results in up to 50 % loss of enzyme activity. AIP is characterised by potentially life-threatening acute neurovisceral attacks, the penetration of the disease is low, and most carriers will never develop symptoms.

In this work, we sought to characterise the AIP disease-associated mutants p.R26H, p.L30P, p.K98R, p.D99N, p.R149Q, p.R167Q, p.R195C, p.E223K, p.L238P, p.E250Q, p.A252P and p.A339D in erythroid-specific HMBS, expressed in *E. coli* and purified. The strategy and methods selected for the characterization, in comparison with the wt-HMBS were a combination of anion-exchange chromatography, native PAGE, circular dichroism, differential scanning fluorimetry, high-resolution mass spectrometry, enzymatic activity assays and X-ray crystallography.

Out of the twelve disease-associated variants chosen for this project, we have obtained biochemical and biophysical characteristics for seven variants: p.R26H, p.L30P, p.D99N, p.R149Q, p.R167Q, p.R195C and p.E223K. Variant p.K98R was partly characterised while for the remaining four variants, p.L238P, p.E250Q, p.A252P and p.A339D, the expression and purification protocol resulted in insoluble and unstable apoenzyme with complete folding defects. We have shown that mutants p.K98R and p.R149Q both form a single enzyme intermediate in the apo-form while mutants p.R26H and p.D99N form a single intermediate in the ES_2-state. Mutants p.L30P, p.R167Q, p.R195C and p.E223K all display a distribution of enzyme intermediates resembling that of wt-HMBS and have both structural and catalytic effects on the enzyme. We demonstrated that p.D99N, p.R167Q and p.R195C resulted in some residual activity (~1 % of relative activity compared to wt-HMBS) whereas the remining mutants cause inactive enzymes unable to produce the final HMB product.

1 Introduction

Biosynthesis is defined by Merriam Webster as "the production of a chemical compound by a living organism" and the products produced by these various biosynthetic processes in the human body are vital for cellular and metabolic function and are thus essential for life. An example of one such process is the haem biosynthesis. Haem is an iron-containing prosthetic group in the porphyrin class, important for biological oxidations and vital for all aerobic organisms. It is most commonly known as a component of haemoglobin, myoglobin and cytochromes, and is important in a wide range of functions in the human body such as transportation and activation of oxygen, electron transportation and signal transduction [1-4].

1.1 The haem biosynthesis

In humans, the haem biosynthesis consists of eight enzymatic steps that take place in the cell cytosol and mitochondria, resulting in a porphyrin ring complex (Figure 1-1) [5, 6]. The first step in the haem biosynthesis starts in the mitochondria with the condensation of glycine and succinyl CoA, catalysed by δ-Aminolaevulinic (ALA) synthase, to ALA molecules [7, 8]. ALA synthase is the rate-limiting step. The ALA molecules exit the mitochondria and move into the cytosol where two ALA molecules are further catalysed by the enzyme ALA dehydratase forming a porphobilinogen (PBG) molecule [9]. Following the production of PBG, hydroxymethylbilane synthase (HMBS) catalyses the formation of a linear tetrapyrrole 1-hydroxymethylbilane (HMB) from four PBG molecules in consecutive steps [10]. HMB is then further modified by uroporphyrinogen synthase (UROS) and URO decarboxylase forming coproporphyrinogen, before it is transported back into the mitochondria where it is converted to protoporphyrinogen IX by coproporphyrinogen oxidase [11, 12]. The final step is ferrochelatase, which catalyses the insertion of a ferrous iron-forming haem [1, 13].

Each of the eight enzymatic steps is associated with metabolic inherited disorders in the porphyria group [14, 15], which is commonly divided into two major subgroups, the acute hepatic porphyrias, and the cutaneous porphyrias [13, 16, 17]. In this work, the protein of interest is the third enzyme in the haem biosynthesis pathway, namely HMBS. HMBS is primarily synthesised by erythrocytes in the bone marrow and hepatocytes in the liver. Mutations in the *HMBS* gene are associated with acute intermittent porphyria (AIP), the most common acute hepatic porphyria disorder, which is characterised by acute attacks of severe abdominal pain along with a range of gastrointestinal and psychiatric symptoms [6, 13, 18].
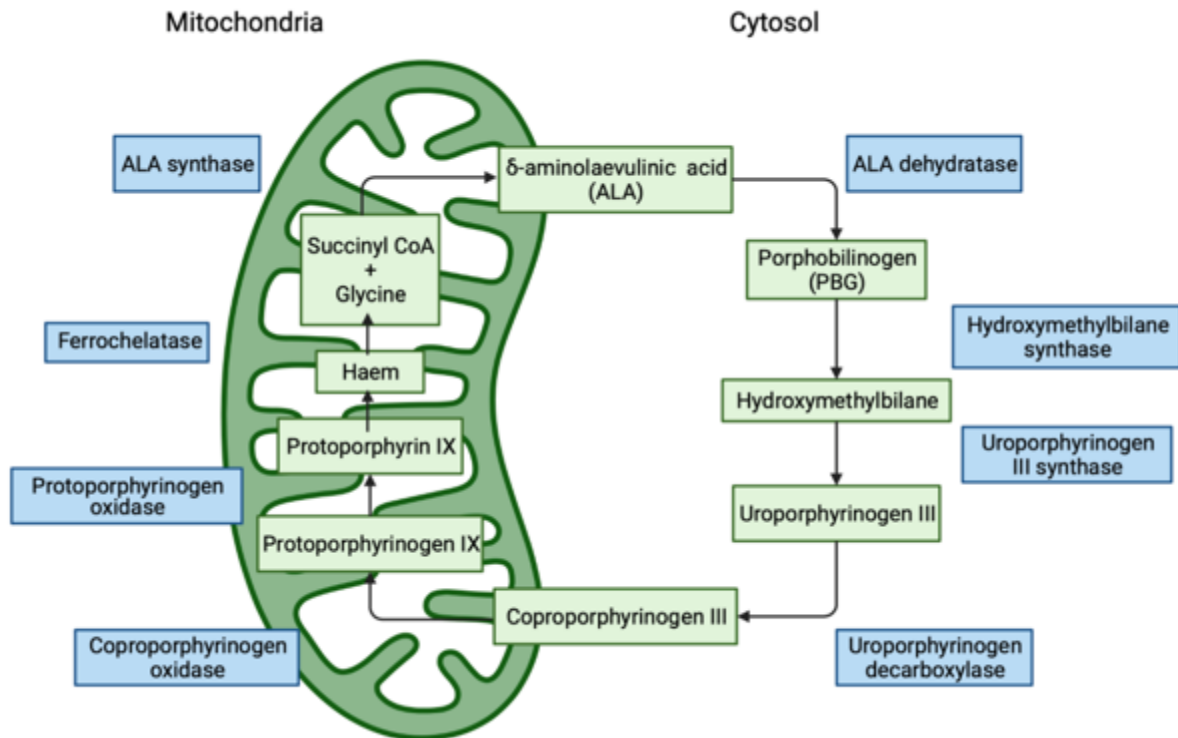
Figure 1-1. The haem biosynthesis. The haem biosynthesis starts in the mitochondria where succinyl CoA and glycine are condensed to form ALA. Continuing in the cytosol, two ALA molecules are catalysed by ALA dehydratase to form PBG. Four PBG molecules are catalysed to form HMB, which is further modified by uroporphyrinogen synthase and uroporphyrinogen decarboxylase forming coproporphyrinogen before it is transported back into the mitochondria. The last three steps are in the mitochondria, resulting in the formation of haem. See main text for further details. The enzymes involved in each step are represented in the blue boxes. Figure is modified from [6, 16].

1.2 Acute intermittent porphyria

Acute intermittent porphyria (AIP, OMIM #17600) is an autosomal dominant inherited disease caused by a mutation in the *HMBS* gene, on chromosome 11.q.24 [19]. The clinical penetrance of AIP is low and most people with a disease-causing variant will never experience any symptoms, indicating that one functioning allele providing 50 % of normal enzyme activity is sufficient [20, 21]. The occurrence of symptomatic AIP in the European population is estimated to be 0.5–1% [22-24]. Some patients will however suffer from recurrent acute attacks and chronic symptoms resulting in a high disease-burden and reduced quality-of-life [25, 26].

Acute AIP attacks can be precipitated by exposure to both endogenous and exogenous factors, an example of the latter is restriction of food intake [27, 28]. Under these circumstances the haem synthesis will be stimulated in the liver and induce cytochrome P450 synthesis, further leading to an increase in the porphyrin precursors ALA and PBG, thus

triggering an acute attack [24, 29]. Other exogenous factors include stress, consumption of alcohol and certain drugs [24]. Infection and inflammation are examples of endogenous triggering factors. For many, despite being exposed to these factors, the half-normal enzymatic activity is usually sufficient for relatively normal enzymatic function. AIP is more common in women than in men, and rare in children, with onset usually occurring after puberty [30-32]. This is likely related to the endogenous hormonal changes and sex steroids, which also act as inducers in the synthesis of haem [30].

Clinically, AIP is characterised by acute attacks of severe abdominal pain [33, 34]. Other symptoms range from nausea, constipation and vomiting, muscle pains and tachycardia to a wide range of neurological and psychiatric symptoms such as seizures, hallucinations, and confusion at variable frequencies [16, 35-38]. Due to the low penetrance of AIP, which results in a high number of asymptomatic carriers, in addition to the wide range of unspecific symptoms, many AIP cases go undiagnosed. Untreated or inappropriately treated AIP may result in serious acute attacks that potentially can be life-threatening and have severe long-term consequences such as respiratory failure, paralysis, seizures and cardiac arrythmia [14, 39]. Both symptomatic and asymptomatic individuals with a disease-causing mutant have an increased risk of developing long-term complications of AIP such as hypertension, increased risk of primary liver cancer and chronic kidney disease [40-42]. It is therefore important that patients diagnosed with AIP undergo regular check-ups such as ultrasound screenings of the liver twice a year after the patient turns 50 years old for early detection.

1.3   Current treatment options for AIP

The wide range of non-specific and heterogeneous symptoms in AIP makes diagnostics challenging [43]. However, due to the elevated levels of the porphyrin precursors ALA and PBG, measuring these in urine is the recommended and most common method of diagnosis [39, 44]. Following diagnosis, the patient is recommended to avoid triggering factors as a preventative measure. Additionally, family members are advised to undergo DNA sequencing to identify possible asymptomatic individuals [45, 46].

Management of acute attacks caused by AIP generally starts by identification and removal of triggering factors, followed by pain management, glucose and haem arginate infusions [47]. The haem arginate infusions are administered intravenously to downregulate ALA, thus reducing the accumulated ALA and PBG in the body and providing relief from acute attacks [6, 39, 48, 49]. Specific symptoms such as nausea and seizures can be treated

with anti-nausea medication and benzodiazepines, respectively. Acute AIP attacks can potentially be life-threatening and require hospitalisation.

For AIP patients who suffer from recurrent attacks, defined as four or more attacks per year, liver transplantation has up until recently been the only cure [50, 51]. Undergoing liver transplantation poses a substantial risk to the patient, with a wide range of complications, and should therefore only be considered as a last resort. The recently approved drug Givosiran, an RNA interference therapy developed to target the ALA synthase in order to reduce attacks in AIP patients, has therefore been welcomed [52]. In recent years, other emerging therapies have evolved using pharmacological chaperones, lipid nanoparticles, gene and enzyme replacement therapy [6]. In the last 20 years there have been several attempts of both enzyme and gene replacement therapies, where human HMBS has been directly administered, and there have been attempts where the correct HMBS DNA or mRNA sequence has been inserted into liver cells [53-56]. There have also been promising results obtained in the use of nanoparticles encapsulating mRNA for human HMBS resulting in a relief during acute attacks [57]. Currently, Cordoba et al. have explored the use of apolipoprotein A-I bound to HMBS in mouse models to correct the enzyme defects in HMBS [58]. If any of these methods prove successful, there will be a substantial improvement in the treatment of AIP.

1.4 Hydroxymethylbilane synthase

HMBS (EC:2.5.1.61, also known as porphobilinogen deaminase), is the third enzyme in the haem biosynthetic pathway, which catalyses the formation of the linear tetrapyrrole 1-hydroxymethylbilane (HMB). Four consecutive PBG molecules are bound in a stepwise condensation to the holoenzyme (E_{holo}), creating the enzyme-intermediate complexes ES, ES_2, ES_3 and ES_4 respectively, before the final product HMB is released [6, 13]

There are two isoforms of HMBS, arising from alternative splicing; the ubiquitous form expressed in all tissues (NM_000190.4) and the erythroid form (NM_001024382.2) [59, 60]. The erythroid HMBS is 344 amino acids long and has a molecular mass of 37 kDa whereas the ubiquitous form contains an additional 17 amino acid residues at the N-terminus and has a molecular mass of 39 kDa [59, 61]. Due to alternative splicing, an *HMBS* mutation located in the common part of the sequence, will affect both isoforms [60].

The first crystal structure of human HMBS was published in 2009 (PDB ID: 3ECR) [62] and provided valuable insight of the structure and function of HMBS. Several crystal

structures have been published since, as reviewed in [63]. Despite this, the catalytic and elongation mechanisms, as well as the function of the 17 additional amino acids at the N terminus in ubiquitous HMBS, are not yet fully understood [64, 65]. Human HMBS consists of three distinct domains; 1: residues 1 to 116 and 216 to 239, 2: residues 117 to 215 and 3: residues 240 to 361 [62]. The active site is located in the cleft between domain 1 and 2 where it binds the dipyrromethane (DPM) cofactor (Figure 1-2A) [66, 67]. The cofactor is formed by two PBG molecules, each consisting of a pyrrole ring, C1 and C2, and is covalently linked to Cys261 by the C1 ring [67, 68] where it acts as a primer in the elongation process [69]. For each substrate bound a pyrrole ring added and there is a change in charge, the substrate is protonated, and an amino group is released (Figure 1-2B). The final product, a tetrapyrrole HMB, is then cleaved via hydrolysis while the cofactor remains attached to the enzyme. The catalytic reaction starts with the cofactor binding to the apoenzyme creating the holoenzyme, before four PBG molecules are bound consecutively and creating the enzyme-intermediate complexes ES, ES_2, ES_3 and ES_4, respectively (Figure 1-2C). In the fourth step of the haem biosynthesis HMB is further synthesised into URO-III in the presence of UROS (Figure 1-1), however, if UROS is absent, HMB spontaneously rearranges into the non-physiological URO-I [70, 71].
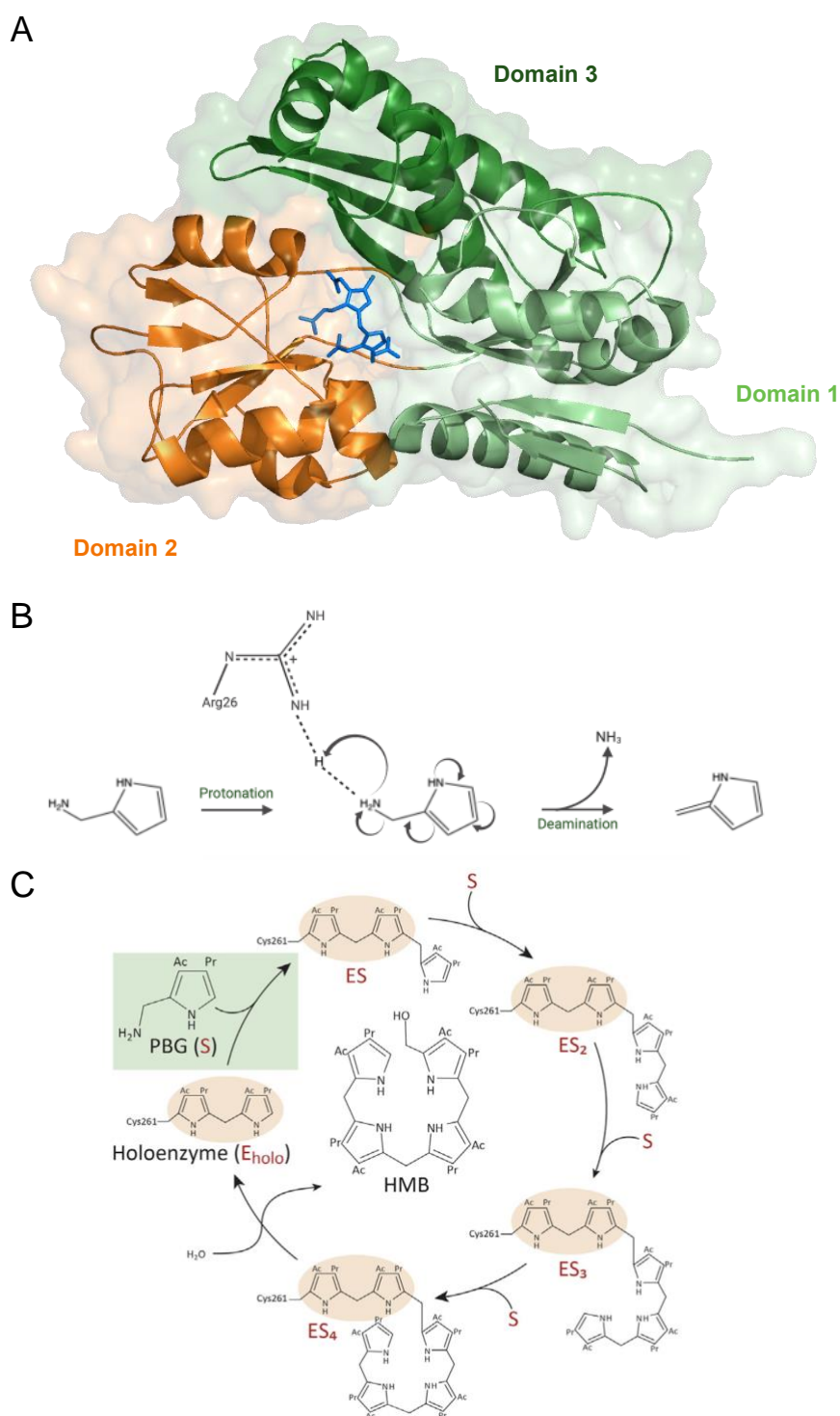
Figure 1-2. HMBS crystal structure and illustration of the polypyrrole elongation process. (A) A cartoon representation of the HMBS crystal structure with a transparent surface (PDB ID: 7AAJ*), showing the* dipyrromethane (DPM) cofactor (blue) and the three domains: domain 1 residues 1 to 116 and 216 to 239 (pale green), domain 2 residues 117 to 215 (orange) and domain 3 residues 240 to 361 (dark green). **(B)** Schematic diagram showing the mechanism of a single step in the elongation process. Figure modified from *[72]*. **(C)** The elongation process of HMBS starting with the holoenzyme followed by the binding of four subsequent PBG substrates (S) creating the enzyme-intermediate complexes ES, ES_2, ES_3 and ES_4 before releasing HMB. Figure reproduced from *[72]* under the terms of the Creative Commons CC-BY.

1.5   Variants in the *HMBS* gene

Currently, there are more than 500 different variants registered in the *HMBS* gene according to the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk, April 2022), 33 % of them are missense/nonsense variants, followed by splicing, small deletions and insertions [6]. Mutations located at or in close proximity to the substrate-binding and active site are more likely to alter binding of the cofactor or substrates and are therefore likely to result in changes in catalytic activity. Mutants located further out are more likely to affect the conformation/structure and folding of the enzyme.

The mutants vary in their degree of clinical penetrance and severity of symptoms, even within families affected by the same mutant. Due to its dominant inheritance, most AIP mutations are family specific, however, exceptions are found in cases where there has been a genetic founder effect [24, 73-75] or with mutants that are located in CpG dinucleotide hotspots [76-78]. For this project, we have selected missense variants from the Norwegian AIP population: p.R26H, p.L30P, p.K98R, p.R167Q, p.R195C, p.E223K, p.L238P, p.E250Q, p.A252P, and p.A339D in the ubiquitous HMBS isoform. In addition, we have selected two variants, p.D99N and p.R149Q, which are not present in the Norwegian AIP population, due to their catalytic importance. All twelve will be described in further detail below.

As previously mentioned, there are two isoforms of HMBS: the ubiquitous and erythroid-specific form. Based on previous reports the 17 extra amino acids in the ubiquitous enzyme do not affect the structural stability or activity and have no known functional difference [64, 65]. In this work we have utilised the erythroid-specific HMBS isoform, however, to keep consistency with the literature the amino acid numbering is kept consistent with the ubiquitous isoform.

1.5.1   Mutant p.R26H

The conserved Arg26 residue is found in domain 1 of HMBS, where it is located near the putative substrate-binding site within hydrogen-bonding distance to the substrate-binding site. It is therefore suggested that this residue is involved in proton donation during the formation of ES and ES_2 [67, 69, 79], where it forms a salt bridge with the acidic side chain in the second pyrrole ring of the cofactor [80, 81]. A single nucleotide substitution in the codon of Arg26 yields histidine at residue 26 (c.77G>A; p.R26H) in a CpG dinucleotide hotspot, and was first described in 1993 [80]. It has since been reported in several AIP families [5, 24, 82].

Two other mutants have also been reported in the same codon, yielding p.R26C and p.R26L [83-86]. Mutations at Arg26 or at its equivalent position in *E. coli*, Arg11, result in variants with little to no enzymatic activity compared to the wild-type (wt) enzyme [87. Due to its location, Arg26 is therefore indicated as a critical residue in the elongation process and thus vital for enzymatic catalysis {Gill, 2009 #82, 88, 89].

1.5.2 Mutant p.L30P

Leu30 is located on an α-helix at the entrance of the active site cleft. A leucine to proline substitution at residue 30 was first described in 2019 [21] as a disease-causing variant. Three other mutants have been described at the same residue with little to no enzyme activity, p.L30I, p.L30R and p.L30F[24, 90, 91]. A substitution to a proline will potentially act as a α-helix breaker, altering the secondary structure [92]. The Leu30 residue is located at the start of the helix, a variant at that position would therefore likely interrupt the start of the helix.

1.5.3 Mutant p.K98R

The positively charged residue Lys98, located at the active site in domain 1, forms a salt bridge with the acetate side chain of the cofactor [79]. A lysine to arginine substitution was first reported as a disease-causing variant in 1995 and predicted to alter the α-helix content [83]. Lys98 is an important residue in assembling of the cofactor and substrates, thus essential in catalysis and a mutation here will likely affect the enzyme activity [67].

1.5.4 Mutant p.D99N

The conserved Asp99 residue located in domain 1 of HMBS, interacts with the nitrogen atoms of the pyrrole ring, probably in all stages of the pyrrole elongation, and is therefore crucial in the overall reaction [67]. An aspartic acid to asparagine substitution at residue 99 (c.295 G>A; p.D99N) was first described in 2001 [93, 94]. Due to its location and involvement in the pyrrole elongation, Asp99 is suggested as a key residue in enzyme catalysis and likely to cause reduced activity if mutated [93, 95, 96]. Guo *et al.* [96] performed structural studies on mutations in the corresponding residue in *B. megaterium*, Asp82, and found that different mutations cause differences in cofactor assembly where p.D82N is missing the cofactor whereas p.D83E assembles the cofactor. Reduced enzyme activity has also been seen in two other missense mutations reported at the same position, p.D99G [95] and p.D99H [77], as well as its corresponding residue in *E. coli*, Asp84 [69, 87].

The p.D99N mutant is not present in the Norwegian AIP-population, however, based on the previous findings indicating this mutant being trapped in the ES$_2$-state [95] and findings in the corresponding residues Asp84 in *E. coli* [87] and Asp82 in *B. megaterium* [96], suggesting its importance in the catalysis, it was included in this work to investigate the behaviour of this mutant in human HMBS.

1.5.5   Mutant p.R149Q

Arg149 in domain 2 of HMBS interacts with the acidic side chain of the first pyrrole ring of the cofactor [67, 97]. An arginine to glutamate substitution at position 149 (c.446G>A; p.R149Q) in the HMBS enzyme was first described in 1991 [98] in a CpG dinucleotide hotspot, and since been reported in several patients [20, 24, 84]. Two other missense mutants, p.R149L and p.R149P in addition to the nonsense mutation p.R149X, have also been reported at the same position [83, 94, 99, 100]. Mutations at this position are proposed to inhibit the binding of the cofactor thereby leading to an apoenzyme [88, 95, 101].

1.5.6   Mutant p.R167Q

Residue Arg167 is a conserved arginine located in domain 2 at the active site cleft of HMBS where it has been proposed as an important residue for substrate binding and catalytic activity. It has recently, through molecular dynamics simulations, been described as an essential residue for proton donation when PBG molecules are bound [102]. An arginine to glutamate substitution at residue 167 (c.500G>A;p.R167Q) was first reported in 1990 [103] and has been indicated to severely affect the catalytic activity by disturbing the elongation mechanism [23, 79, 95]. Another mutant has been described at the same location, p.R167W, and is also described as causing severely impaired enzyme activity [104-108].

1.5.7   Mutant p.R195C

The conserved Arg195 is located at the active site cleft in domain 2 [67]. Structural studies of HMBS have shown that Arg195 likely interacts with the C2 ring of the cofactor and mutations located here will likely affect the enzymatic activity [79]. A single nucleotide substitution in the codon of Arg195 yielding cysteine at residue 195 (c.583C>T; p.R195C) was first described in 1995 [83] and has since been reported to occur at a CpG dinucleotide hotspot [109]. p.R195C has been described several times in the literature and is reported to markedly reduce the enzymatic activity [20, 23, 24, 110]. Two other mutants have been

described at the same location, p.R195H and p.R195A, and have also been described as having no activity [67, 90].
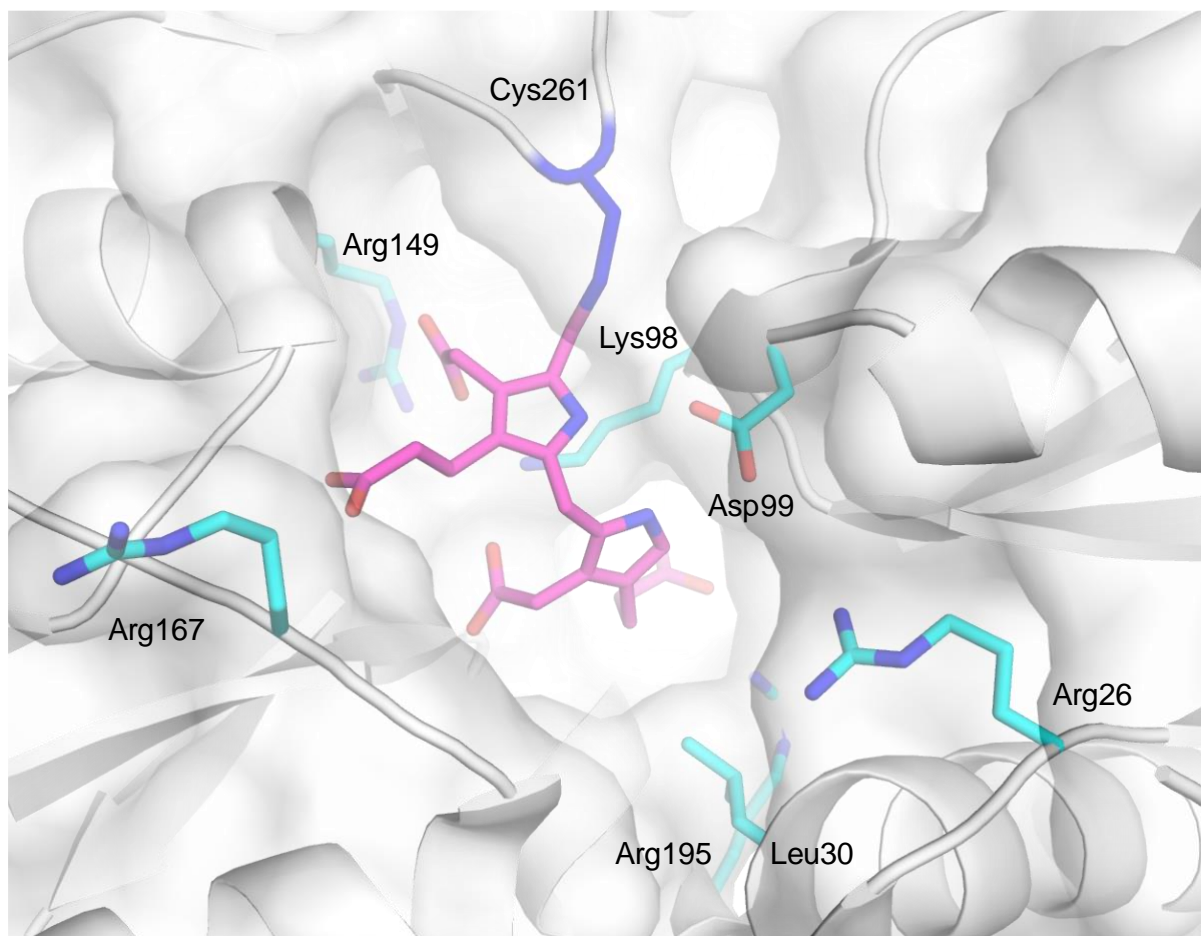
Figure 1-3. The active site cavity of HMBS. Structure of HMBS, zoomed in on the active site cavity. The protein surface is added to better illustrate the cavity with the DPM cofactor (pink) linked to Cys261 and seen with the residues where mutations have been characterised in this work, located in and around the active site, Arg26, Leu 30, Lys98, Asp99, Arg149, Arg167 and Arg195, as sticks (light blue).

1.5.8 Mutant p.E223K

Glu223 is located in domain 1 of HMBS and highly conserved across species. A p.E223K variant was first described in 1994 [99] as a disease-causing variant, but no functional studies has been conducted.

1.5.9 Mutant p.L238P

Residue Leu238 is located in domain 1 and is described as a hinge residue positioned away from the active site [67]. A leucine to proline substitution at residue 238 was first described in 2006 [86] and since been described a few times but no function studies have been

conducted [94, 111, 112]. One other variant has been described at the same position, p.L238R [83].

1.5.10 Mutant p.E250Q

Glu250 is located in domain 3 of HMBS and has no direct interactions with the cofactor or substrates. It does however form a salt-bridge with Arg116 which is an important hinge residue linking domain 1 and 2 [67]. A variant at Glu250 would therefore possibly break an important interaction with Arg116 and could potentially have severe structural effect. A glutamic acid to glutamate substitution at residue 250, p.E250Q, was first reported in 1995 [113]. Functional studies have found the variant p.E250Q to have only 2 % enzyme activity, and two other variants have been described at the same residue, p.G250K [99] and p.G250V [5].

1.5.11 Mutant p.A252P

Ala252 is located close to the active site loop in domain 3. An alanine to proline substitution at residue 252 was first reported in 1997 [114] but no functional studies have been conducted so far. As mentioned for p.L30P, a substitution to a proline is undesirable as it could potentially act as a α-helix breaker and likely disturb the orientation of the α-helix thus is likely to affect the secondary structure of the enzyme.

1.5.12 Mutant p.A339D

Ala339 is located in the $\alpha 2_3$ helix in domain 3 distal from the active site. The alanine to aspartic acid substitution at residue 339 (p.A339D) has not previously been reported in in the literature, however through structural investigations of its interaction network it implicates an important role in stabilisation of the $\alpha 1_3$ helix.
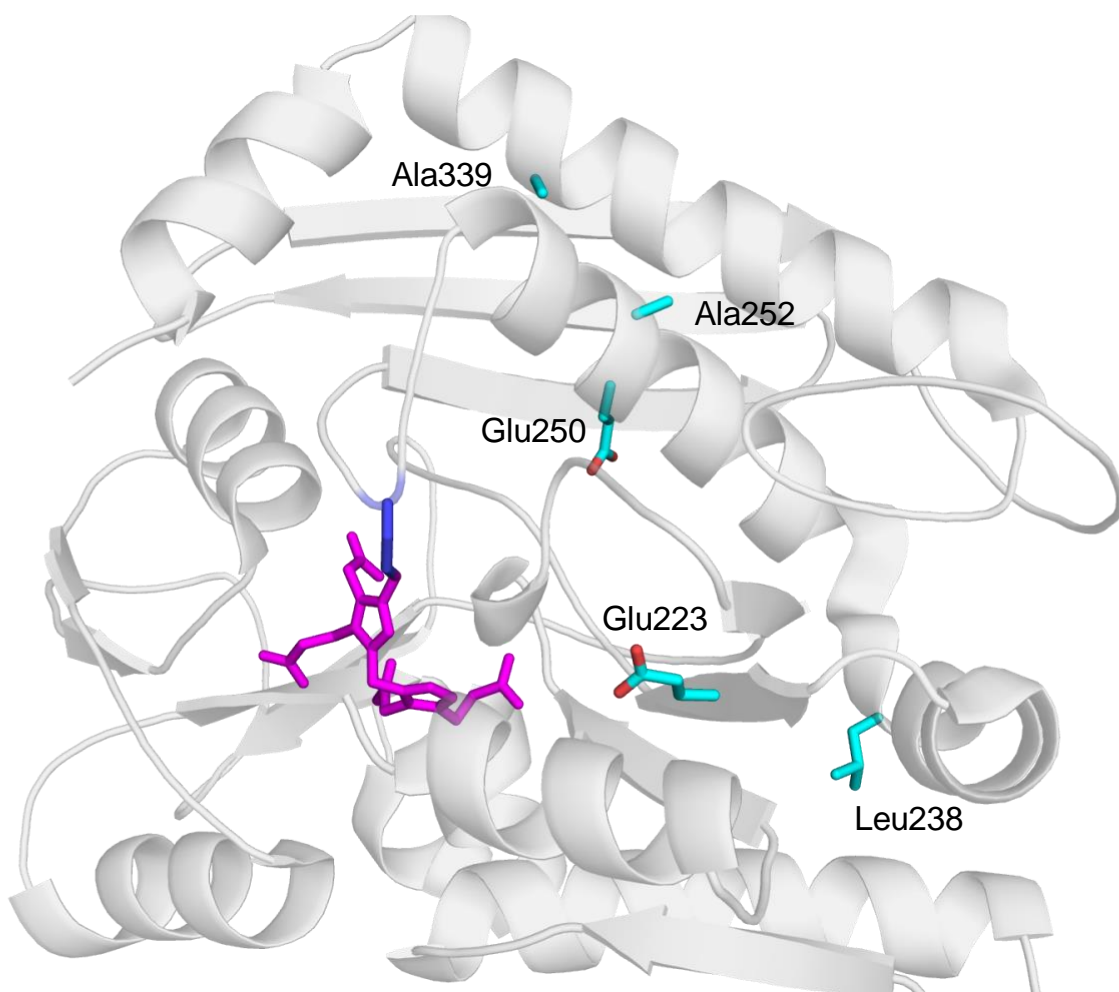
Figure 1-4. The HMBS residues with variants analysed in this project that are located outside of the active site. The structure of HMBS, illustrated with the residues where mutations have been characterised in this work that are located outside the active site cleft, p.E223K, p.L238P, p.E250Q, p.A252P and p.A339D.

2 Aims

Mutations in *HMBS* give rise to acute intermittent porphyria. The overall aim of this research project was to gain understanding of selected disease-associated mutants and their enzyme-intermediate complexes at the protein level. To achieve this, the following objectives were defined:

I. Characterise the effect of twelve selected disease-associated mutations regarding the structure, stability, and activity of HMBS, through following activities:
   a. To recombinantly express and purify the HMBS mutants
   b. To perform biochemical and biophysical characterisation of the mutants
   c. To screen for crystallisation conditions, and X-ray crystallography of selected HMBS mutants in the different intermediate stages if the timeframe allows.
II. Increasing the knowledge of relevant disease-associated mutations on a molecular level, contributing to the fundament of genotype descriptions, thus facilitating for the establishment of future genotype-phenotype relationships in AIP.

These objectives will be achieved using a combination of biochemical and biophysical methods: anion-exchange chromatography, native PAGE, circular dichroism, differential scanning fluorimetry, high-resolution mass spectrometry, X-ray crystallography and enzymatic activity assays, and one of the purposes of this master project has been that the candidate gets proficiency in these techniques

This master's thesis forms a part of the prospective clinical study *Predictors of symptomatic disease and long-term complications in AIP* (PredPor) collaboration with the Norwegian Porphyria Centre (NAPOS) at Haukeland University Hospital.

3 Methodological considerations

3.1 Plasmid and expression

To characterise a protein of interest *in vitro*, it must first be expressed and purified in high quantity and purity. To do this, the gene coding for the protein is first cloned into an expression vector before it is induced into an expression host that can be grown at larger levels, facilitating the production of mg-quantities of the target protein. In this project, the plasmids were purchased from GeneScript as pET vectors cloned with wt-HMBS and selected mutants. An expression vector is a construct containing the gene of interest to be used for protein expression. The pET vector also contains a kanamycin antibiotic resistance gene as a selectable marker, and a lac operon and a T7 promoter used to induce expression by the addition of isopropyl beta-D-1-thiogalactopyranoside (IPTG). HMBS is expressed as a His-tagged fusion protein in *Escherichia coli* (*E. coli*), which is a well-established expression system known to produce high amounts of proteins. The affinity tag is cleavable by Tobacco Etch Virus (TEV) protease at the TEV cleavage site.
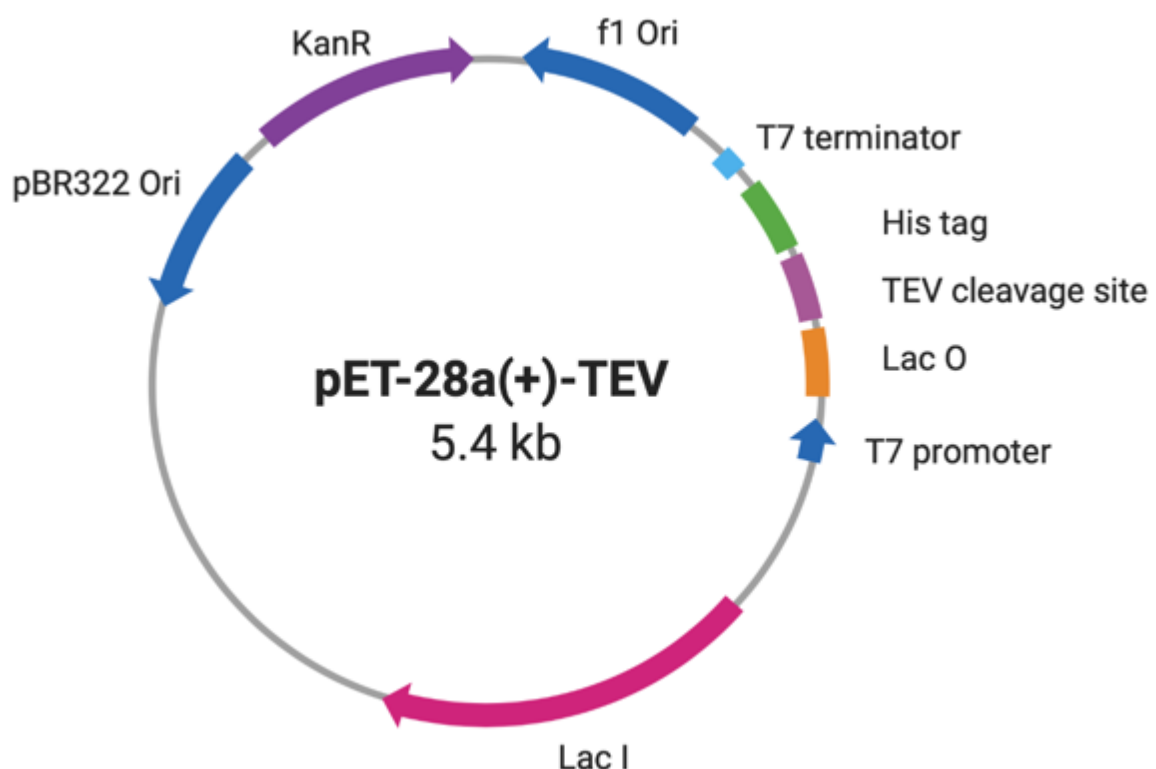

Figure 3-1. Simplified pET vector map. The vector used to express wild-type human HMBS and the selected mutants during this project. Made up of a lac operon, a kanamycin antibiotic resistance gene and a T7 promoter and a His-tag with a TEV cleavage site. The figure is modified from GeneScript and illustrated using BioRender.

3.2 Protein purification and chromatography

Following expression, the protein of interest, here HMBS, is purified by consecutive steps aiming to separate it from all the other proteins, cell debris and unwanted components from the lysate. The purification process depends largely on the characteristics of the target protein, such as stability and solubility. The purification starts with lysis of the cells, often done by e.g., sonication or homogenisation, followed by centrifugation to separate the soluble fraction from insoluble fraction and cell debris. At this stage, the target protein is in the soluble fraction (mobile phase) and is separated based on its interaction with a stationary material such as a coated column or a resin (solid phase).

To isolate the target protein from the soluble fraction after sonication, affinity chromatography is often applied. For His-tagged proteins, the sample can be passed through a gravity flow column containing a nickel resin where the His-tagged protein will have affinity for the nickel resin. When the His-tagged protein is trapped in the resin, the remaining sample matrix can be washed off before the pure His-tagged protein is eluted using imidazole. Imidazole is used as eluant due to its higher affinity for nickel than histidine. Affinity chromatography is used in combination with size-exclusion chromatography (SEC), commonly called gel filtration. SEC is a separation method where proteins in a solution are separated based on their size. The column used in the separation is packed with porous beads allowing the sample mixture to pass through at different speed depending on their size.

3.3 Ion-exchange chromatography

In ion-exchange chromatography (IEX) molecules are separated based on their net charge either by cation or anion exchange. The isoelectric point (pI) is the pH at which the net charge of a protein carries no charge. By increasing the pH in the sample, the protein will become negatively charged and by decreasing it, the protein will carry a positive net charge, and thus pH is a parameter that can be used to modulate the protein interaction with IEX chromatographic resins. Wt-HMBS has a pI of 6.97, while the pI for the selected mutants in this project is in the range of 6.71–7.36. At pH 8, HMBS will therefore carry a net negative charge thus resulting in binding affinity to an anion (positively charged) resin. A Mono Q 4.6/100 PE column (Cytiva) packed with polystyrene and divinyl benzene beads is used with a gradient buffer elution with increasing salt concentration, allowing the enzyme-intermediate complexes in HMBS to be separated based on their net charge once bound to the resin. The

negative salt ions compete with the protein in binding affinity for the stationary phase in the column when the salt concentration increases.
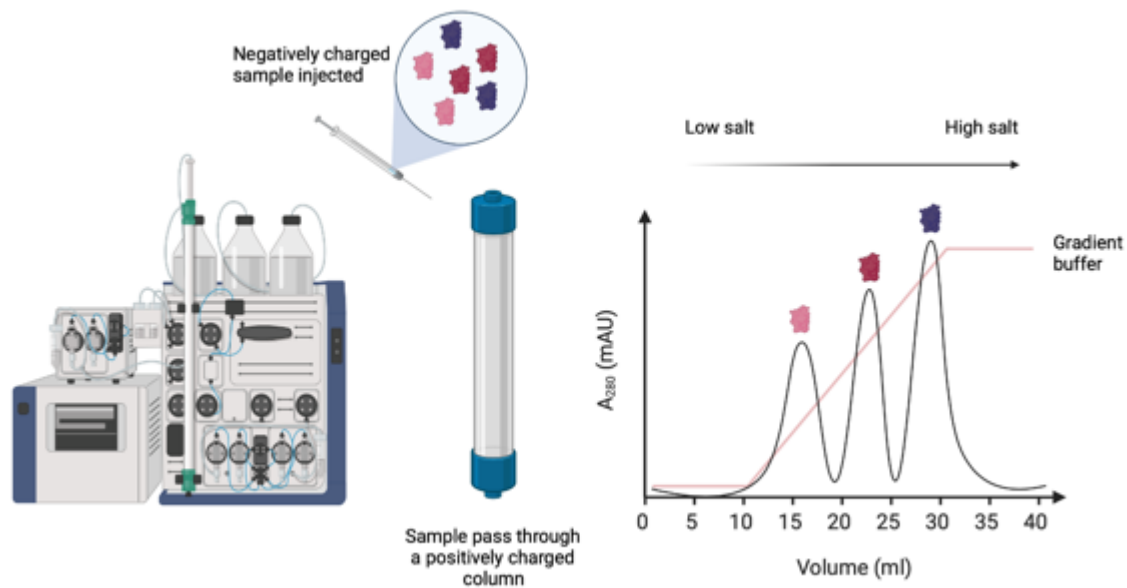

Figure 3-2. Anion-exchange chromatography. The negatively charged proteins pass through a positively charged column. As the salt concentration increases with a gradient elution, proteins in the sample are separated based on their interaction with the stationary phase, due to competition from the salt ions. The proteins elute in the order of increasing charge. Illustrated using BioRender.

3.4 Fourier-transform ion cyclotron resonance

Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is a high-resolution MS method with a high resolving power that allows for the separation and identification of ions with very low mass differences [115]. An electrospray ionisation (ESI) source is commonly used to convert the liquid sample into a gaseous phase by exciting the ions, before they are transferred into a mass analyser, the cyclotron [116]. The cyclotron functions as an ion-trap with a magnet, causing the ions to oscillate based on their mass to charge ratio. Thus, allowing the mass analyser to record a signal used to calculate the specific mass which can identify the ions [115].

3.5 Polyacrylamide gel electrophoresis

Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) is the most common gel electrophoresis technique. Here, the proteins are separated based on their molecular weight independent on charge. To do this, SDS is used to denature the protein. The

binding of SDS gives the protein a negative charge and causes it to unfold. In addition, heating at 95 ºC and the addition of a reducing agent, e.g., dithiothreitol (DTT), further ensure denaturation of the proteins by cleaving the disulphide bonds [117]. An electric current is then applied to the gel with the samples, and the proteins migrate down the matrix in the gel towards the positively charged electrode only influenced by their molecular weight [118]. The migration rate of proteins coated with SDS is the opposite of its molecular weight logarithm, thus the smaller proteins will migrate faster than the larger ones.

Native PAGE is a technique used to analyse proteins by *non-denaturing* gel electrophoresis, allowing the analysis of the protein while it is still in its folded state. Unlike SDS-PAGE, a native-PAGE analysis does not use denaturing or reducing agents, and the migration of the protein therefore depends not only on its molecular weight but also on its folded structure and net charge.

3.6   Circular dichroism

By measuring the difference in absorbance of left and right polarised light, circular dichroism spectroscopy (CD) is a technique that provides the estimation of the secondary structure content and degree of folding of proteins [119]. The proteins are scanned in the far-UV region, commonly at 190–260 nm, upon which the data from the spectrum is used to estimate the content of secondary structure elements in the protein: α-helices, β-sheets, and disordered regions [120]. The different elements of the secondary structure have characteristic spectrums allowing for an interpretation of its folding and the far-UV region absorbs light from the peptide bonds of a protein by signals from two types of electron transitions: $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ [121]. Far-UV CD can also be applied to show conformational changes resulting from a gene variant, or a protein at different conditions such as temperature, pH and buffer, as well as observing the unfolding of the protein as a function of temperature at a set wavelength.

3.7   Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) is a high-throughput method that monitors the thermal denaturation of a protein. As the temperature increases, proteins start to unfold, and the hydrophobic area of a protein becomes exposed. A fluorescent dye such as SYPRO orange, which is added to the sample prior to analysis, will bind to these hydrophobic areas of the protein during unfolding and become unquenched while a fluorescent probe records the

changes in fluorescence [122-124]. A plate reader, such as RT-PCR, can then be used to monitor and record this signal. The data obtained from the analysis presents the fraction of unfolded protein as a function of temperature, and the melting temperature (T_m) is defined as the temperature at which 50 % of the protein sample is unfolded. The transitions in these analyses can often be complex, we therefore utilised the first derivative (dF/dT) to allowing for a better visualisation of the T_m-values (at the inflexion points of the transitions), and differences between transitions.

3.8 Enzyme activity assays

Enzymatic activity assays are important to study the enzymatic catalytic properties. The specific activity of an enzyme can be measured by adding a fixed amount of substrate and letting the reaction run for a set time. The reaction starts when the enzyme interacts with the substrate(s) to form an enzyme-substrate complex before it releases the product and regenerate the free enzyme. When characterising pathogenic variants of an enzyme, the wt enzyme can be used for comparison, and the relative activity between wt and mutant can be calculated. As HMB is spontaneously rearranged and cyclised into URO-I, this can be used as a product to measure the enzymatic activity in vitro. The addition of benzoquinone to the reaction will oxidise URO-I to uroporphyrin, whose absorbance can be measured at 405 nm by a spectrophotometer. Making a standard curve using URO-I will therefore allow us to measure the enzymatic activity of HMBS, and the results are reported as nmol of URO-I per hour per mg of enzyme.

Furthermore, enzyme kinetic assays can be used to study the reaction rate of an enzyme providing insights into the catalytic mechanism. One way of doing this is using the Michaelis-Menten model to analyse the steady-state kinetics of an enzyme. Using a range from low to high concentration of substrate it is possible to calculate the maximum velocity (V_{max}) and the K_m, which is the substrate concentration providing a reaction rate of half V_{max}.

3.9 X-ray crystallography

X-ray crystallography is the oldest method used study the X-ray diffraction patterns of a protein crystal to determine the 3-dimensional structure of protein. The first step in this process is to purify high-quality protein in solution to prepare the protein crystals, starting by screening for crystallisation conditions by varying protein concentration, precipitant, salts,

and pH [125]. Commercial screens such as Molecular Dimensions JCSG plus HT containing a range of conditions such as different salts, organic acids and precipitants varying in pH are available for this. One of the most common methods for screening crystallisation conditions for proteins is vapour diffusion by either using hanging or sitting drops. The basic principle here is to allow the drop, containing protein and precipitant, to release water vapour which ends up in the reservoir solution in a closed system, in order to reach equilibrium between the precipitant in the drop and in the reservoir. This causes supersaturation of the protein in the drop, resulting in formation of crystal nuclei and crystal growth at favourable conditions [125]. Any hits identified from the screening conditions can be further optimised by adjusting the pH, temperature, precipitant or protein concentration to gain good quality crystal which can be preserved and analysed by an X-ray diffraction source to obtain a diffraction pattern used to uncover the electron density and further structural elements of the protein [126].

4 Materials and Methods

If not otherwise stated, chemicals were purchased from Merck (Darmstadt, Germany). Wt-HMBS and mutants were expressed in *E. coli* BL21(DE3) strains as His-tagged fusion proteins with a TEV protease cleavage site (NLYFQ/G) in a pET-28a (+)-TEV vector. HMBS DNA constructs were purchased from GeneScript Biotech (Piscataway, New Jersey, USA). In this project, we have utilised the recombinant wt erythroid *HMBS* gene construct (NM_001024382.2), and selected missense variants listed in Table 1. To keep consistency with the literature we will refer to the constructs in this work as the corresponding constructs in the ubiquitous HMBS isoform (NM_000190.4), namely wt-HMBS and the mutants as p.R26H, p.L30P, p.K98R, p.D99N, p.R149Q, p.R167, p.R195C, p.E223K, p.L238P, p.E250Q, p.A252P and p.A339D.

Table 1. The HMBS variants prepared in this work, with the numeration of the erythroid HMBS construct utilised in this project, and their corresponding numeration in the ubiquitous HMBS isoform.

Residue	Ubiquitous HMBS		Erythroid HMBS	
	Protein	Nucleotide	Protein	Nucleotide
Arg26	p.R26H	c.77G>A	p.R9H	c.26G>A
Leu30	p.L30P	c.89T>C	p.L13P	c.38T>C
Lys98	p.K98R	c.293A>G	p.K81R	c.242A>G
Asp99	p.D99N	c.295G>A	p.D82N	c.244G>A
Arg149	p.R149Q	c.446G>A	p.R132Q	c.395G>A
Arg167	p.R167Q	c.500G>A	p.R150Q	c.449G>A
Arg195	p.R195C	c.583C>T	p.R178C	c.532C>T
Glu223	p.E223K	c.667G>A	p.E206K	c.616G>A
Leu238	p.L238P	c. 713T>C	p.L221P	c.662T>C
Glu250	p.E250Q	c.748G>C	p.E223Q	c.697G>C
Ala252	p.A252P	c.754G>C	p.A235P	c.703G>C
Ala339	p.A339D	c.1016C>A	p.A322D	c.965C>A

4.1   Transformation of protein

In brief, wt-HMBS or mutant DNA was transformed into BL21 (DE3) cells by adding 2 µL plasmid to 50 µL cells and incubated on ice for 30 min before heat shock at 42 °C for 45 sec. After two min on ice, 950 µL of SOC media (New England BioLabs, Massachusetts, USA) was added and the samples were incubated at 37 °C for 45 min. 50 µL of the cells were then spread onto agar plates (1.5 % (w/v) Agar-Agar) and incubated overnight (ON) at 37 °C to grow colonies.

4.2   Plasmid purification

Purification of plasmids was done according to the Monarch Plasmid MiniPrep Kit protocol (New England BioLabs, Massachusetts, USA). In brief, 5 mL LB cultures were incubated overnight before pelleted and resuspended in lysis buffer. Neutralisation buffer was added, and the tube was gently inverted. The sample was incubated for 2 min at room temperature (RT), before it was centrifuged for 5 min, and the supernatant transferred to a miniprep spin column where the sample was washed twice before the column was transferred to a new Eppendorf tube and the plasmid was eluted using Monarch® DNA Elution Buffer. The concentration of the plasmid was determined by measuring the absorbance at 260, using a NanoDrop One/OneC microvolume UV/Vis Spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA).

4.3   Protein expression and purification

E. coli BL21 (DE3) expressing wt-HMBS and mutants were added to 50 mL LB broth (EZmix™) starter cultures supplemented with 50 µg/mL kanamycin, grown overnight in a shaking incubator at 200 RPM and 30 °C. The starter cultures were diluted in 1 L LB broth supplemented with kanamycin and incubated further at 200 RPM and 37 °C until OD_{600} of 0.5–0.6 was obtained before they were induced with 0.5 mM IPTG and incubated ON at 200 RPM and 27 °C. Cells were harvested at $6000 \times g$ for 30 min and pellets were resuspended in 15 mL wash buffer (25 mM Tris pH 8, 400 mM NaCl, 20 mM imidazole at pH 7.5 and 0.5 mM tris (2-carboxyethyl) phosphine (TCEP) with added protease inhibitors (1 tablet cOmplete™ (Roche Diagnostics GmbH, Mannheim, Germany) w/EDTA, 10 mM benzamidine and 0.2 mM PMSF). Resuspended pellets were frozen at –20 °C until use.

The resuspended pellet was thawed on ice and lysis buffer was added to 40 mL before sonication on ice for 3×2 min at 25 W output with 10 sec pulses and pauses on ice, then centrifuged at $20\,000 \times$ g for 45 min. The supernatant was loaded onto a nickel-charged affinity resin (Ni-NTA Agarose, Qiagen) pre-equilibrated in wash buffer, using a gravity flow purification column for His-tag affinity chromatography. The resin was washed $3 \times$ column volume (CV) with wash buffer before eluting the protein with 25 mM Tris pH 8, 400 mM NaCl, 400 mM imidazole at pH 7.5 and 0.5 mM TCEP.

1:100 TEV protease was added to the eluted His-tagged protein to cleave off the fusion tag, during dialysis with 25 mM Tris pH 8, 400 mM NaCl and 0.5 mM TCEP in a cellulose membrane dialysis tubing with MW cut-off at 14 000 ON at 4 °C. Cleaved protein was run through Ni-NTA pre-equilibrated with 25 mM Tris pH 8, 150 mM NaCl and 0.5 mM TCEP, collected and concentrated using Amicon Ultra centrifugal 30 kDa cut-off filters (Millipore) at $4000 \times$ g. A NanoDrop Spectrophotometer (Thermo Fisher Scientific) was used to measure the protein concentration at 280 nm using the sequenced-derived theoretical extinction coefficient (ε) of 15720 $M^{-1}cm^{-1}$ for wt-HMBS and all mutants.

4.4 Size Exclusion Chromatography

A calibration curve was made for the GE HiLoad Superdex 75 16/60 PG column using a gel filtration calibration kit including conalbumin (75 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa), and blue dextran as a void volume marker. Concentrated wt-HMBS and mutants were further purified by SEC on a GE HiLoad Superdex 75 16/60 PG column connected to an Äkta Pure Protein Purification System (Cytiva Europe GmbH, Freiburg, Germany) at 4 °C, with 25 mM Tris pH 8, 30 mM NaCl as running buffer at flow rate 1 mL/min. The fraction containing protein was collected and concentrated using Amicon Ultra centrifugal 30 kDa cut-off filters (Millipore) at $4000 \times$ g.

4.5 Anion-exchange chromatography

Anion-exchange chromatography using a Mono Q 4.6/100 column (Cytiva) connected to an Äkta Pure (Cytiva) system was used to separate the enzyme-intermediate complexes in wt-HMBS and mutants. The Mono Q column was equilibrated in 25 mM Tris pH 8, 30 mM NaCl (buffer A) at 4 °C. The protein was eluted using a gradient from 0 % to 65 % using 25 mM Tris pH 8, 400 mM NaCl (buffer B) and peak fractions were collected. The isolated

intermediates were concentrated using Amicon Ultra centrifugal 30 kDa cut-off filters (Millipore) for further analyses.

4.6 Mass spectrometry

Purified enzyme-intermediate complexes from wt-HMBS and mutants were sent to collaborators Dr. Mikko Laitaoja and Prof. Janne Jänis, at the University of Eastern Finland where the samples were analysed as previously reported in [72]. Briefly, the enzymes were buffer-exchanged to 20 mM NH_4OAc pH 6.8 using PD Miditrap G-25 column (Cytiva). Further dilutions were made using HPLC quality acetonitrile, water, and acetic acid. Mass spectra were measured using a Bruker 12-T solariX XR FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). All samples were measured by direct infusion, at 2 µL/min, using standard electrospray source. The instrument was calibrated using NaPFHA clusters before measurements. The mass spectrometer was operated using ftmsControl 2.2 software, and the data were processed using the Bruker DataAnalysis 5.1 software. Deconvolution (i.e., zero-charge) spectra were calculated using a Maximum Entropy (MaxEnt) deconvolution algorithm. All masses are reported as most abundant isotopic masses and compared with their theoretical masses sequence-derived and predicted by the ProtParam Expasy software. The theoretical molecular weight (Da) for the mutants we were able to characterise are listed in Table 2.

Table 2. Theoretical molecular weight (Da) of wt-HMBS and analysed missense mutants.

Protein	Apo	E_{holo}	ES	ES_2	ES_3	ES_4	pI
Wild-type	37756	38175	38384	38593	38802	39012	6.97
p.R26H	37737	38156	38365	38574	38783	38993	6.75
p.L30P	37683	38101	38311	38520	38729	38938	6.90
p.D99N	37755	38174	38383	38592	38801	39011	7.36
p.R149Q	37728	38147	38356	38565	38774	38984	6.72
p.R167Q	37728	38147	38356	38565	38774	38984	6.72
p.R195C	37703	38122	38331	38540	38749	38959	6.72
p.E223K	37698	38116	38326	38535	38744	38953	7.78

4.7   Polyacrylamide Gel Electrophoresis

Wt-HMBS and mutants were analysed by SDS-PAGE with a 2× Laemmli sample buffer (Bio-Rad Laboratories, USA) with final concentration of 200 µM DTT. Samples were heated at 95 °C for 3 min and loaded on a 10 % Mini-PROTEAN® TGX^{TM} Precast Protein Gel (Bio-Rad). The gel ran for approximately 30 min at 200 V with a Precision Plus™ Dual Colour protein standard (Bio-Rad) using SDS-PAGE running buffer (25 mM Tris, 200 mM glycine, 0.1% SDS) at RT before staining with Coomassie blue staining and visualised by Molecular Imager ChemiDoc XRS+ with ImageLab software (Bio-Rad). Native PAGE was performed likewise, however, without the denaturing agents SDS, DTT and heating, and ran at 140 V for 2 ½ h at 4 °C with native PAGE sample buffer (Bio-Rad) and native PAGE running buffer (50 mM Tris, 1920 mM glycine at pH 8.3).

4.8   Circular Dichroism

Far-UV spectra were obtained using a J-810 Jasco spectropolarimeter with a CDF-426S Peltier element for temperature control and a 300 µL quartz cuvette with a path length of 1 mm. Wt-HMBS and mutant samples at 5 µM were prepared in 10 mM K_2HPO_4 and 100 mM NaF. Three scans were obtained for each spectrum and buffer scans were subtracted using the integrated Jasco software. BeStSel web server was used for secondary structure prediction [120], and statistics analysed using unpaired t-test (GraphPad Prism version 9.2.0 for MacOS, GraphPad Software, San Diego, California USA, www.graphpad.com).

4.9   Differential Scanning Fluorimetry

A Light Cycler 480 Real-Time PCR (Roche) was used to obtain thermal denaturing scans for wt-HMBS and mutants. 5 µM protein was prepared in PBS with $5 \times$ SYPRO Orange (Agilent technologies, USA) and analysed in 384-well plates as indicated [127]. The thermal denaturing curves were recorded at a scan rate of 2 °C/min with 0.2 °C intervals from 20 °C to 99 °C. The half-denaturing temperature (T_m) values were obtained from the maximum of the 1st derivative in the raw data. Data were analysed using HTSDSF Explorer [123]. The T_m-values provided for wt-HMBS result from average of four measurements from four different plates, whereas the values for the mutants are obtained from average of four in-plate measurements.

4.10 Enzymatic Activity Assays

Prior to the activity assay, a standard curve was made using URO-I where URO-I was dissolved in 6 M HCl to make a stock solution at 1:6 dilution in ddH$_2$O. A dilution series was prepared from 100 to 1.56 % stock in 1 M HCl, and a NanoDrop Spectrophotometer was used to measure URO-I concentration at A$_{405}$ (ε_{URO-1}= 0.547 mmol^{-1}cm^{-1}) calculated as concentration:

$$c = \frac{A_{405}}{l \times \varepsilon_{URO-1}}$$ (Eq.1)

The measurements were analysed by a linear regression model to make a standard curve, from which the regression equation was extracted. This equation was used to calculate the specific activity of the enzyme and reported as nmol of URO-I/hr per mg enzyme.

As previously reported in [108], the enzyme activity of recombinant wt-HMBS and mutants was assayed by adding 5 µg enzyme to 50 mM Na-Hepes pH 8.2, and the sample was pre-incubated for 3 min at 37 °C To start the enzyme reaction 100 µM PBG was added, and the reaction was stopped by adding 5 M HCl and 0.1 % benzoquinone after 4 min. Samples were incubated on ice for 30 min before the absorbance was measured at 405 nm by a NanoDrop Spectrophotometer, defined as the formation of 1 nmol URO-1 per hour. Michaelis-Menten kinetics was used to determine the kinetic parameters; K_M and V_{max} were obtained using a range of PBG concentrations from 3.125 to 1000 µM and analysed using the Michaelis-Menten enzyme kinetics model in GraphPad Prism.

4.11 X-ray Crystallography

Screening of crystallisation conditions for HMBS mutants p.R26H and p.R195C was carried out by vapour diffusion using 96-well MRC SD2 and TTP LabTech sitting drop plates using a Mosquito LCP crystallisation robot (SPTLabTech, Melbourn, UK). Based on results acquired during the project, mutant p.R26H was screened as the isolated ES$_2$ intermediate, since results show that it is trapped at this stage in the elongation process. Mutant p.R195C was screened as the purified enzyme sample incubated with PBG, as results show that this mutant accumulates the ES$_4$ intermediate in the presence of PBG substrate, and the structure of the ES$_4$ intermediate has not been described before.

HMBS mutants p.R26H and p.R195C was concentrated using Amicon Ultra centrifugal filters with 30 kDa cut-off (Millipore) to 10 mg/mL and 8 mg/mL concentration, respectively, and filtered using Costar Spin-X centrifuge filters prior to screening. The initial screening

was conducted using the commercial screens PACT premier HT and JSCG plus HT (Molecular Dimensions, Sheffield, UK). The plates were stored either in a 20 ºC incubator and examined manually under a microscope or in the 20 ºC Rock Imager 182 (Formulatrix, Massachusetts, USA), where they were automatically imaged at a set time interval.

TCEP was added to p.R26H prior to screening to prevent oxidation of the ES$_2$ substrate complex. p.R26H was optimised using a small gradient from the initial hit from JCSG plus C5 (0.8 M NaH2PO4, 0.8 M KH2PO4, 0.1 M HEPES pH 7.5), by altering the salt concentration, plus screening around conditions where crystals have been obtained previously for wt-HMBS [72, 128]. Crystals were observed for p.R26H after 2 days. From the first round of optimisation for p.R26H, six single crystals were picked for diffraction testing, from conditions containing 20 % PEG 3350 and 0.1 M Ammonium Citrate pH 5.5. The crystals were picked using litho loops (Molecular Dimensions) and briefly soaked in crystallisation solution implemented with 30 % glycerol for cryoprotection prior to flash freezing with liquid nitrogen. Crystals were tested for X-ray diffraction using remote access on the EMBL beamline P13 at PETRAIII at DESY, Hamburg, Germany. Crystal picking and diffraction studies were done by Dr. Juha Kallio.

4.12 Structural investigations

Figures prepared from crystal structures were done using PDB ID: 7AAJ and Pymol (Schrödinger, L. & DeLano, W., 2022. *PyMOL 2.5.2 for Mac*, Available at: http://www.pymol.org/pymol). Structural Interactions were investigated using Discovery Studio (BIOVIA, Dassault Systèmes, BIOVIA Discovery Studio, San Diego: Dassault Systèmes, 2022.

4.13 Statistics and reproducibility

For DSF experiments the measurement are a result of four in-plate replicates for all mutants and four between-plates replicates for wt-HMBS and analysed as mean ± standard deviation (SD) by Excel (Microsoft Office 365, Excel for Mac).

CD experiments were analysed in triplicates and buffer scans were subtracted using the integrated Jasco software. BeStSel web server was used for secondary structure prediction, and statistics analysed using unpaired t-test (GraphPad Prism version 9.2.0 for MacOS, GraphPad Software, San Diego, California USA (http://www.graphpad.com).

Activity assays were performed with wt-HMBS as an assay control for each experiment, all measurements were analysed in triplicates and the Michaelis-Menten enzyme kinetics model and K_M and $V_{max} \pm$ the standard error of mean (SEM) were calculated using GraphPad Prism.

5 Results

In this project, we have utilised the shorter human recombinant erythroid wt-HMBS and the selected missense variants listed in Table 1 (Chapter 4). To keep consistency with the literature we will refer to the construct and the numeration of the mutants as the corresponding ubiquitous HMBS isoform, namely wt-HMBS and mutants as p.R26H, p.L30P, p.K98R, p.D99N, p.R149Q, p.R167Q, p.R195C, p.E223K, p.L238P, p.E250Q, p.A252P and p.A339D (see details in chapter 1.5).

5.1 Summary of results

All 12 selected HMBS mutants were expressed and attempted to be purified following a standard protocol using affinity- and SEC developed for wt-HMBS. The mutants yielding sufficient amounts of protein were further characterised using a combination of IEX, native PAGE, CD, DSF, high-resolution FT-ICR MS and enzyme activity assays.

The first step in characterising a mutant was to analyse the folding using CD to see if it folded correctly, and the results were compared to wt-HMBS. IEX was then used to separate the enzyme intermediates followed by native PAGE to analyse the isolated peaks from the IEX. To confirm which enzyme-intermediate states the mutant corresponds to, the isolated main peak from SEC (denoted as mix hereafter) was further analysed by FT-ICR MS.

As previously demonstrated, the thermostability of HMBS decreases with each addition of a PBG substrate [108], therefore, to investigate if the mutants have any effect on the structural stability of the enzyme, we analysed the isolated enzyme-intermediate complexes obtained from IEX by DSF. Finally, analysis of the enzymatic activity and determination of kinetic parameters was done as previously reported [108].

5.2 Expression and purification

Competent BL21(DE3) cells were transformed with plasmids. Colonies were successfully grown on LB agar plates containing kanamycin before they were expressed and purified following the standard protocol using affinity- and size-exclusion chromatography developed for wt-HMBS.

5.2.1 Expression test

A small-scale expression test was performed to assess the length of induction where samples were taken after 3 and 15 h, and subsequently analysed by SDS-PAGE (Figure 5-1). To be able to directly compare the bands on the gel, the cell density for the samples were diluted and normalised to an optical density of 4. All mutants were successfully expressed. As seen in Figure 5-1 on the expression test, the majority of the mutants show better expression after 15 h, and an overnight expression was therefore chosen for optimal yield. Mutants p.E223K, p.L238P, p.E250Q and p.I113N show a weaker intensity compared to the rest of the mutants.
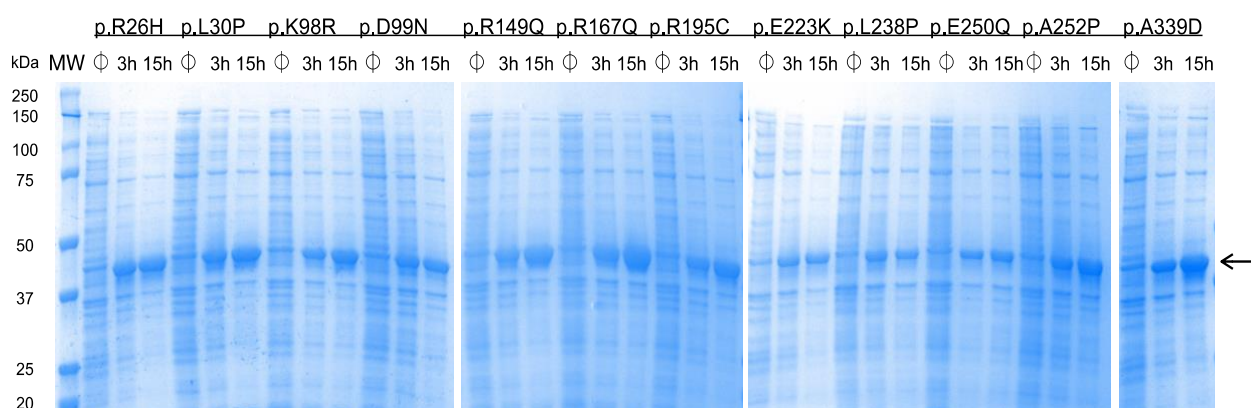

Figure 5-1. Protein expression test of the selected HMBS variants. SDS-PAGE showing the uninduced (Φ) and induced cultures after 3 and 15 h, for mutants p.R26H, p.L30P, p.K98R, p.D99N, p.R149Q, p.R167Q, p.R195C, p.E223K, p.L238P, p.E250Q, p.A252P and p.A339D. A Precision Plus™ Dual Colour protein standard was included in lane 1 of the first gel (MW) and an arrow indicating where the protein is.

5.2.2 Purification

Following expression, wt-HMBS and the mutants were purified by affinity chromatography and SEC. Firstly, wt-HMBS was purified in high yield by nickel-affinity chromatography on a gravity flow column and resulted in ~50 mg fusion protein from 1 L culture. The purification of mutants p.R26H and p.R167Q yielded amounts similar to wt-HMBS, p.R149Q yielded ~80 % of the wt-HMBS whereas p.L30P, p.K98R, p.D99N, p. R195C and p.E223K yielded ~50 % compared to wt-HMBS. Mutants p.L238P, p.E250Q, p.A252P and p.A339D yielded little to no protein, and it was therefore decided to not continue any further with the biochemical and biophysical characterisation for these. The His-tagged fusion proteins were cleaved overnight at 4 °C by TEV protease and the samples were further purified by SEC. The purified proteins presented strong bands at ~38 kDa on SDS-PAGE corresponding to their molecular weight (results not shown). Secondly, cleaved wt-HMBS and mutants were further purified by SEC

and the isolated peaks were collected. The elution profiles from SEC are presented in Figure 5-2A-H. As shown, the elution profiles for wt- HMBS and mutants; p.R26H, p.R149Q, p.R167Q, p.R195C and p.E223K, all present three peaks, 1, 2 and 3 respectively. However, the middle peak is absent in mutant p.D99N and the overall elution profile for p.L30P is slightly different from the rest with only two peaks eluting a little earlier compared to the rest of the mutants.
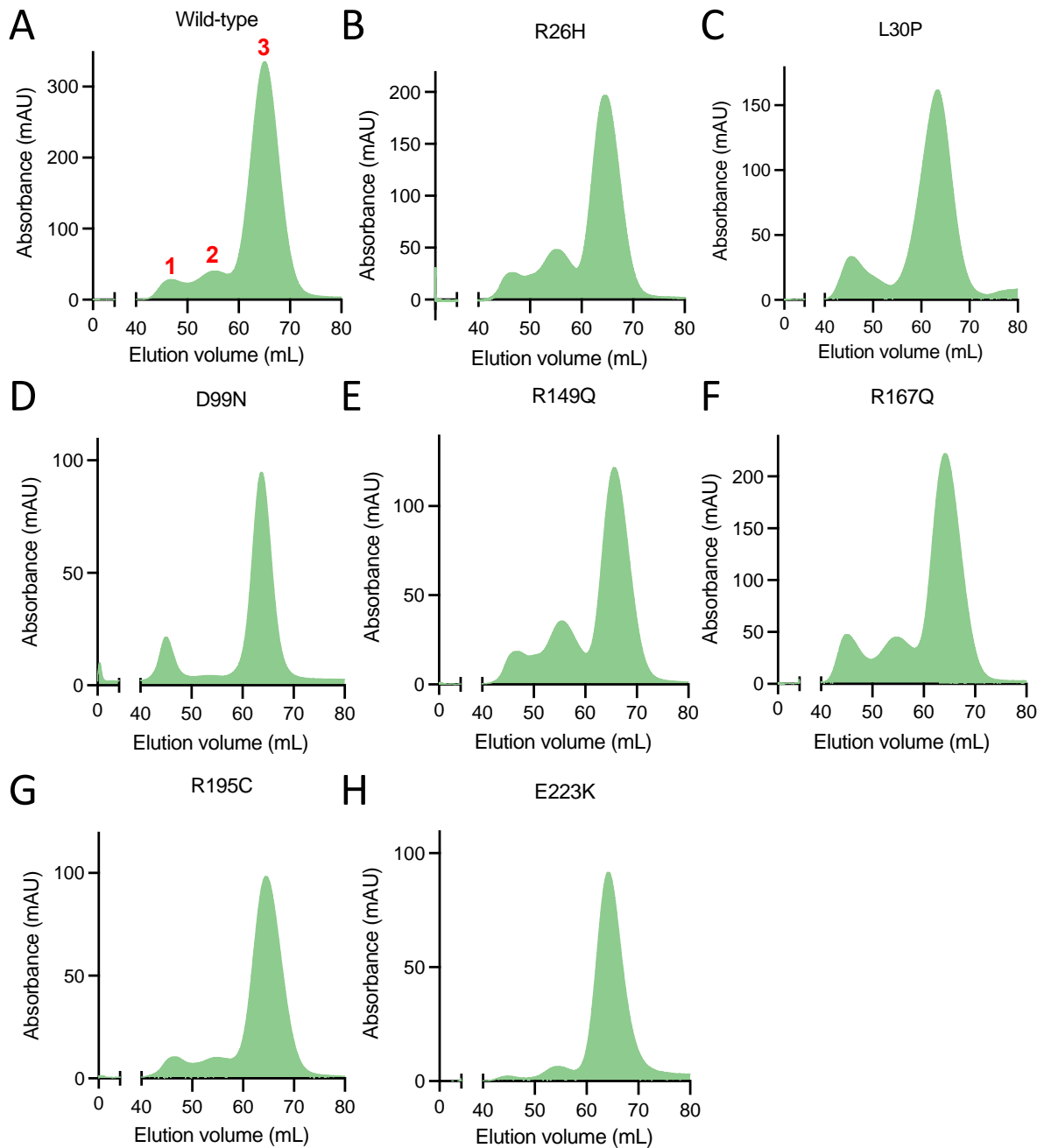

Figure 5-2. Final purification step of wt-HMBS and mutants by size exclusion chromatography. The purification of wt-HMBS and mutants by SEC on a GE HiLoad Superdex 75 16/60 PG column. **(A)** wt-HMBS **(B)** p.R26H **(C)** p.L30P **(D)** p.D99N **(E)** p.R149Q **(F)** p.R167Q, **(G)** p.R195C and **(H)** p.E223K.

5.2.3 Molecular weight estimation

To estimate molecular weight from the size-exclusion chromatograms, a calibration curve was made using a gel filtration calibration kit including conalbumin (75 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa), and blue dextran as a void volume marker. The estimation of molecular weight of an analyte in size exclusion chromatography is determined as a function of the ratio between the elution volume and the total volume of the column used (K_{av}), and calculated using the equation:

$$K_{av} = \frac{(V_e - V_0)}{(V_t - V_0)} \qquad \text{(Eq. 2)}$$

where V_e is the elution volume, V_0 is the void volume, and V_t is the total bed volume (120 mL). V_0 is obtained from blue dextran, which yielded 45.88 mL.
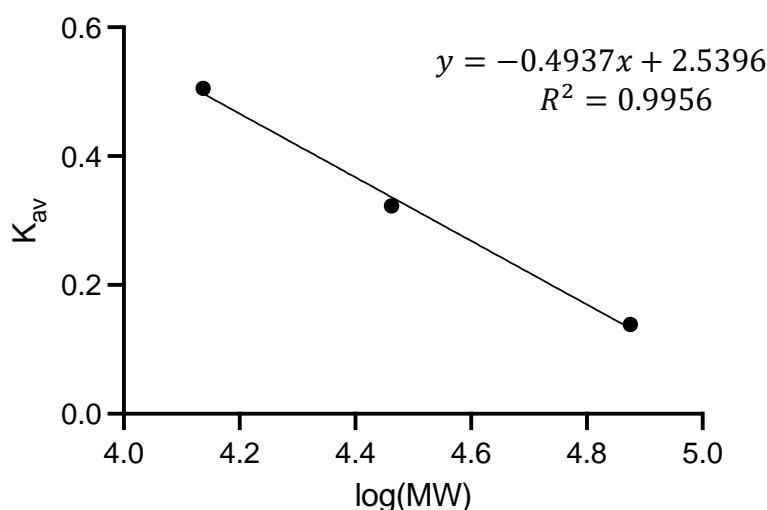

Figure 5-3. GE HiLoad Superdex 75 16/60 PG column calibration curve. A calibration curve made using a gel filtration calibration kit containing conalbumin (75 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa). The data points were plotted on a graph and fitted to a linear regression model that provided the calibration curve equation, y = –0.4937x + 2.5396 with a R^2 = 0.9956.

Eq. 2 was subsequently used to calculate the K_{av} for each component in the gel filtration calibration kit and fitted to a linear regression model, yielding the calibration curve equation:

$$y = -0.4937x + 2.5396 \qquad \text{(Eq. 3)}$$

This equation (Eq. 3) was then used to estimate the size of each peak from the SEC chromatograms and the results are presented in Table 3. The peak eluting first at ~45 mL

(peak 1) contains the population with the largest size and elutes in the void, suggesting the presence of large oligomers, or aggregates. The following peak (peak 2) corresponds to approximately the size of a HMBS dimer, but it has also been associated with a partially unfolded monomer [96]. The third peak (peak 3) corresponds to the approximate size of a HMBS monomer. The species eluting in peak 3 for wt-HMBS and all mutants ranges from ~37 to 43 kDa, which is close to the theoretical molecular weight of HMBS as a monomer (~38 kDa), indicating that the protein elutes mainly as a monomer.

Table 3. Estimated molecular weight from size-exclusion chromatography peaks calculated based on the elution volume of the peak using Eq. 2.

Protein	Peak 1		Peak 2		Peak 3	
	Size (kDa)	Area (%)	Size (kDa)	Area (%)	Size (kDa)	Area (%)
wt	Void	5	71.5	9	39.0	86
p.R26H	Void	6	70.8	16	40.3	78
p.L30P	Void	17	ND	ND	44.9	83
p.K98R	Void	9	70.1	11	41.1	80
p.D99N	Void	14	ND	ND	42.9	86
p.R149Q	Void	6	69.8	20	37.5	74
p.R167Q	Void	10	72.7	12	41.2	78
p.R195C	Void	6	71.8	6	40.2	88
p.E223K	Void	1	79.9	5	42.5	94

*ND = Not detected

5.3   Enzyme-intermediate complexes of wt-HMBS

Following SEC, the pure protein was further analysed by IEX to separate the enzyme-intermediate complexes. Due to a small difference in the net charge of the enzyme - intermediates, arising from the addition of 1, 2 or 3 covalently bound PBG substrates to E_{holo}, it is possible to separate them by IEX using a buffer gradient going from low (buffer A) to high (buffer B) salt concentration. The small difference in net charge also allows the enzyme-intermediate complexes to be separated by native PAGE. As can be seen in the IEX chromatogram for wt-HMBS (Figure 5-4A), the enzyme intermediates are separated from A– E; apoenzyme, E_{holo}, ES, ES_2 and ES_3 respectively, as the percentage of buffer B increases. Similarly, the native PAGE analysis of the isolated IEX peaks also shows bands for the enzyme intermediates (A–E) where they migrate stepwise as another substrate is bound (Lane 2–6, Figure 5-4B). Lane 1 includes the purified wt-HMBS from the mix sample. As can be seen in the gel, smears are observed in several of the lanes suggesting the presence of more than one single intermediate in the isolated peak samples.
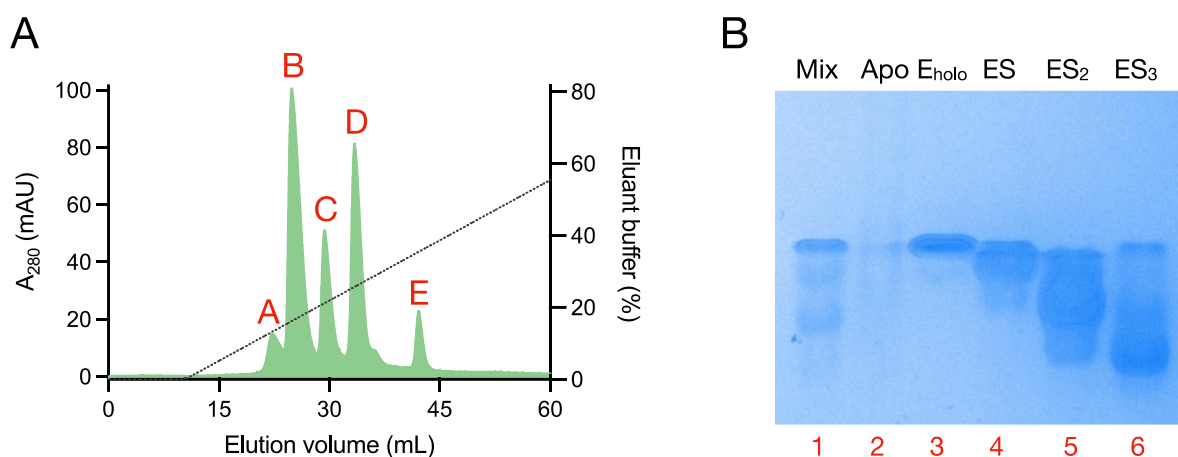
Figure 5-4. Separation of enzyme-intermediate complexes in wt-HMBS. (A) Recombinant wt-HMBS separated by anion-exchange chromatography, presenting peaks for A, apoenzyme; B, E_{holo}; C, ES; D, ES_2 and E, ES_3. **(B)** Native PAGE analysis of wt-HMBS and its isolated anion-exchange peaks. Lane 1, Mix; Lanes 2–6, Isolated peaks A–E from IEX chromatography (Panel A).

FT-ICR-MS data show that the isolated peaks from IEX chromatography do not include purified species, but still exhibit a distribution of enzyme intermediates (Figure 5-5). This demonstrates that even when isolated, the sample still contains an overlap of intermediates, and the analysis of intermediate distribution is therefore not clear cut as also seen in the native PAGE (Figure 5-4B). These results are included to show the complexity of the enzyme intermediate distribution in HMBS and based on this knowledge, we thereby chose to perform the CD and MS analyses of all the mutants from the mix sample and not as isolated intermediates.
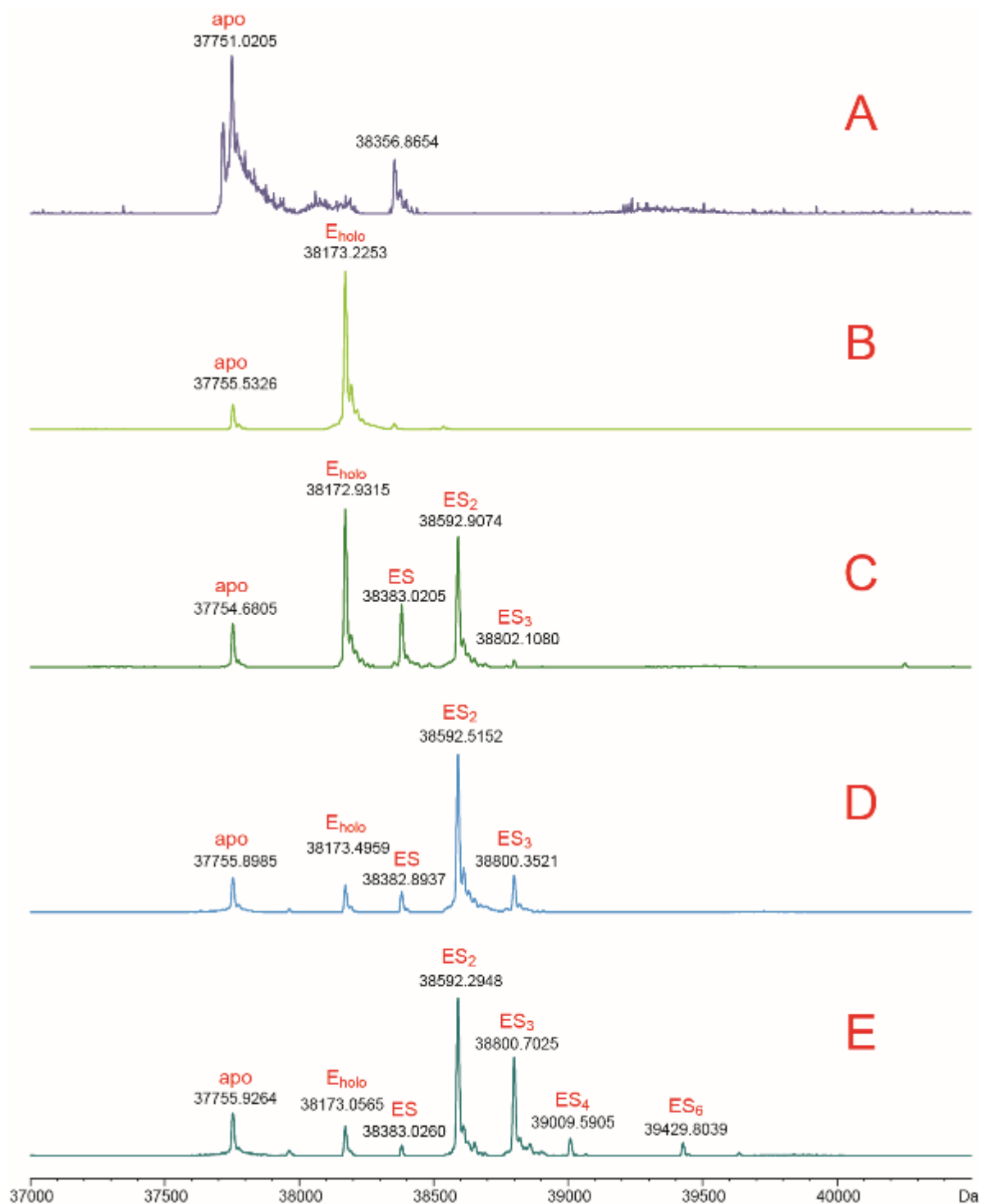
Figure 5-5. FT-ICR-MS of the separated enzyme-intermediate complexes in wt-HMBS. Charged-deconvoluted ESI FT-ICR mass spectrum of the separated enzyme-intermediate complexes in wt-HMBS at denaturing conditions with 5 µM protein. The chromatograms present the isolated enzyme intermediates starting at the top with A, apoenzyme; B, E_{holo}; C, ES; D, ES_2 and E, ES_3.

5.4 Preliminary results for activity measurements

To be able to calculate the specific enzymatic activity of the mutants a standard curve was made using URO-I. As HMB spontaneously rearranges into the non-physiological URO-I, in the absence of the UROS, this can be utilised to measure the enzymatic activity of HMBS. A dilution series of 1.56–100 % URO-I was made, measured by a spectrophotometer, and fitted to a linear regression model. From this, the regression equation was extracted and further used when calculating the specific enzyme activity and the steady-state kinetics for wt-HMBS and mutants.
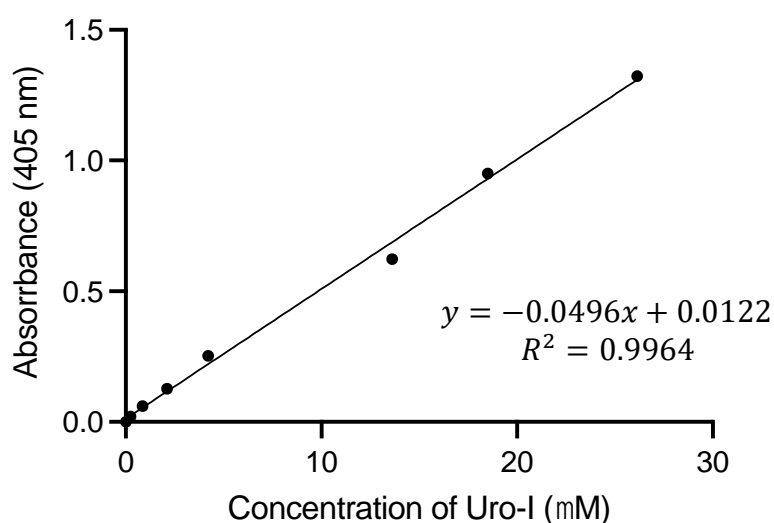

Figure 5-6. Standard curve made using URO-I. In absence of UROS, HMB is spontaneously cyclised into URO-I, which can be utilised to measure the enzyme activity of HMBS. A dilution series was prepared, i.e.,1.56–100 % URO-I, and fitted to a linear regression model to make the standard curve. The regression equation was extracted and used to calculate the specific activity of wt-HMBS and all mutants.

5.5 Biochemical and biophysical characterisation of HMBS mutants

The schematic overview of the workflow for the 12 selected HMBS mutants for which biochemical and physical characterisation was attempted in this project is presented in Figure 5-7.
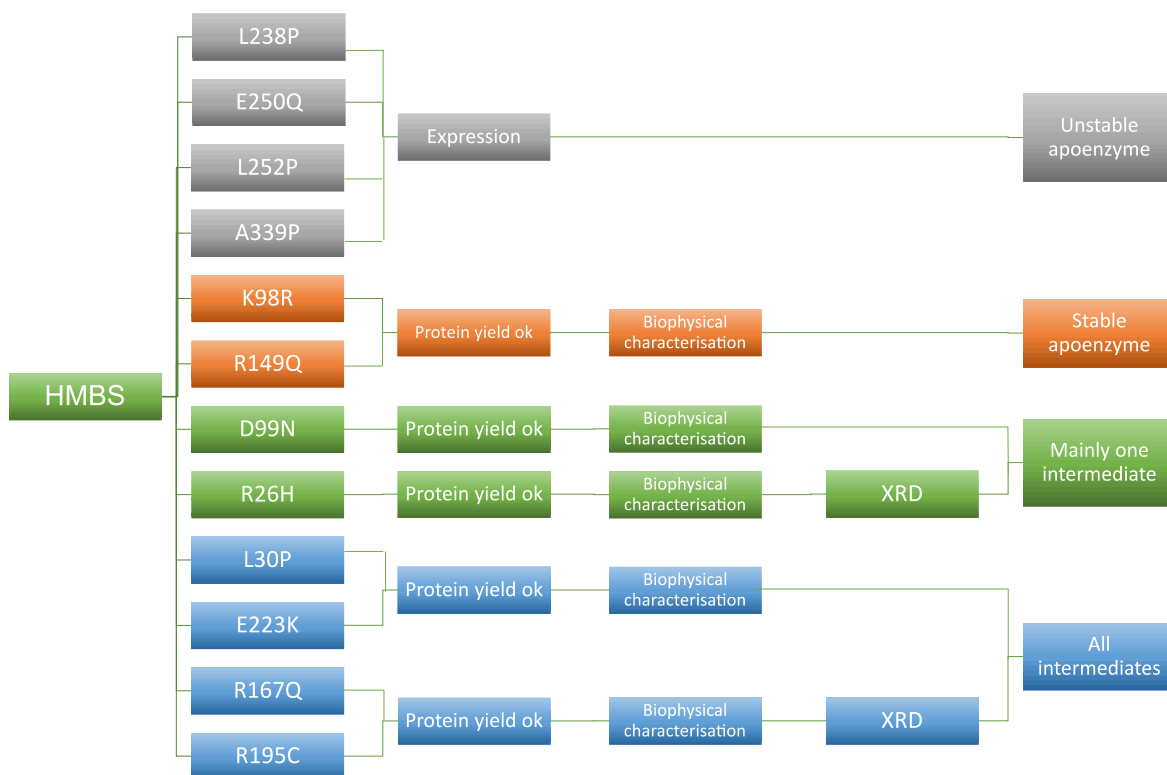

Figure 5-7. Summary of all the HMBS variants analysed in this project. According to the results along the workflow, the mutants are divided into four main classes: complete folding defect (grey), stable apoenzyme (orange), mutants yielding a distribution of intermediates (green) and mutants yielding mainly one intermediate (blue).

5.5.1 Mutants unable to be purified using standard protocol

Like all mutants, p.L238P, p.E250Q, p.A252P and p.A339D were expressed and analysed using the standard expression and purification protocol (see chapter 4.3 in materials and methods). The expression test (Figure 5-1) indicated that they express at slightly lower intensity compared to the previously described mutants. The purification resulted in very little to no amounts protein, and not enough to perform any further experiments.

5.5.2 Mutants unable to assemble the cofactor

5.5.2.1 Mutant p.R149Q

Wt-HMBS presents the characteristic CD spectrum with a double dip at 208 and 222 nm, indicating a large presence of α-helices, and the p.R149Q mutant indicates that its folding is very similar to wt-HMBS (Figure 5-8A). The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). The IEX chromatogram shows the elution profile for mutant p.R149Q and it is apparent that this mutant mainly elutes as a single intermediate corresponding to the apoenzyme in wt-HMBS at ∼24 % buffer B (Figure 5-8B). The native PAGE analysis (Figure 5-8C) shows wt-HMBS (lane 1) presenting all bands while p.R149Q both mix (lane 2) and isolated IEX peak (lane 3) migrates as a smear at the top with one band further down above the wt-HMBS E_{holo} band. The mix from the mutant was further analysed by FT-ICR-MS and the analysis confirmed that p.R149Q present a single intermediate in the apo-state (Figure 5-8D). The DSF analysis of p.R149Q shows a very low thermostability with a T_m-value of 42.30 ± 0.07 ℃ which is ∼32 ℃ lower compared to wt-HMBS intermediates ranging from 78.19 ± 0.70 ℃ to 71.01 ± 0.18 ℃ (Figure 5-8E). Wt-HMBS had a specific activity of 757.70±67.67 nmol×mg-1×h-1 whereas the mutant p.R149Q was completely inactive (Figure 5-8F).
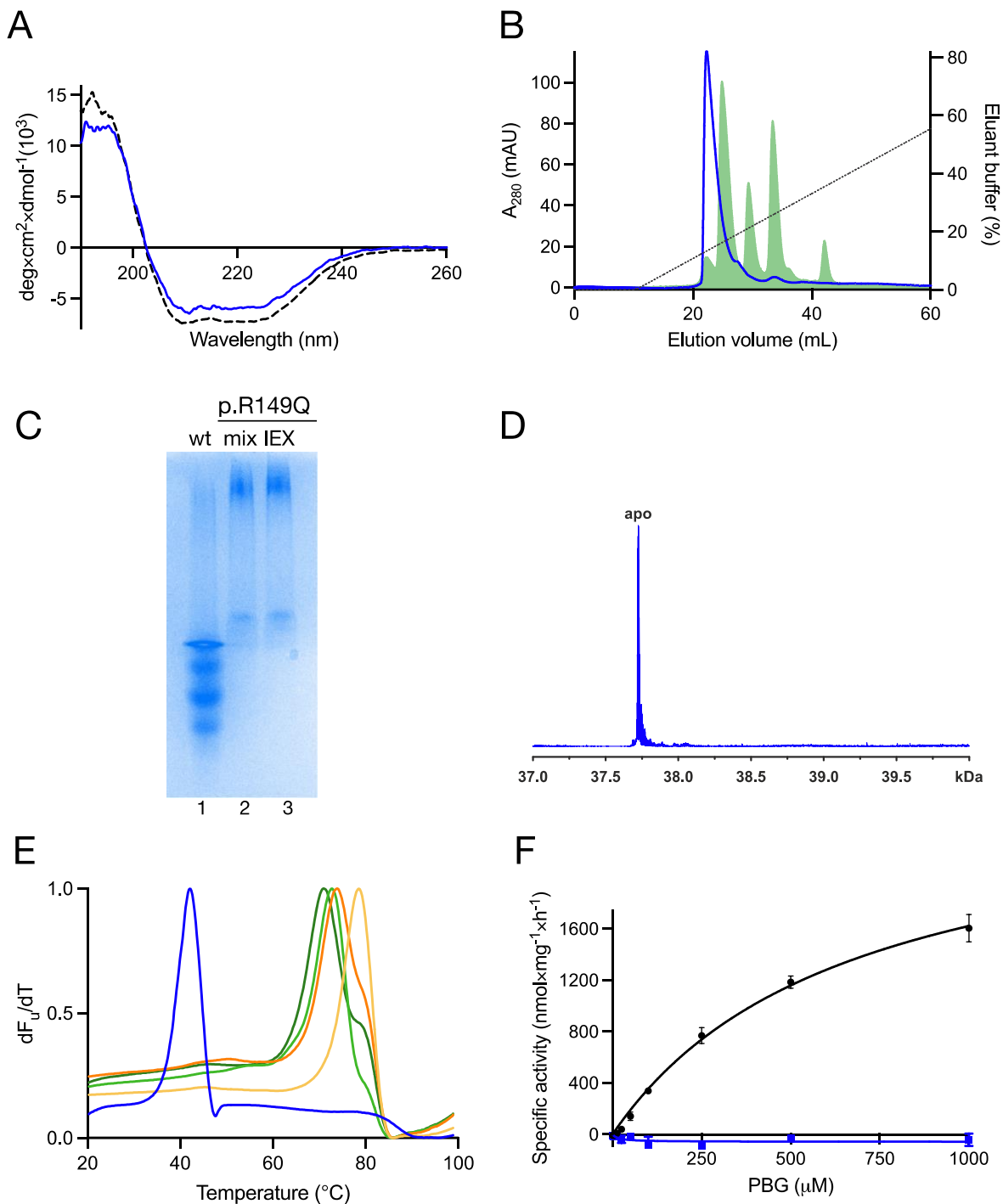
Figure 5-8. Biophysical characterisation of HMBS mutant p.R149Q. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.R149Q (blue) measured in triplicates; unpaired t-test shows no significant difference between the two spectra. **(B)** Isolation of enzyme-intermediate complexes by IEX. Recombinantly expressed wt-HMBS (green) with mutant p.R149Q (blue). **(C)** Native PAGE analysis of wt-HMBS (lane 1, wt), p.R149 (lane 2, SEC) and the isolated main peak of p.R149Q (lane 3, IEX). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.R149Q in denaturing conditions at 1 µM protein concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of isolated p.R149Q IEX peaks (blue) with wt-HMBS enzyme intermediates for reference; E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.E223K (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.2.2   Mutant p.K98R

Due to the low amounts of protein available following purification, FT-ICR MS and activity assays were prioritised in the analysis of p.K98R as this would provide the most informational data. The FT-ICR MS spectrum show that the mutant presents only the apoenzyme (Figure 5-9A) and the enzyme activity assays shows that the mutant is inactive (Figure 5-9B).
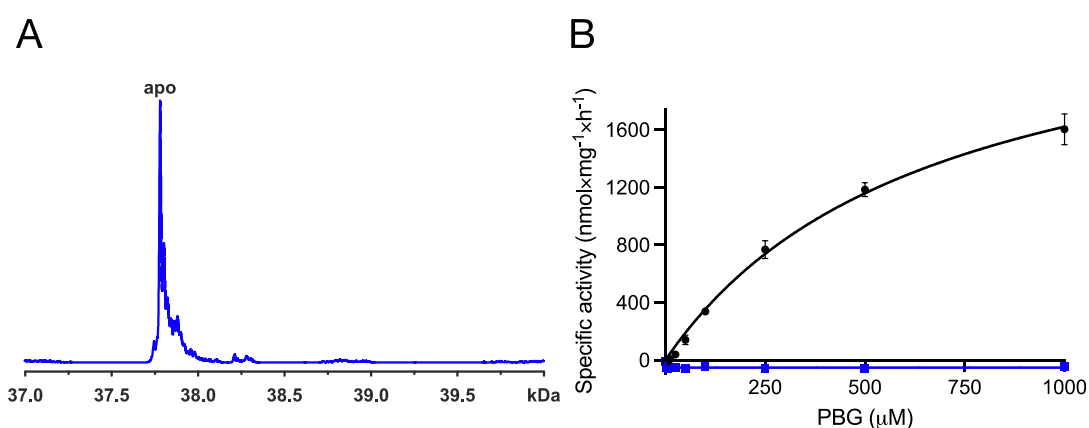

Figure 5-9. Biophysical characterisation of p.K98R. (A) Charge-deconvoluted ESI FT-ICR mass spectrum of p.R149Q in denaturing conditions at 1 µM concentration. **(B)** The catalytic activity of wt-HMBS (black) and p.K98R (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.3   Mutants yielding mainly one intermediate

5.5.3.1   Mutant p.R26H

Wt-HMBS presents the characteristic CD spectrum with a double dip at 208 and 222 nm, indicating a large presence of α-helices, and a very similar CD spectrum was obtained for the p.R26H mutant (Figure 5-10A). The CD results indicate that the mutant folds in a similar manner as wt-HMBS. The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). IEX was then used to separate the enzyme-intermediate complexes and when comparing the elution profile for wt-HMBS with p.R26H, it is apparent that the mutant mainly elutes as a single intermediate at ~26 % buffer B (Figure 5-10B). Wt-HMBS and p.R26H were further analysed by native PAGE and as can be seen in lane 1 (Figure 5-10C), all bands are present for wt-HMBS while p.R26H migrates as one major band (lane 2) with smaller bands corresponding to the peaks in

the chromatogram. As the isolated main peak of p.R26H (lane 3) migrates further than the corresponding band for the wt-HMBS ES_2 peak, the results are inconclusive as to which intermediate accumulates in this mutant. To confirm the identity of the intermediate, the mix sample was further analysed by FT-ICR MS, which confirmed that p.R26H mainly presents as a single intermediate in the ES_2-state with a small amount of ES_3 (Figure 5-10D).

To investigate if the Arg→His substitution in p.R26H has any effect on the stability of the enzyme we analysed the isolated wt-HMBS and mutant intermediates from IEX by DSF. As expected, based on previous reports [108], the results show a gradual decrease in stability and T_m-values from E_{holo} (78.19 ± 0.70 ℃) to ES_3 (71.01 ± 0.18 ℃) (Figure 5-10E). For the mutant, the ES_2 intermediate presented a T_m-value of 72.3 ± 0.1 ℃, corresponding to wt-HMBS in the ES_2-state (T_m = 72.44 ± 0.10 ℃). Furthermore, analysis of the enzymatic activity and determination of kinetic parameters was completed as previously reported [108]. Wt-HMBS had a specific activity of 757.70 ± 67.67 $nmol \times mg^{-1} \times h^{-1}$ whereas the mutant p.R26H was completely inactive (Figure 5-10F).
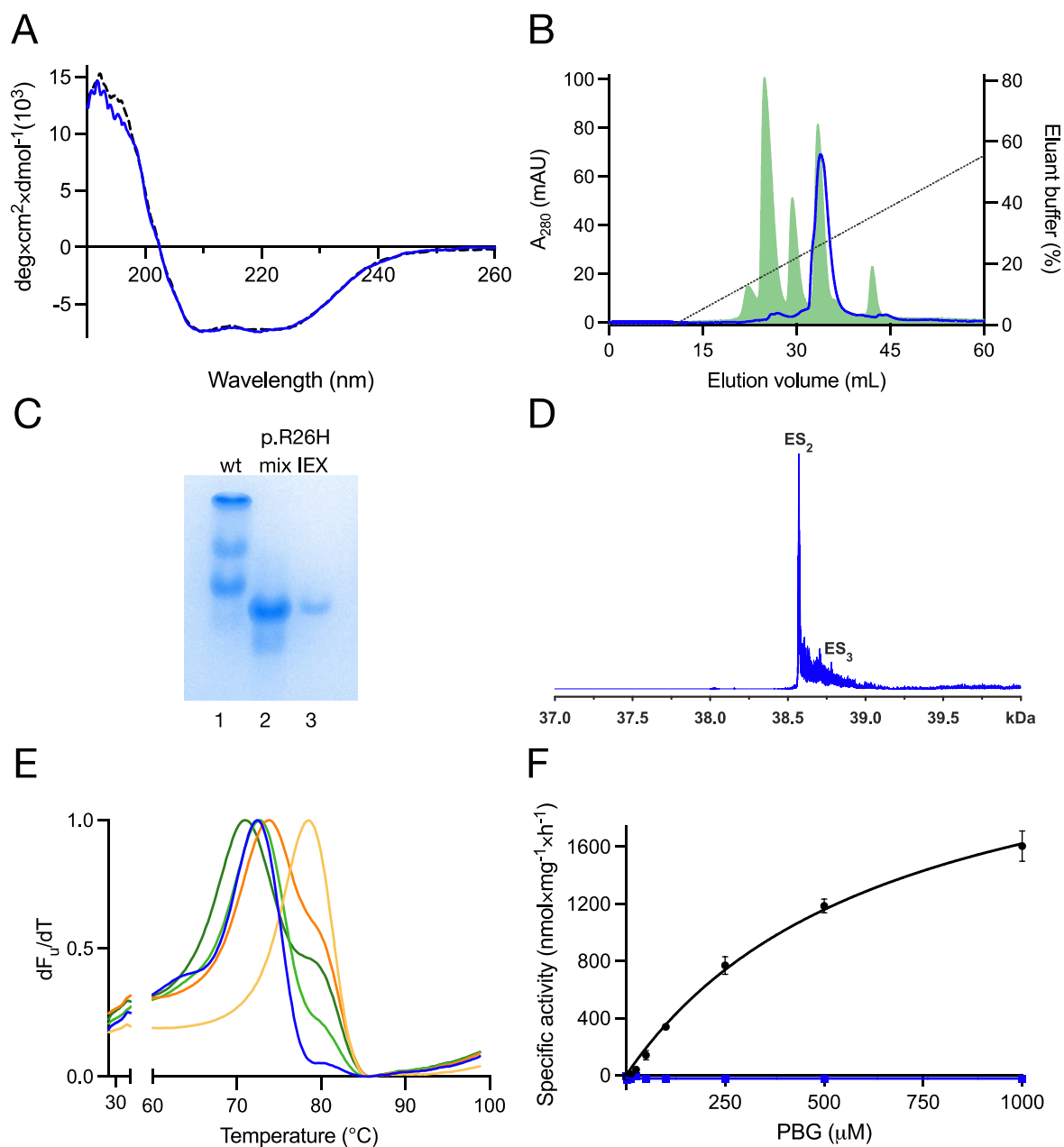
Figure 5-10. Biophysical characterisation of HMBS mutant p.R26H. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.R26H (blue) measured in triplicates; an unpaired t-test showed no significant difference between the two spectra. **(B)** IEX of wt-HMBS (green) and mutant p.R26H (blue), showing the separation and isolation of the enzyme-intermediate complexes. **(C)** Native PAGE analysis of wt-HMBS (lane 1, wt), p.R26H (lane 2, mix) and the isolated main IEX-peak of p.R26H (lane 3; IEX). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.R26H in denaturing conditions at 1 µM concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of the isolated IEX peak of p.R26H (blue) with wt-HMBS enzyme intermediates (obtained by IEX) for reference; E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.R26H (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.3.2 Mutant p.D99N

CD analysis to investigate p.D99N mutant folding indicated similarity to wt-HMBS, the small difference observed in the spectra is likely due to a difference in concentration (Figure 5-11A). The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). IEX chromatography was then used to separate the enzyme-intermediate complexes and by comparing the elution profile for wt-HMBS with mutant p.D99N, it is apparent that also this mutant mainly elutes as a single intermediate. The identification of this intermediate is however inconclusive as the peak falls between the ES and ES_2 intermediate in wt-HMBS at ~25 % buffer B (Figure 5-11B). As can be seen in lane 1 of the native PAGE (Figure 5-11C), all bands are present for wt-HMBS while the p.D99N from mix migrates as one major band (lane 2), which can also be seen for the isolated main IEX-peak of p.D99N (lane 3). The native PAGE analysis is also inconclusive concerning the identity of the intermediate, as the bands do not correspond to the wt-HMBS intermediate bands. The p.D99N mix was further analysed by FT-ICR-MS to determine the accumulating intermediate. The FT-ICR-MS analysis of p.D99N confirmed that it mainly presents a single intermediate in the ES_2-state (Figure 5-11D). Moreover, the DSF analysis of p.D99N shows a T_m-value of 70.04 ± 0.04 ℃ which is lower than all the wt-HMBS intermediates, ranging from 78.19 ± 0.70 ℃ to 71.01 ± 0.18 ℃ (Figure 5-11E). Finally, the mutant had a specific activity of 7.91 ± 4.10 nmol×mg^{-1}×h^{-1}, compared with 757.70±67.67 nmol×mg^{-1}×h^{-1} for wt-HMBS, i.e., showing a relative activity of 1 % (Figure 5-11F) with a V_{max} = 38.09 ± 5.45 and K_m = 170.50 ± 71.95 (Table 6).
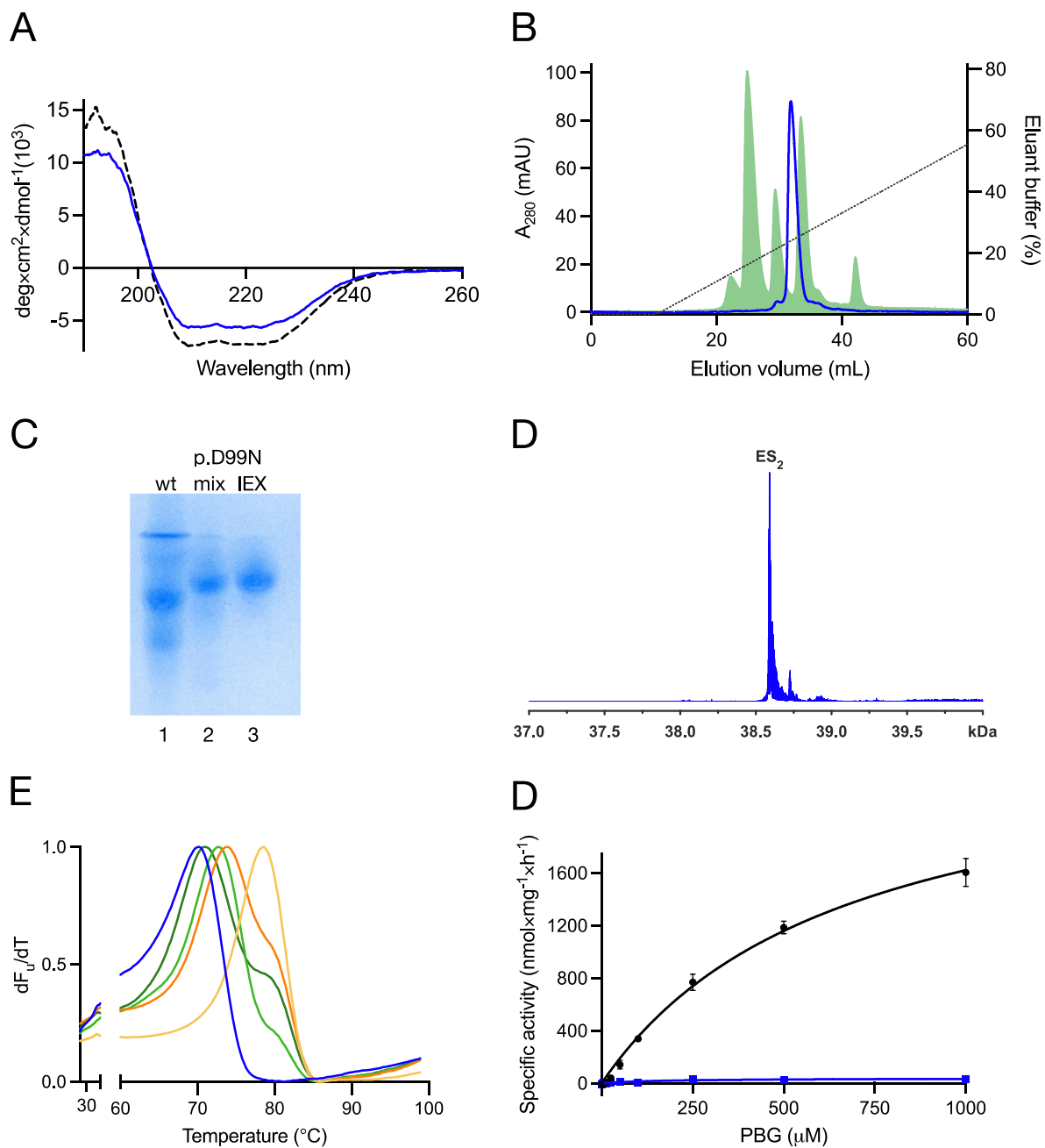
Figure 5-11. Biophysical characterisation of HMBS mutant p.D99N. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.D99N (blue) measured in triplicates; an unpaired t-test showed no significant difference between the two spectra. **(B)** Wt-HMBS (green) and mutant p.D99N (blue). Enzyme-intermediate complexes are separate, and denoted peaks isolated by IEX. **(C)** Native PAGE analysis of wt-HMBS (lane 1; wt), p.D99N (lane 2; SEC) and the isolated main peak of p.D99N (lane 3; IEX). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.D99N in denaturing conditions at 1 µM concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of the isolated IEX peak of p.D99N (blue) with wt-HMBS enzyme intermediates for reference; E_{holo} (yellow), ES (orange), ES2 (light green) and ES3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.D99N (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.4 Mutants yielding a distribution of various intermediates

5.5.4.1 Mutant p.L30P

CD analyses for p.L30P showed different spectra (Figure 5-12A) and significant difference in secondary structure content, as compared to wt-HMBS. The higher α-helical content relative to wt-HMBS, can be seen in Figure 5-12A, where the mutant demonstrates distinct deeper dips at 208 and 222 nm. The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). The IEX chromatogram for p.L30P (Figure 5-12B) shows one main peak seemingly corresponding with ES_2 in wt-HMBS with two smaller peaks on each side. The two main peaks from IEX, B and C, were isolated and further analysed by native PAGE and as seen in the gel they seem to correspond to the ES_2 (lane 3) and ES_3 (lane 4) (Figure 5-12C). The mix (lane 2) demonstrates a distribution of intermediates.

p.L30P mix was further analysed by FT-ICR-MS (Figure 5-12D) and the analysis confirmed that it presents a distribution of the enzyme intermediates apoenzyme, E_{holo}, ES_2 and ES_3, no ES was observed as also seen in the native PAGE for mix (lane 2). DSF show that mutant p.L30P is displaying a lower thermostability compared to wt-HMBS for the intermediates E_{holo}, ES_2 and ES_3 with a temperature decrease of ~4 ºC for all intermediates (Figure 5-12E). The activity assays show no measurable activity, the mutant p.L30P was therefore inactive (Figure 5-12F).
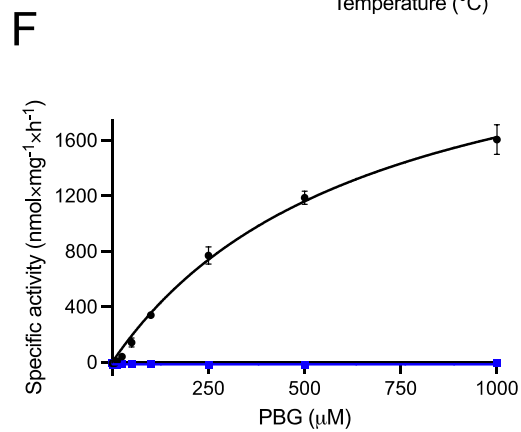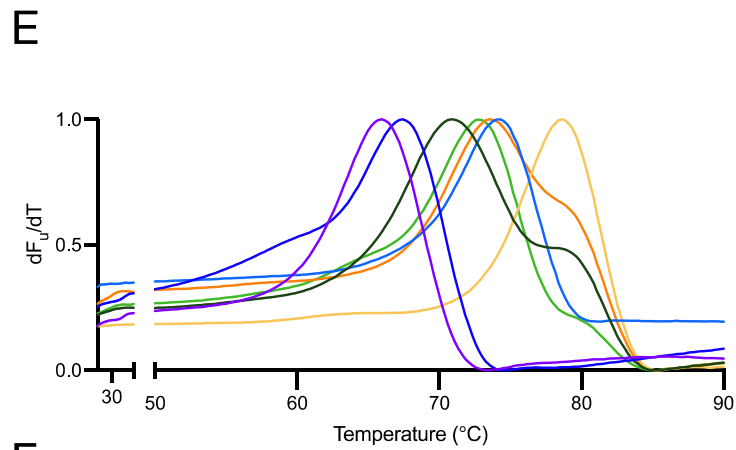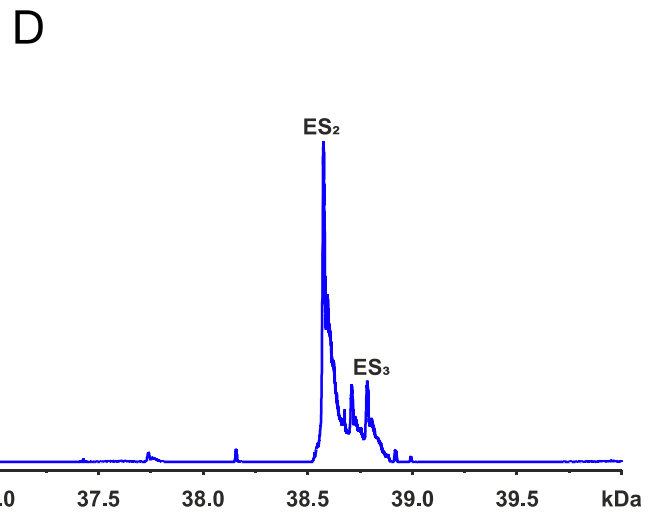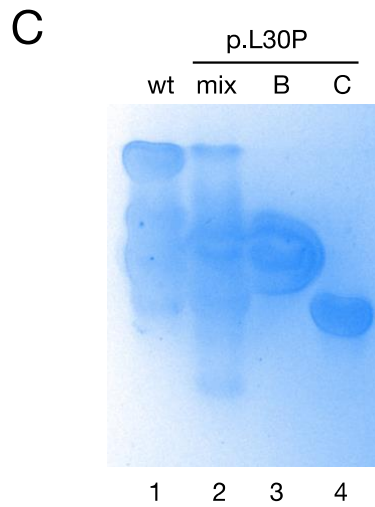
A

B

C

wt | p.L30P
mix | B | C

1   2   3   4

D

ES₂

ES₃

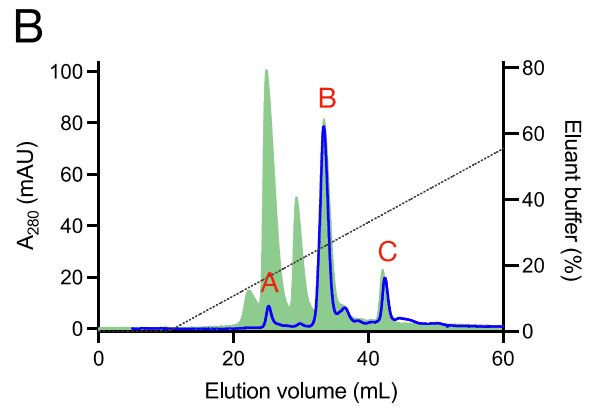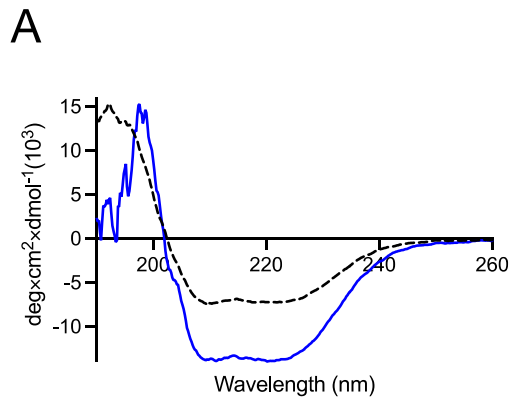37.0   37.5   38.0   38.5   39.0   39.5   kDa

E

F

47

Figure 5-12. Biophysical characterisation of HMBS mutant p.L30P. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.L30P (blue) measured in triplicates. T-test showed a significant difference in α-helix and β-sheet content (Table 5). **(B)** Enzyme-intermediate complexes are separated for wt-HMBS (green) and mutant p.L30P (blue) by IEX, and the isolated peaks for the mutant are labelled (A,B,C). **(C)** Native PAGE analysis of wt-HMBS (lane 1; wt), p.L30P mix (lane 2; SEC), isolated IEX peak B (lane 3) and peak C (lane 4). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.L30P in denaturing conditions at 1 μM concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of isolated p.L30P IEX peaks A (light blue), B (dark blue) and C (purple), with wt-HMBS enzyme intermediates for reference; E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.L30P (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.4.2 Mutant p.R167Q

CD analysis to investigate folding indicates that the α-helical content is lower in the mutant compared to wt-HMBS (Figure 5-13A, Table 5). The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). The mutant p.R167Q was further analysed by IEX and as can be seen in the chromatogram, it has a similar elution profile as wt-HMBS with well separation in the distribution of enzyme-intermediate complexes, showing peaks corresponding to the apoenzyme, E_{holo}, ES, ES_2 and ES_3 (Figure 5-13B). The apoenzyme of Wt-HMBS starts eluting at ~22 % buffer B however the mutant starts eluting at ~26 % buffer B. A similar trend can be seen in the native PAGE analysis where the p.R167Q mix (lane 2) migrates further compared to wt-HMBS (lane 1) (Figure 5-13C). The four main IEX peaks were isolated and analysed by native PAGE, producing well separated bands corresponding to A; E_{holo}, B; ES, C; ES_2 and D; ES_3.

p.R167Q mix was further analysed by FT-ICR-MS and the analysis confirmed that it presents a distribution of the enzyme intermediates apoenzyme, E_{holo}, ES, ES_2 and ES_3 (Figure 5-13D). The DSF analysis confirms very stable intermediates corresponding to the thermostability of the wt-HMBS intermediates with T_m-values for wt-HMBS ranging from 78.19 ± 0.70 °C to 71.01 ± 0.18 and T_m-values for R167Q 78.75 ± 0.07 °C to 70.41 ± 0.03 °C (Figure 5-13E). Wt-HMBS had a specific activity of 757.70 ± 67.67 nmol$\times$mg$^{-1}\times$h^{-1} whereas the mutant p.R167Q was completely inactive (Figure 5-13F).
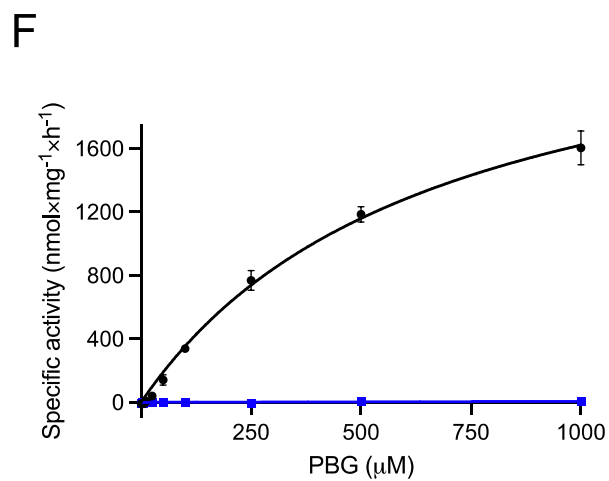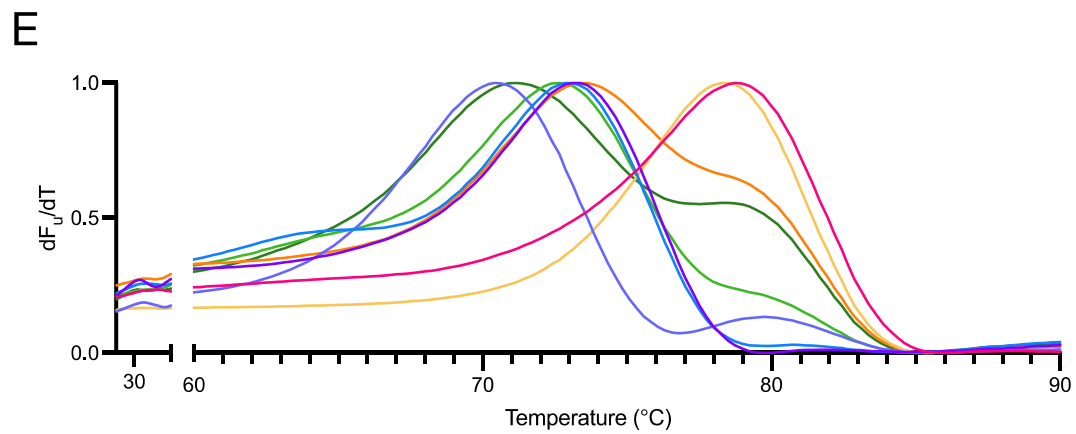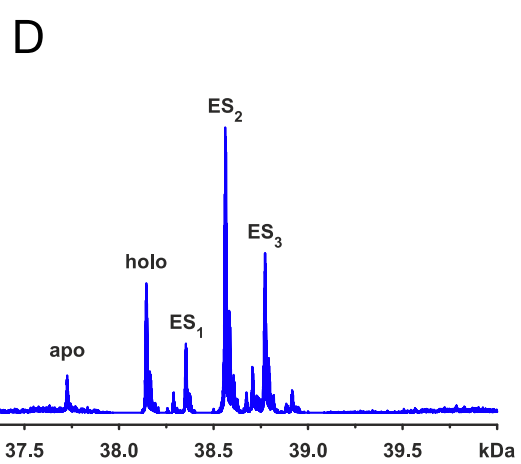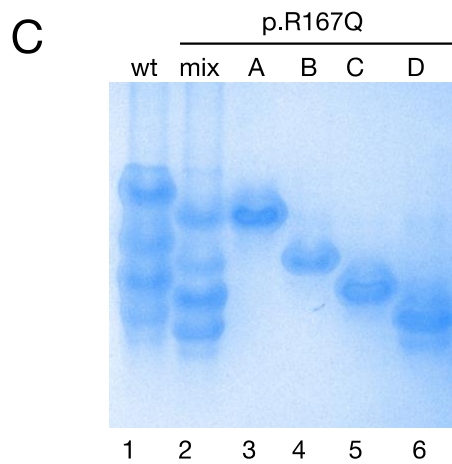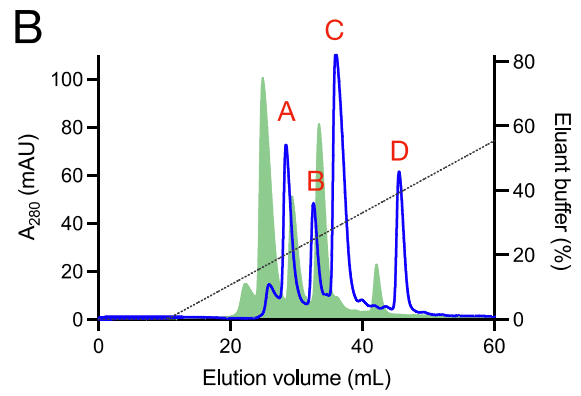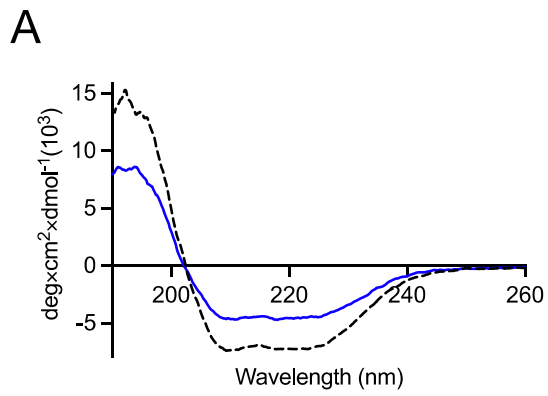
49

Figure 5-13. Biophysical characterisation of HMBS mutant p.R167Q. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.R167Q (blue) measured in triplicates. Unpaired t-test shows a statistically significant difference ($p < 0.001$) for α-helix content. **(B)** Enzyme-intermediate complexes are separated, wt-HMBS (green) and mutant p.R167Q (blue), and denoted peak isolated by IEX. **(C)** Native PAGE analysis of wt-HMBS (lane 1; wt), p.R195C (lane 2; SEC), IEX peak A (lane 3; A), peak B (lane 4; B), peak C (lane 5; C) and peak D (lane 6; D). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.R195C in denaturing conditions at 1 µM concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of isolated p.R167Q IEX peaks A (pink), B (purple), C (light blue) and D (light purple), with wt-HMBS enzyme intermediates for reference; E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.R167Q (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.4.3 Mutant p.R195C

CD analysis to investigate mutant folding indicate folding very similar to wt-HMBS (Figure 5-14A). The estimated secondary structure content as calculated by BestSel for wt-HMBS and analysed mutants is presented in Table 5 (Chapter 5.8). IEX analysis to separate the intermediates presented an elution profile displaying a distribution of enzyme-intermediate complexes, it is however inconclusive in determining which enzyme-intermediate complexes are present in comparison to wt-HMBS due to overlapping of peaks (Figure 5-14B). The three main peaks from IEX (A, B and C, respectively) were isolated and further analysed by native PAGE and as seen in from the gel (Figure 5-14C), peak A (lane 2) presents a smear at the top with a band further down, above the wt-HMBS bands, peak B (lane 3) corresponds to the band for E_{holo} in wt-HMBS and peak C (lane 4) is inconclusive as it migrates in between ES2 and ES3 in wt-HMBS.

p.R195C mix was further analysed by FT-ICR-MS and the analysis confirmed that it presents a distribution of the enzyme intermediates apoenzyme, E_{holo}, ES, ES_2 and ES_3 (Figure 5-14D). The three main IEX peaks were then further analysed by DSF (Figure 5-14E). The T_m-values for wt-HMBS ranges from 78.19 ± 0.70 ºC to 71.01 ± 0.18 ºC whereas p.R195C is displaying a lower thermostability with 41.98 ± 0.07 ºC for the E_{apo}, 71.79 ± 0.07 ºC for the E_{holo} intermediate and 69.09 ± 0.02 ºC for ES_2. Wt-HMBS had a specific activity of 757.70 ± 67.67 nmol$\times$mg$^{-1}\times$h^{-1} whereas the mutant p.R195C has very little residual activity at 5 nmol$\times$mg$^{-1}\times$h^{-1} (Figure 5-14F) with a $V_{max} = 35.34 \pm 11.19$ and $K_m = 279.60 \pm 226.11$ (Table 6).
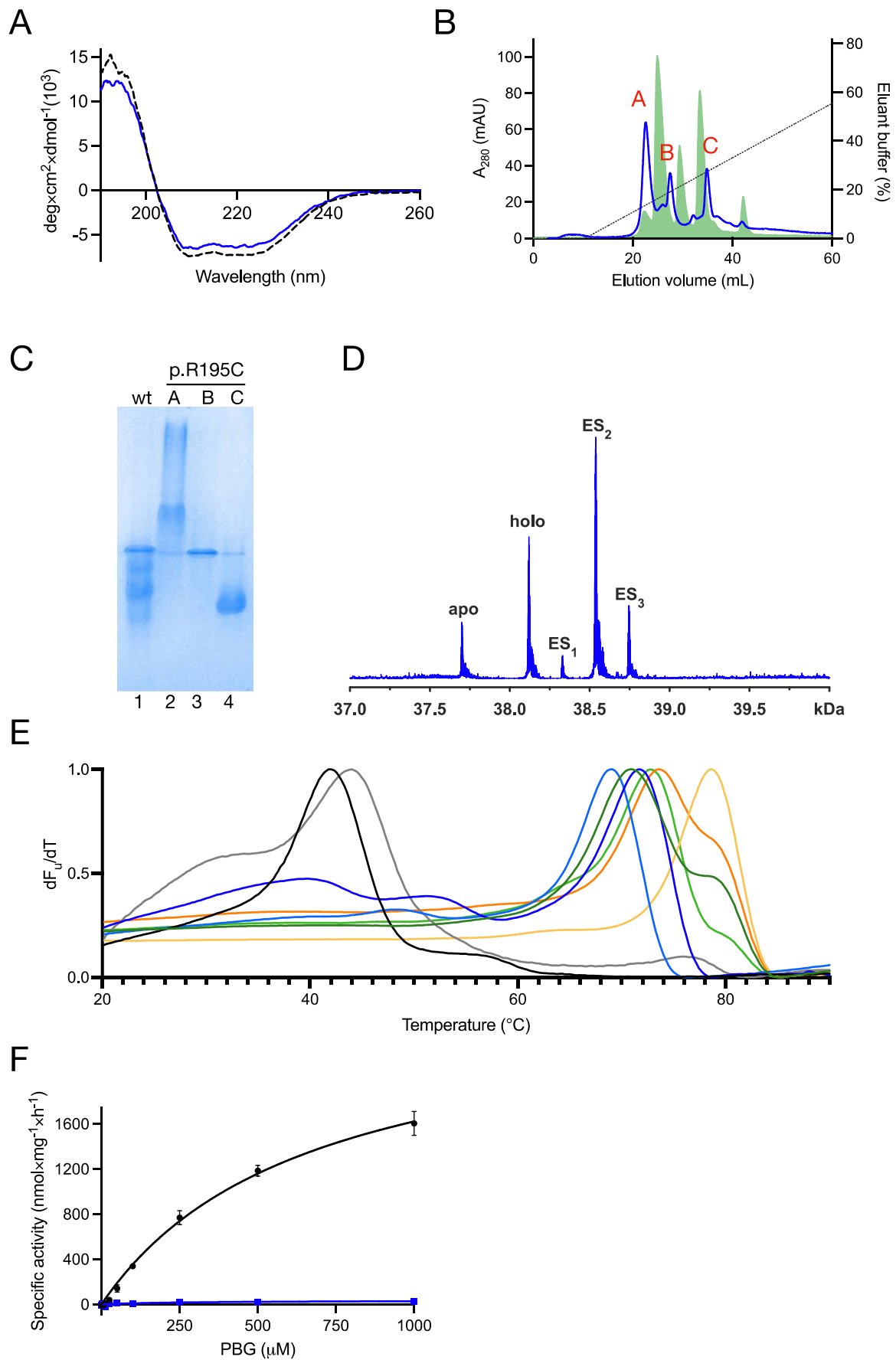
Figure 5-14. Biophysical characterisation of HMBS mutant p.R195C. **(A)** Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.R195C (blue) measure in triplicates. Unpaired t-test showed no significant difference between the two spectra. **(B)** Enzyme-intermediate complexes are separated, wt-HMBS (green) and mutant p.R195C (blue), and denoted peak isolated by IEX. **(C)** Native PAGE analysis of wt-HMBS (lane 1), p.R195C isolated peak A (lane 2), peak B (lane 3) and peak C (lane 3). **(D)** Charge-deconvoluted ESI FT-ICR mass spectrum of p.R195C in denaturing conditions at 1 μM concentration. **(E)** Thermal denaturing profile presented as the first derivative (dF_u/dT) of isolated p.R195C IEX peaks A (black), B (dark blue) and C (light blue), with wt-HMBS enzyme intermediates for reference; E_{apo} (grey), E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). **(F)** The catalytic activity of wt-HMBS (black) and p.R167Q (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.4.4 Mutant p.E223K

The Far-UV CD analysis of mutant p.E223K indicates a slight change in secondary structure content as compared to wt-HMBS (Figure 5-15A). The IEX chromatogram displays a distribution of enzyme intermediates in mutant p.E223K with some correspondence to the distribution of intermediates in wt-HMBS (Figure 5-15B). The four main IEX peaks (A-D) were isolated and further analysed by native PAGE seemingly corresponding to E_{apo}, E_{holo}, ES_2 and ES_3 in wt-HMBS (Figure 5-15C).

p.E223K mix was further analysed by FT-ICR-MS and the analysis confirmed that it presents a distribution of the enzyme intermediates apoenzyme, E_{apo}, E_{holo}, ES, ES_2 and ES_3 (Figure 5-15D). p.E223K is displaying a slightly lower thermostability compared to the corresponding intermediate in wt-HMBS as seen in Figure 5-15E. Steady-state kinetic assays demonstrate that the mutant is completely inactive (Figure 5-15F).
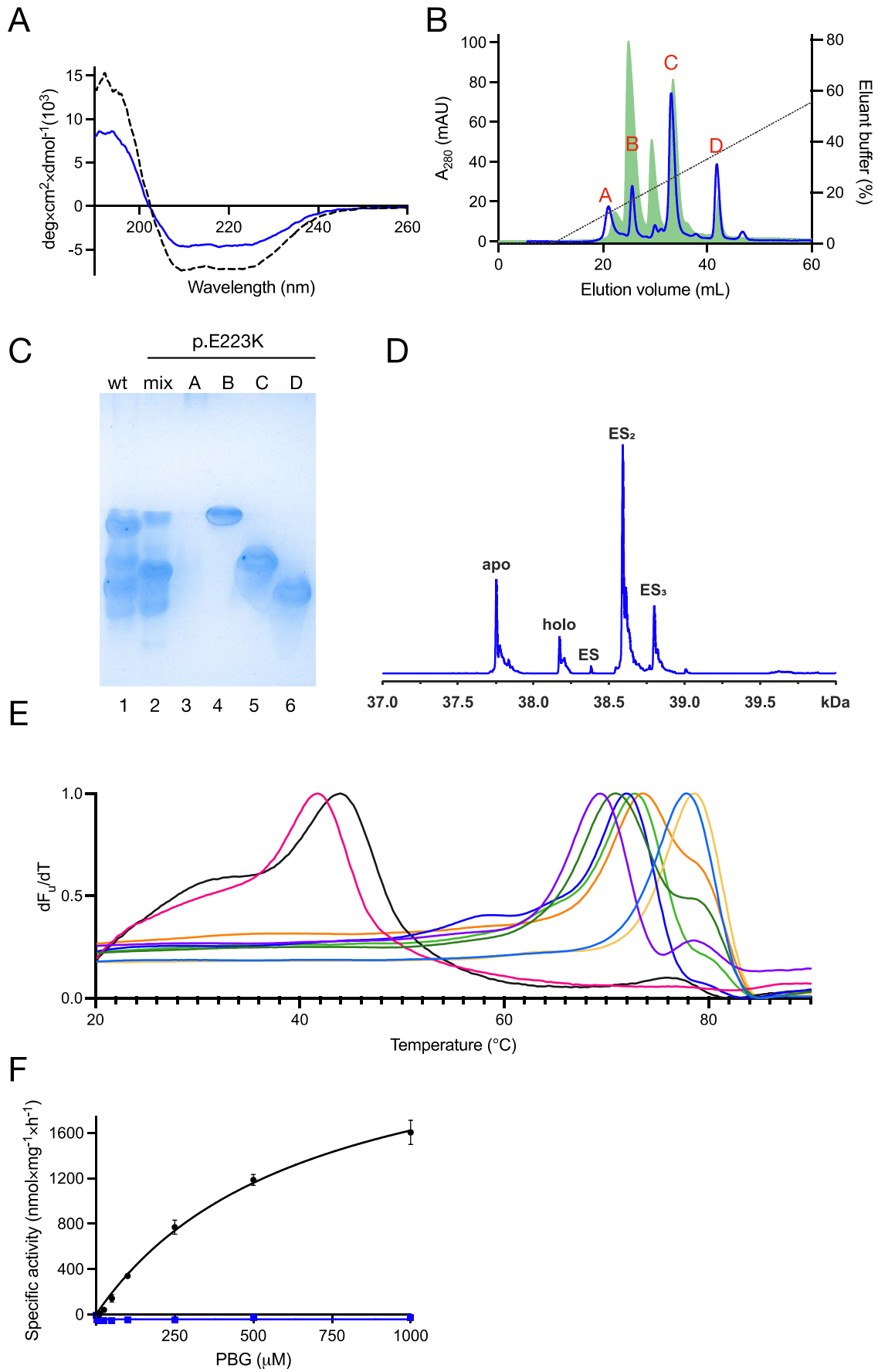
Figure 5-15. Biophysical characterisation of HMBS mutant p.E223K. (**A**) Far-UV CD spectroscopy of wt-HMBS (grey dotted line) and p.RE223K (blue) analysed in triplicates. An unpaired t-test shows significant differences, i.e. $p<0.004$ for α-helix, $p<0.001$ for β-sheets and others an $p<0.008$ for turns. (**B**) Enzyme-intermediate complexes are separated, wt-HMBS (green) and mutant p.E223K (blue), and denoted peak isolated by IEX. (**C**) Native PAGE analysis of wt-HMBS (lane 1), p.E223K mix (lane 2), isolated peak A (lane 3), peak B (lane 4), peak C (lane 5) and peak C (lane 6). (**D**) Charge-deconvoluted ESI FT-ICR mass spectrum of p.E223K in denaturing conditions at 1 μM concentration. (**E**) Thermal denaturing profile presented as the first derivative (dF_u/dT) of isolated p.E223K IEX peaks A (pink), B (light blue), C (dark blue) and D (purple), with wt-HMBS enzyme intermediates for reference; E_{apo} (black), E_{holo} (yellow), ES (orange), ES_2 (light green) and ES_3 (dark green). (**F**) The catalytic activity of wt-HMBS (black) and p.E223K (blue) as a function of PBG concentration, fitted to Michaelis-Menten enzyme kinetics model using GraphPad Prism.

5.5.5 Incubation of mutant p.R195C with PBG substrate

Due its direct contact with the cofactor through a salt bridge, we were interested to investigate what would happen to mutant p.R195C and its enzyme intermediate distribution upon incubation with excess PBG. An attempt was therefore made to incubate p.R195C with excess PBG substrate to observe if any changes occurred to the enzyme-intermediate distribution. To our surprise, the FT-ICR MS analysis of p.R195C after incubation with excess PBG shows that it accumulates one intermediate in the ES_4-state (Figure 5-16A), suggesting that there is turnover in the intermediates, but that HMB product release is hindered. Furthermore, as seen in Figure 5-16B, p.R195C incubated with PBG at 4 ºC overnight and analysed by native PAGE (lane 3), showed a unique band, lower than those for wt-HMBS mix (lane 1) and p.R195C mix (lane 2), implying an additional charge or a different size or conformation of the mutant as it migrates further on the gel. This PBG-affected intermediate is important to shed light on the HMB exit mechanism and X-ray crystallography was therefore pursued in order to gain valuable structural information, see chapter 5.6.
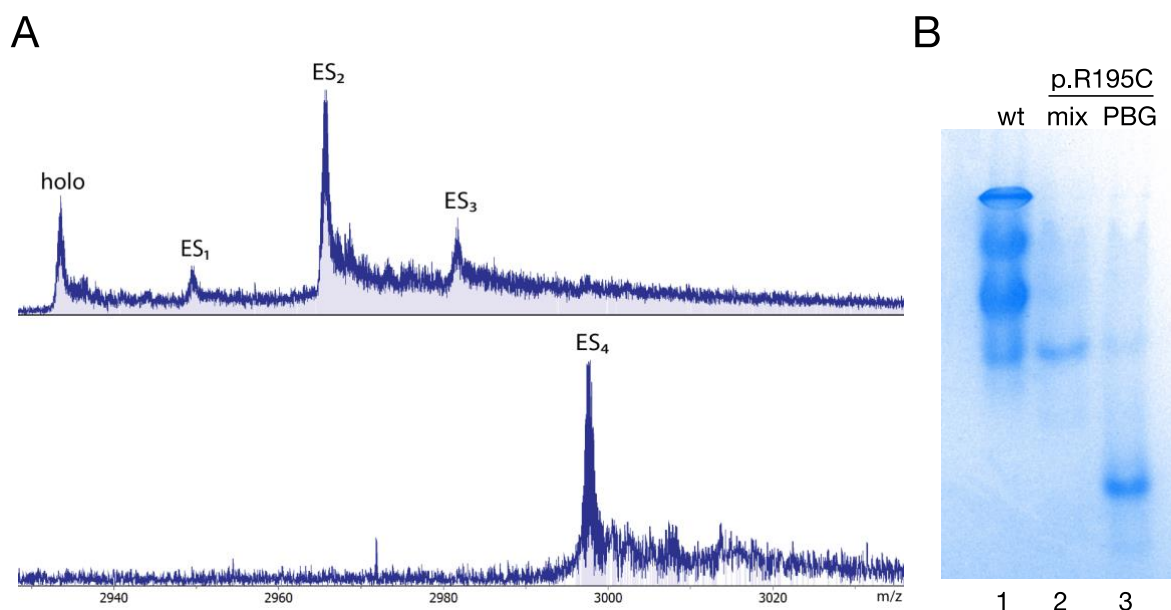
Figure 5-16. Mutant p.R195C incubated with PBG substrate. **(A)** Native PAGE analysis of wt-HMBS (lane 1), p.R195C (lane 2) and p.R195C incubated with excess PBG substrate. **(B)** Native ESI FT-ICR mass spectrum of p.R195C incubated with PBG, the mass spectrum is not deconvoluted.

5.6 X-ray Crystallography

Screening for crystallisation conditions was started for mutants p.R26H and p.R195C due to their interest to shed light on the catalytic and elongation mechanism of HMBS. The biophysical characterisation of mutant p.R26H shows that it stops in the ES_2 state of the elongation process, this mutant was therefore screened using the isolated intermediate in the attempt to get a crystal structure for p.R26H in the ES_2 state (Chapter 5.5.3.1). The crystallisation of mutant p.R195C was screened using a mix sample incubated with excess PBG because the biophysical characterisation results showed that this mutant accumulates the ES_4 intermediate in presence of PBG (Chapter 5.5.5).

Currently our crystallisation attempts for p.R195C have been unsuccessful. On the other hand, the initial screening of p.R26H presented prominent conditions of 0.1 M HEPES pH 7.5, 0.8 M NaH_2PO_4, 0.8 M KH_2PO_4, for crystal growth and optimisation was conducted by altering the salt and precipitate concentrations along with testing conditions where HMBS have previously been crystallised [72, 128]. From the first round of optimisation for p.R26H six single crystals were picked for diffraction testing from condition B1 (9.35 mg/mL p.R26H, 20% PEG 3350 and 0.1 M Ammonium Citrate pH 5.5) and E1 (6.35 mg/mL p.R26H, 20% PEG 3350 and 0.1 M Ammonium Citrate pH 5.5) on the optimisation plate, similar conditions as HMBS crystals were previously obtained [72]. Thin plate shape crystals,

having average size of ~50 µm on the longest edge, were mostly growing as bundles but were possible to separate from each other. Crystals diffracted up to 3 Å resolution but only from one orientation. Diffraction drastically changed upon rotating the crystal reaching only 4–4.5 Å resolution and being highly anisotropic. Although no data was collected crystals showed potential, and therefore data collection will be pursued after further optimisation of the crystals.
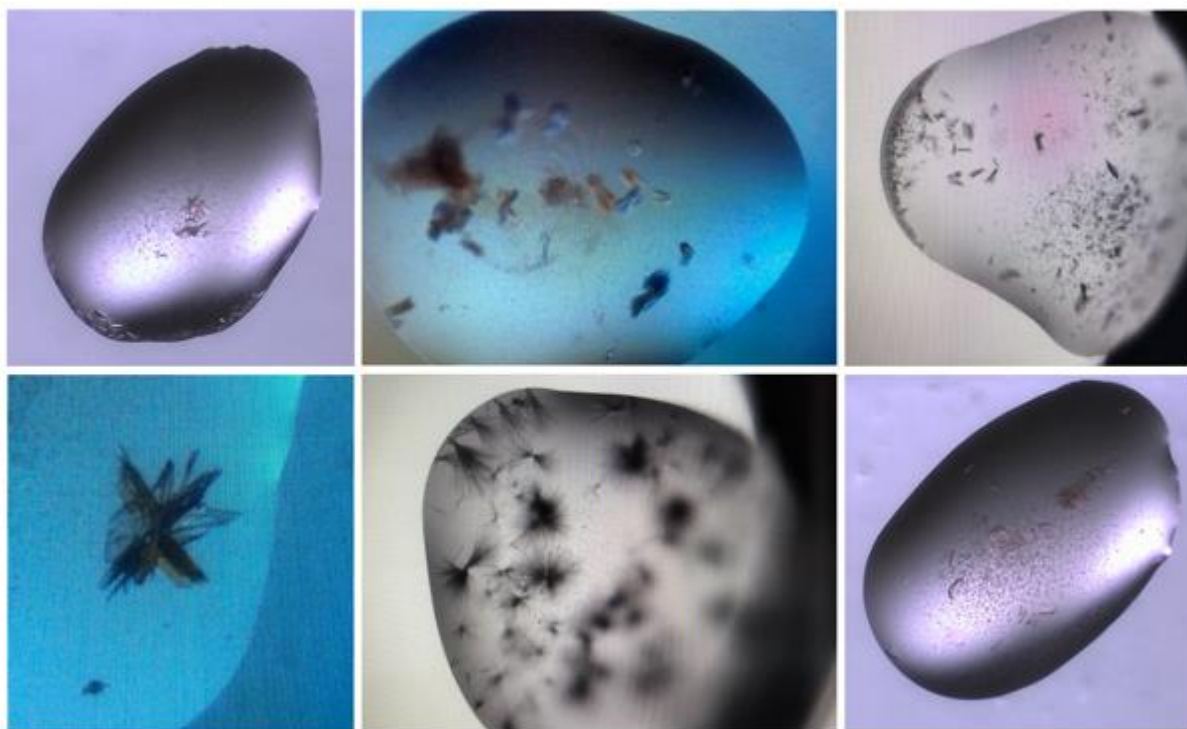

Figure 5-17. Crystals grown in the initial crystallography screenings of HMBS mutant p.R26H.
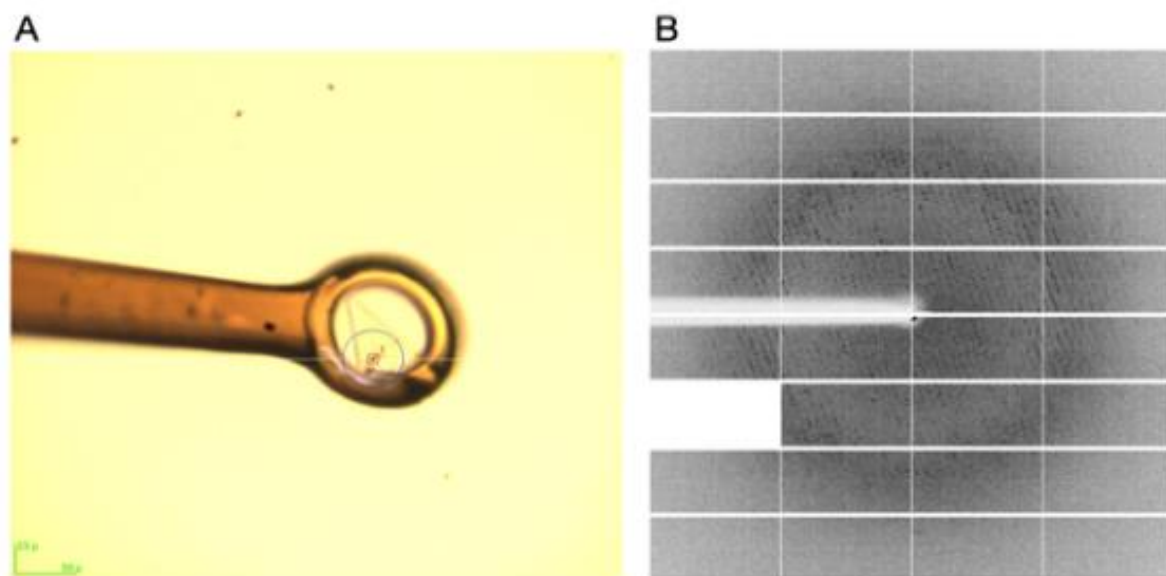

Figure 5-18. Diffraction data from crystal picked from p.R26H. (A) HMBS crystal picked from the optimisation of p.R26H. (B) X-ray diffraction pattern of HMBS crystal, showing diffraction to ~3Å resolution. Detector limit was set to 2.8Å.

5.7 Summary of biochemical and biophysical characterisation

Using a combination of IEX, native PAGE and FT-ICR MS we have been able to identify which enzyme-intermediate complexes are present for each mutant, and a summary of these is listed in Table 4, together with their T_m-values, as determined through DSF analysis.

Table 4. Summary of enzyme-intermediate complexes present in each mutant with the thermostability in °C (T_m-value) for each intermediate as measured by DSF.

	T_m (°C)[a]					
Protein	*Apo*	*E_{holo}*	*ES*	*ES_2*	*ES_3*	*ES_4*
Wild-type	43.90 ± 0.32	78.19 ± 0.70	73.33 ± 0.68	72.32 ± 0.72	71.01 ± 0.18	
p.R26H				72.44 ± 0.10	X	
p.L30P	X	74.27 ± 0.03		67.42 ± 0.12	65.14 ± 0.26	
p.K98R	X					
p.D99N				70.04 ± 0.04		
p.R149Q	42.30 ± 0.07					
p.R167Q		78.75 ± 0.07	73.24 ± 0.03	72.97 ± 0.09	70.41 ± 0.03	
p.R195C	41.98 ± 0.07	71.79 ± 0.07	X	69.09 ± 0.02	X	X
p.E223K	41.80 ± 0.28	77.96 ± 0.04	X	72.08 ± 0.04	69.48 ± 0.04	

X = intermediate present but not able to measure thermostability due to too low protein concentration.
[a]T_m-values are provided as an average ± SD of four in-plate measurements for the mutants and four between-plates measurements for wt-HMBS.

Table 5. BestSel prediction of secondary structure content based on Far-UV CD spectrum at 190–250 nm.

Protein	*α-helix*	*β-sheets*	*Turn*	*Others*
Wild-type	18.6 ± 1.7	28.2 ± 2.1	12.2 ± 0.8	40.9 ± 0.8
p.R26H	15.5 ± 3.5	30.6 ± 3.1	11.9 ± 0.9	41.9 ± 0.9
p.L30P	26.7 ± 0.2*	17.2 ± 3.3**	13.6 ± 0.8	42.4 ± 1.6
p.D99N	15.2 ± 0.6	31.2 ± 1.1	12.5 ± 0.5	41.2 ± 2.1
p.R149Q	16.6 ± 1.6	30.8 ± 3.0	12.6 ± 0.6	39.9 ± 1.1
p.R167Q	10.6 ± 0.7*	33.3 ± 3.6	13.0 ± 0.7	43.0 ± 2.5
p.R195C	17.5 ± 2.6	29.2 ± 4.1	13.4 ± 0.1	40.0 ± 2.5
p.E223K	31.3 ± 2.6**	13.3 ± 1.9*	8.6 ± 1.0***	46.8 ± 1.1*

CD analysis to calculate secondary structure as predicted by BestSel and statistical analyses using multiple unpaired t-tests between wt-HMBS and mutants.
T-tests indicate no significant difference between wt-HMBS and mutants p.R26H, p.D99N, p.R149Q and p.R195C. * P-value < 0.001, ** P-value < 0.004, *** P-value < 0.008 for significant differences as compared to wt-HMBS and calculated by unpaired t-test. Data presented as mean ± SD.

Table 6. Summary of results from the enzyme activity assays for wt-HMBS and mutants.

Protein	Relative activity (%)	Specific activity (nmol×mg^{-1}×h^{-1})	$V_{max} \pm$ SEM (nmol×mg^{-1}×h^{-1})	$K_m \pm$ SEM (PBG in µM)
Wild-type	100	757.70 ± 67.67	2697.00 ± 140.50	663.30 ± 66.71
p.R26H	0	ND	ND	ND
p.L30P	0	ND	ND	ND
p.K98R	0	ND	ND	ND
p.D99N	1.0	7.91 ± 4.10	38.09 ± 5.45	170.50 ± 71.95
p.R149Q	0	ND	ND	ND
p.R167Q	0.1	0.73 ± 3.70	ND	ND
p.R195C	0.7	5.00 ± 0.00	35.34 ± 11.19	279.60 ± 226.11
p.E223K	0	ND	ND	ND

ND = not detected. The catalytic activity was measured as a function of PBG concentration. Enzyme kinetic analysis were conducted using a range of 3.126–1000 µM PBG with 0.5 mg/mL enzyme. Steady-state kinetic parameters were analysed using the Michaelis-Menten enzyme kinetics model in GraphPad Prism.

6   Discussion

Biochemical and biophysical characterisation of proteins offers valuable information on protein conformation, stability, and function. Furthermore, using protein variants provides insights into pathomechanisms on a protein level, which helps to better understand the molecular background of a disease and the possibility to establish genotype-phenotype relationships in future work. In this thesis, biochemical and biophysical characterisation of several disease-associated *HMBS* missense variants has been conducted along with attempts of X-ray crystallisation of HMBS in different intermediate states, which could potentially provide valuable structural information and is essential to fully understand its function. All except two of the selected mutations are found in the Norwegian AIP population, where some are already described in the literature, and some are reported without any functional characterisation or not described at all. Characterisation of disease-associated variants can also provide a fundament for and valuable insight aiding in the development for more effective AIP treatments in the future.

6.1   HMBS mutants with complete folding defects

In order to characterise the selected *HMBS* variants, recombinant wt-HMBS and mutants were expressed in *E. coli* according to the standard protocol developed for wt-HMBS. An expression test carried out using SDS-PAGE analysis of bacterial cultures indicated that all mutants were expressed, however mutants p.E223K, p.L238P, p.E250Q and p.A252P show a slightly lower intensity relative to the uninduced sample, compared to the others (Figure 5-1). Following expression and purification using affinity- and size-exclusion chromatography, these mutants also yielded very little to no protein. The initial expression test does not differentiate between soluble and insoluble protein, however, since the protein purification did not produce any yield, this suggests that the above-mentioned mutants are insoluble. It was decided not to pursue the characterisation of these mutants in the context of this thesis. To further investigate why they did not purify, the expression and purification protocol could be optimised to better fit each mutant to increase the solubility and yield.

We hypothesise that these mutants are apoenzymes, i.e., unable to assemble the DPM cofactor. HMBS in the apo-form can be relatively stable [129], however, if the interaction network between the clearly defined domains is disrupted either by breaking of important interactions (e.g., p.E250Q) or the interdomain architecture (e.g., p.A339D) is disturbed, the overall stability of the enzyme is lost, as described below.

Leu238 sits far away from the active-site and therefore cannot interact directly with the DPM cofactor [67]. However, due to its position as a hinge residue, and with interactions with other residues within domain 1, a leucine to proline substitution at residue 238 could cause severe structural effects as these interactions are lost. Case reports for individuals with a p.L238P variant present inconsistent phenotypes, where Yarra *et al.* [112] present a patient requiring monthly hospitalisation whereas Fu *et al.* [94] predicts a mild clinical phenotype.

Residue Ala252 is in close proximity to the cofactor-binding loop and is involved in loop mobility [65]. The reported p.A252P variant [114] likely causes unfavourable proline interactions in the $\alpha 1_3$-helix, and in doing so disrupts the orientation of the helix, thereby interfering with the loop mobility, restricting the movement in the enzyme, and thus altering the secondary structure.

Glu250 is located in domain 3 and does not directly interact with the cofactor, however, it forms a salt-bridge with Arg116, another important hinge residue. A previously characterised variant of Arg116, p.R116W, results in a stable apoenzyme [108]. This suggests that mutants at Glu250 cause more damaging effects on the overall stability with a larger folding defect compared to p.R116W.

Ala339 is a residue placed in the $\alpha 2_3$-helix. It is located far away from the active site. Structural investigations displayed mainly interactions to residues Gly335, Ile336, Leu342 and Leu343 through hydrogen bonds, indicating that Ala339 plays an important role in the stabilisation of the $\alpha 2_3$-helix. An alanine to aspartic acid does not interrupt inter-helical interactions, however, the introduction of charge will cause repulsion between helices $\alpha 2_3$ and $\alpha 1_3$, following a disruption of the interaction network between these two helices that affects the overall domain 3 architecture and thus the binding of the cofactor, leading to detrimental folding effects. Mutants p.E223K, p.L238P, p.E250Q and p.A252P all affect various important interdomain interactions and therefore, despite not being directly involved with binding of the DPM cofactor, lead to destructive folding effects.

6.2 HMBS mutants as stable apoenzymes

Residues located in and around the active site cleft and near the substrate-binding site, are likely to be involved the enzyme ability to bind the cofactor or substrates. The large interaction network around the cofactor includes a number of residues, amongst them are the conserved residues Lys98 and Arg149 [67]. Both are positively charged residues forming hydrogen bonds with the cofactor. Lys98 forms a salt bridge with the acetate cofactor of the

C1 and C2 rings [79]. The stability decreases when the cofactor is not assembled, and it has been suggested that this occurs due to domain movements and 'open' and 'closed' confirmations of the enzyme in response to cofactor binding [96, 129]. Variants hindering the binding of the cofactor will therefore have a more structural impact on the enzyme where the enzyme loses the interaction network mediated by the binding of the cofactor and possibly hinder the movement between the domains resulting in a loss of flexibility.

As demonstrated by IEX, native PAGE and MS, both p.K98N and p.R149Q mutants present as a single intermediate in the apo-form, suggesting they are incapable of binding the cofactor. However, in the cases of both p.K98R and p.R149Q, the lack of cofactor did not cause detrimental folding effects, as seen for the mutants p.E223K, p.L238P, p.E250Q and p.A252P, indicating that the latter mutants have a larger impact on general conformation and folding.

Investigating the thermostability of wt-HMBS demonstrates that the apo-form has a melting temperature at ~43 ℃ (Table 4). The thermostability of p.R149Q as measured by DSF was ~42 ℃. The same low stability can also be seen in the isolated apoenzyme fraction of p.R195C from IEX (Table 4). The apoenzyme have a ~50 % loss in stability compared to the isolated p.R195C E_{holo} intermediate (i.e., with bound cofactor) with melting temperature of ~78 ℃. A melting temperature of ~43 ℃ is not low for a protein compared to most proteins in the body, as most human proteins ranges from 40 to 50 ℃ in melting temperature [130]. However, it is relatively unstable compared with E_{holo} in wt-HMBS. The steady-state kinetics assays showed that both p.R149Q and p.K98R were completely inactive (Table 6), as expected.

6.3   HMBS variants accumulating single intermediate

Unlike the apoenzyme, some mutants are able to bind the cofactor followed by a number of PBG substrates but are catalytically blocked at one step in the elongation process. The result is a mutant trapped in one intermediate state, which then accumulates. Based on the IEX, native PAGE and MS results, we demonstrate that both p.R26H and p.D99N are catalytically blocked and accumulate enzyme intermediates in the ES_2-state. Interestingly, both mutants stop in ES_2. Despite numerous crystal structures, the complete understanding of the elongation mechanism is still unknown [63]. We do not have evidence of the exact catalytic steps for pyrrole elongation mechanism and cannot with certainty say if the ES_2 is formed by the addition of PBG substrates from E_{holo} and then ES, or if the endogenous HMBS from *E.*

coli produces HMB that is inserted in the apo-form leading to an ES_2 intermediate directly. Previous reports have found that apo-HMBS has a higher affinity for the HMB product compared to PBG, indicating that endogenous HMB can yield ES_2 directly [131].

To investigate if the Arg→His substitution in p.R26H and the Asp→Asn substitution in p.D99N have any effect on the stability of the enzyme we analysed the isolated ES_2-intermediate by DSF. For wt-HMBS, as expected and as previously reported [108], the results show a decrease in stability upon substrate incorporation, with T_m-values declining from ~78 ºC to 71 ºC for E_{holo} to ES_3 (Table 4 and Figure 5-10E). The mutant p.R26H has a T_m-value of 72 ºC, corresponding to wt-HMBS in the ES_2-state (72.3 ºC), suggesting that the mutation has no major structural effects on the structural integrity. This indicates that the protein is potent and ready for catalysis in its ES_2-form, however, the missing Arg26 is crucial for substrate binding and the elongation is blocked. Arg26 is a highly conserved arginine residue located at the putative substrate-binding site and has been found within hydrogen-bonding distance to the substrate in crystal structures [69, 79].

A previous study investigating the effects of p.R173W, found that this variant was also trapped in the ES_2-state, highlighting its importance in the formation of ES_3 [72]. As demonstrated by Sato *et al*. the three arginine residues Arg26, Arg167 and Arg173 all interact with the substrate analogue 2-I-PBG positioned in the substrate-binding site [128], also indicating that at least Arg26 and Arg173, possibly in concert, are crucial for the formation of ES_3.

The other mutant trapped in a single intermediate state, p.D99N, has a T_m-value of 70.0 ºC, this is lower than all intermediates in wt-HMBS. This slightly lower melting temperature ($\Delta T_m = 2.4$ ºC), indicates some structural effects making it less stable than in wt-HMBS and p.R26H. Its lower melting temperature is an indication that the structural defect because of the Asp→Asn substitution resembles what is observed for p.R173W [108]. Where residues such as Arg173 and Asp99, yielding lower thermostability when mutated, may have roles as structural coordinators, Arg26 probably participates mainly in directing incoming substrates through proton donation. This is also supported by the residual activity of p.D99N (~1 % relative to wt-HMBS), whereas p.R26H is completely inactive as it is the direct action of Arg26 that is absent.

6.4 Variants yielding a distribution of enzyme-intermediate complexes

The other group of variants characterised in this project are displaying a distribution of enzyme-intermediate complexes resembling those of wt-HMBS, suggesting there is residual activity and turnover. Based on IEX, native PAGE and MS data, mutants p.L30P, p.R167Q, p.R195C and p.E223K all exhibit a distribution of enzyme intermediates. Calculating the secondary structure content from CD analysis indicated that p.R195C is not significantly different from wt-HMBS whereas mutant p.R167Q results in a lower α-helix content ($p>0.001$, Table 5). Mutants p.L30P and p.E223K are both significantly different in secondary structure content with relatively high amount of α-helix content and lower amounts of β-sheets compared to wt-HMBS (Table 5). From IEX and native PAGE results it is apparent that p.R167Q has more distinctly separated intermediates like those observed in wt-HMBS, compared to p.R195C. The difference in p.R167Q and wt-HMBS is that the mutant elutes slightly later in IEX (21 ml vs 26.5 ml, respectively) and migrates further in the native PAGE suggesting a loss of a positive charge. Unlike p.R167Q, the peaks for mutant p.R195C are not fully resolved and separated in IEX, causing an overlap.

In HMBS with normal function, the HMB product is readily hydrolysed and cleaved, and for mutants with a distribution of enzyme intermediates like wt-HMBS, it would be natural to hypothesise some turnover and release of HMB. Interestingly, and despite having an intermediate distribution, the results show that the mutants do not necessarily have any catalytic activity, and the activity therefore does not reveal much information about the intermediates. Mutant p.L30P demonstrated no measurable activity indicating a complete inactive mutant. Mutants p.R167Q and p.R195C demonstrate a relative activity of 0.1 % and 0.7 % respectively. Interestingly, the previously reported p.R167W has residual activity [108] while p.R167Q in our studies does not, indicating that the Arg$\rightarrow$Glu substitution has a more severe catalytic effect on the enzyme. Mutants p.E223K and p.L30P both have a distribution of enzyme intermediates as well but have no enzyme activity indicating inactive mutants.

To further investigate the product turnover, p.R195C and p.R167Q were both incubated with excess substrate, resulting in an accumulation of p.R195C in the ES$_4$ state, however, no such accumulation was seen for p.R167Q. For p.R195C it is obvious that the mutation causes a hinder or an extreme delay in product release, whereas the glutamine substitution in Arg167 does not have the same effect. Through molecular dynamics simulations, Arg167 has been identified as a gatekeeper for HMB exit [88]. The results obtained in this work highlight its importance in the exit mechanism, however if that was the

only function of Arg167, the ES_4 intermediate should have been observed upon incubation with PBG, as was observed for p.R195C. These findings suggest Arg167 have other functions as well.

The thermostability of the p.R167Q enzyme intermediates is similar to those of wt-HMBS, with a decreasing thermal stability for each substrate bound. However, this is not the case for p.R195C, which has a lower thermostability with a 6.4 °C decrease in E_{holo}-state. The loss in thermostability indicates that this mutation affects the structural architecture to an extent that both decreases its overall stability and hinder product release. A crystal structure of p.R195C in the ES_4 could possibly provide details for this mechanism however the initial crystallisation attempts have been unsuccessful. Gaining a structure for the final intermediate would therefore be valuable in understanding the product release mechanism, and despite molecular dynamics simulations to map the exit mechanism attempts have been done, little is still known about the exit of HMB [89].

Leu30 mostly interacts through hydrophobic interactions, and with hydrogen bonds to the main chain, a mutation here would therefore not cause a big structural effect. However, it is located at the edge of an α-helix and a substitution to proline, such as the p.L30P variant, will possibly disturb the α-helix. This leads to a conformational change that could possibly distort the Arg26 residue by i.e., changing the orientation and thereby affect its role as proton donor. Glu223 is a residue positioned in domain 1 where it interacts with Val93 where it is involved with binding the two domains together. Introducing a lysine here will alter the charge from positive to negative and lead to a weaker stabilisation when connecting the domains together. Glu223 does not directly alter the active site and gives rise to a distribution of intermediates, however it is still an inactive mutant.

6.5   Limitations of this study

One of the limitations with the steady-state kinetics experiments performed in this project is that a plateau is not reached for wt-HMBS with the Michaelis-Menten kinetics analysis. The assay conditions in this project were adapted from Bustad *et al*. [108], expect for the use of erythroid instead of the ubiquitous isoform. The literature is inconsistent with regards to assay conditions, such as varying buffer, temperature, and time of reaction [110, 132, 133]. Additionally, the results are difficult to compare with the existing literature due to the use of different units.

7 Conclusion

In recent years, new emerging therapies have evolved using pharmacological chaperones, lipid nanoparticles, and gene and enzyme replacement therapy [6, 58], and the now established RNA interference therapy with Givosiran [134, 135]. The more we know and understand on a protein at the molecular level, the better we can aid in developing new mechanistic therapies for treatment of AIP. The main objective of this thesis was to gain biochemical and biophysical characteristics of disease-associated variants of HMBS. Although the findings from the biophysical characterisation alone do not explain the genotype-phenotype correlation directly, it provides valuable insights into the mutational effect which could be important when exploring and selecting alternative treatment options for AIP.

Out of the twelve disease-associated variants chosen for this project, we have obtained biochemical and biophysical characteristics for seven variants: p.R26H, p.L30P, p.D99N, p.R149Q, p.R167Q, p.R195C and p.E223K. p.K98R was partly characterised while the remaining four variants, p.L238P, p.E250Q, p.A252P and p.A339D, resulted in insoluble and unstable enzymes, probably in the apo-form.

It has proven difficult to classify the mutants into distinct groups based on their mutational effects as the results show that there is an overlap in folding defects and catalytic effects. However, p.L238P, p.E250Q, p.A252P and p.A339D, which we were not able to purify and hence deemed unstable and insoluble, can be classified as mutants causing folding defects. Mutants p.K98R and p.R149Q results in stable apo-forms. The results for mutant p.R26H show that it is not structurally affected, thus it rather appears as a pure catalytic mutant. Furthermore, we have the mutants that are a mixture of both folding defects and reduced catalytic activity, these we have further split in two categories: those that have a disruption in elongation leading to accumulation of one intermediate i.e., p.D99N, and those that have a distribution of enzyme-intermediate complexes with reduced catalytic activity i.e., p.L30P, p.R167Q, p.R195C and p.E223K, which are variant that can be catalytically inactive with a different cause of action.

8   Future perspectives

Even though HMBS has been extensively studied in the last two decades some details of the elongation mechanism are still unknown, and there is still much to learn from the biochemical and biophysical characterisation of disease-associated variants. Elucidation of the structure of the ES_3 and ES_4 intermediates is therefore vital to understanding its function.

In this study we attempted to characterise 12 HMBS variants, however, purification for four of the variants were unsuccessful, further investigation and optimisation of the expression and purification protocol for these insoluble and unstable apo variants is needed.

Additionally, SEC coupled with multi-angle light scattering could be utilised to analyse the different populations of HMBS that elute during the final purification using SEC in order to confirm the identity of the first two peaks, as their identity as aggregates and dimer is just speculation at this point. It has been indicated that HMBS can occur as a dimer [21, 136], however one could speculate that is caused by misfolded monomeric protein. The domain movement that occurs when the active site is empty could potentially cause it to oligomerise [96]. This issue was not relevant to this project, and there is limited information on the enzyme regulation, however, it would be interesting to investigate in the future the identity of the species larger than the monomer, as well as their biological relevance.

Further research and optimisation using X-ray crystallography will be necessary to obtain crystal structures for HMBS in the ES_2, ES_3 and ES_4-state. Although no crystal structures were solved during this project, we found that p.R26H crystals in the ES_2-state diffracted and diffraction data collection will be pursued after further optimisation of the crystals. Furthermore, as there are no structures for HMBS in the ES_3 and ES_4-state, we will pursue this further as structures for the ES_3 and ES_4 intermediates will provide valuable insight into the mechanisms of the elongation process, and eventually provide snapshot-like details for each state of elongation.

9 References

1. Hamza, I. and H.A. Dailey, *One ring to rule them all: trafficking of heme and heme synthesis intermediates in the metazoans.* Biochim Biophys Acta, 2012. **1823**(9): p. 1617-32.
2. Bonkovsky, H.L., et al., *Porphyrin and Heme Metabolism and the Porphyrias*, in *Comprehensive Physiology.* 2013. p. 365-401.
3. Celis, A.I. and J.L. DuBois, *Making and breaking heme.* Curr Opin Struct Biol, 2019. **59**: p. 19-28.
4. Shimizu, T., et al., *Heme: emergent roles of heme in signal transduction, functional regulation and as catalytic centres.* Chemical Society Reviews, 2019. **48**(24): p. 5624-5657.
5. Puy, H., et al., *Molecular epidemiology and diagnosis of PBG deaminase gene defects in acute intermittent porphyria.* Am J Hum Genet, 1997. **60**(6): p. 1373-83.
6. Bustad, H.J., et al., *Acute Intermittent Porphyria: An Overview of Therapy Developments and Future Perspectives Focusing on Stabilisation of HMBS and Proteostasis Regulators.* International journal of molecular sciences, 2021. **22**(2): p. 675.
7. Granick, S., *Porphyrin biosynthesis in erythrocytes. I. Formation of gamma-aminolevulinic acid in erythrocytes.* J Biol Chem, 1958. **232**(2): p. 1101-17.
8. Beale, S.I., S.P. Gough, and S. Granick, *Biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley.* Proc Natl Acad Sci U S A, 1975. **72**(7): p. 2719-23.
9. Elder, G.H., C.H. Gray, and D.C. Nicholson, *The porphyrias: a review.* J Clin Pathol, 1972. **25**(12): p. 1013-33.
10. Battersby, A.R., *The discovery of nature's biosynthetic pathways.* Experientia, 1978. **34**(1): p. 1-13.
11. Bogorad, L. and S. Granick, *The Enzymatic Synthesis of Porphyrins from Porphobilinogen.* Proc Natl Acad Sci U S A, 1953. **39**(12): p. 1176-88.
12. Battersby, A.R., et al., *Biosynthesis of the pigments of life: formation of the macrocycle.* Nature, 1980. **285**(5759): p. 17-21.
13. Phillips, J.D., *Heme biosynthesis and the porphyrias.* Mol Genet Metab, 2019. **128**(3): p. 164-177.
14. Karim, Z., et al., *Porphyrias: A 2015 update.* Clin Res Hepatol Gastroenterol, 2015. **39**(4): p. 412-25.
15. Yasuda, M., B. Chen, and R.J. Desnick, *Recent advances on porphyria genetics: Inheritance, penetrance & molecular heterogeneity, including new modifying/causative genes.* Mol Genet Metab, 2019. **128**(3): p. 320-331.
16. Puy, H., L. Gouya, and J.C. Deybach, *Porphyrias.* Lancet, 2010. **375**(9718): p. 924-37.
17. Wang, B., *The acute hepatic porphyrias.* Transl Gastroenterol Hepatol, 2021. **6**: p. 24.
18. Meyer, U.A., et al., *Intermittent acute porphyria--demonstration of a genetic defect in porphobilinogen metabolism.* N Engl J Med, 1972. **286**(24): p. 1277-82.
19. Namba, H., et al., *Assignment of human porphobilinogen deaminase to 11q24.1----q24.2 by in situ hybridization and gene dosage studies.* Cytogenet Cell Genet, 1991. **57**(2-3): p. 105-8.
20. Gouya, L., et al., *Modulation of penetrance by the wild-type allele in dominantly inherited erythropoietic protoporphyria and acute hepatic porphyrias.* Hum Genet, 2004. **114**(3): p. 256-62.

21. Chen, B., et al., *Identification and characterization of 40 novel hydroxymethylbilane synthase mutations that cause acute intermittent porphyria.* J Inherit Metab Dis, 2019. **42**(1): p. 186-194.
22. Elder, G., et al., *The incidence of inherited porphyrias in Europe.* J Inherit Metab Dis, 2013. **36**(5): p. 849-57.
23. Chen, B., et al., *Acute Intermittent Porphyria: Predicted Pathogenicity of HMBS Variants Indicates Extremely Low Penetrance of the Autosomal Dominant Disease.* Human mutation, 2016. **37**(11): p. 1215-1222.
24. Lenglet, H., et al., *From a dominant to an oligogenic model of inheritance with environmental modifiers in acute intermittent porphyria.* Hum Mol Genet, 2018. **27**(7): p. 1164-1173.
25. Simon, A., et al., *Patient Perspective on Acute Intermittent Porphyria with Frequent Attacks: A Disease with Intermittent and Chronic Manifestations.* Patient, 2018. **11**(5): p. 527-537.
26. Gouya, L., et al., *EXPLORE: A Prospective, Multinational, Natural History Study of Patients with Acute Hepatic Porphyria with Recurrent Attacks.* Hepatology, 2020. **71**(5): p. 1546-1558.
27. Sassa, S., *Diagnosis and therapy of acute intermittent porphyria.* Blood Rev, 1996. **10**(1): p. 53-8.
28. Solis, C., et al., *Identification and expression of mutations in the hydroxymethylbilane synthase gene causing acute intermittent porphyria (AIP).* Mol Med, 1999. **5**(10): p. 664-71.
29. Thunell, S., *Porphyrins, porphyrin metabolism and porphyrias. I. Update.* Scand J Clin Lab Invest, 2000. **60**(7): p. 509-40.
30. Innala, E., et al., *Evaluation of gonadotropin-releasing hormone agonist treatment for prevention of menstrual-related attacks in acute porphyria.* Acta Obstet Gynecol Scand, 2010. **89**(1): p. 95-100.
31. Balwani, M., et al., *Acute Intermittent Porphyria in children: A case report and review of the literature.* Mol Genet Metab, 2016. **119**(4): p. 295-299.
32. Pulgar, V.M., et al., *Sex differences in vascular reactivity in mesenteric arteries from a mouse model of acute intermittent porphyria.* Mol Genet Metab, 2019. **128**(3): p. 376-381.
33. Meyer, U.A., M.M. Schuurmans, and R.L. Lindberg, *Acute porphyrias: pathogenesis of neurological manifestations.* Semin Liver Dis, 1998. **18**(1): p. 43-52.
34. Bonkovsky, H.L., N. Dixon, and S. Rudnick, *Pathogenesis and clinical features of the acute hepatic porphyrias (AHPs).* Mol Genet Metab, 2019. **128**(3): p. 213-218.
35. Yang, J., et al., *Clinical and Laboratory Features of Acute Porphyria: A Study of 36 Subjects in a Chinese Tertiary Referral Center.* Biomed Res Int, 2016. **2016**: p. 3927635.
36. Balwani, M., et al., *Acute hepatic porphyrias: Recommendations for evaluation and long-term management.* Hepatology, 2017. **66**(4): p. 1314-1322.
37. Gerischer, L.M., et al., *Acute porphyrias - A neurological perspective.* Brain Behav, 2021. **11**(11): p. e2389.
38. Oliveira Santos, M. and M. Leal Rato, *Neurology of the acute hepatic porphyrias.* J Neurol Sci, 2021. **428**: p. 117605.
39. Anderson, K.E., et al., *Biochemical Diagnosis of Acute Hepatic Porphyria: Updated Expert Recommendations for Primary Care Physicians.* Am J Med Sci, 2021. **362**(2): p. 113-121.

40. Pallet, N., et al., *High prevalence of and potential mechanisms for chronic kidney disease in patients with acute intermittent porphyria.* Kidney Int, 2015. **88**(2): p. 386-95.

41. Baravelli, C.M., et al., *Acute hepatic porphyria and cancer risk: a nationwide cohort study.* J Intern Med, 2017. **282**(3): p. 229-240.

42. Lissing, M., et al., *Risk of primary liver cancer in acute hepatic porphyria patients: A matched cohort study of 1,244 individuals.* J Intern Med, 2022.

43. Marcacci, M., et al., *Challenges in diagnosis and management of acute hepatic porphyrias: from an uncommon pediatric onset to innovative treatments and perspectives.* Orphanet J Rare Dis, 2022. **17**(1): p. 160.

44. Anderson, K.E., *Acute hepatic porphyrias: Current diagnosis & management.* Mol Genet Metab, 2019. **128**(3): p. 219-227.

45. Whatley, S.D. and M.N. Badminton, *Role of genetic testing in the management of patients with inherited porphyria and their families.* Ann Clin Biochem, 2013. **50**(Pt 3): p. 204-16.

46. Baumann, K. and R. Kauppinen, *Penetrance and predictive value of genetic screening in acute porphyria.* Mol Genet Metab, 2020. **130**(1): p. 87-99.

47. Bissell, D.M., K.E. Anderson, and H.L. Bonkovsky, *Porphyria.* N Engl J Med, 2017. **377**(21): p. 2101.

48. Marsden, J.T., et al., *Audit of the Use of Regular Haem Arginate Infusions in Patients with Acute Porphyria to Prevent Recurrent Symptoms.* JIMD Rep, 2015. **22**: p. 57-65.

49. Schmitt, C., et al., *Recurrent attacks of acute hepatic porphyria: major role of the chronic inflammatory response in the liver.* J Intern Med, 2018. **284**(1): p. 78-91.

50. Soonawalla, Z.F., et al., *Liver transplantation as a cure for acute intermittent porphyria.* Lancet, 2004. **363**(9410): p. 705-6.

51. Lissing, M., et al., *Liver Transplantation for Acute Intermittent Porphyria.* Liver Transpl, 2021. **27**(4): p. 491-501.

52. Balwani, M., et al., *Phase 3 Trial of RNAi Therapeutic Givosiran for Acute Intermittent Porphyria.* Zeitschrift Fur Gastroenterologie, 2020. **58**(08): p. 785-785.

53. Sardh, E., et al., *Safety, pharmacokinetics and pharmocodynamics of recombinant human porphobilinogen deaminase in healthy subjects and asymptomatic carriers of the acute intermittent porphyria gene who have increased porphyrin precursor excretion.* Clin Pharmacokinet, 2007. **46**(4): p. 335-49.

54. Yasuda, M., et al., *AAV8-mediated gene therapy prevents induced biochemical attacks of acute intermittent porphyria and improves neuromotor function.* Mol Ther, 2010. **18**(1): p. 17-22.

55. Unzu, C., et al., *Helper-dependent adenoviral liver gene therapy protects against induced attacks and corrects protein folding stress in acute intermittent porphyria mice.* Hum Mol Genet, 2013. **22**(14): p. 2929-40.

56. D'Avola, D., et al., *Phase I open label liver-directed gene therapy clinical trial for acute intermittent porphyria.* J Hepatol, 2016. **65**(4): p. 776-783.

57. Jiang, L., et al., *Systemic messenger RNA as an etiological treatment for acute intermittent porphyria.* Nat Med, 2018. **24**(12): p. 1899-1909.

58. Cordoba, K.M., et al., *Recombinant porphobilinogen deaminase targeted to the liver corrects enzymopenia in a mouse model of acute intermittent porphyria.* Sci Transl Med, 2022. **14**(627): p. eabc0700.

59. Grandchamp, B., et al., *Tissue-specific expression of porphobilinogen deaminase. Two isoenzymes from a single gene.* Eur J Biochem, 1987. **162**(1): p. 105-10.

60. Chretien, S., et al., *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, 1988. **85**(1): p. 6-10.

61. Raich, N., et al., *Molecular cloning and complete primary sequence of human erythrocyte porphobilinogen deaminase.* Nucleic Acids Res, 1986. **14**(15): p. 5955-68.

62. Song, G.J., et al., *Structural insight into acute intermittent porphyria.* Faseb Journal, 2009. **23**(2): p. 396-404.

63. Helliwell, J.R., *The crystal structures of the enzyme hydroxymethylbilane synthase, also known as porphobilinogen deaminase.* Acta Crystallogr F Struct Biol Commun, 2021. **77**(Pt 11): p. 388-398.

64. Brons-Poulsen, J., et al., *Characterization of two isoalleles and three mutations in both isoforms of purified recombinant human porphobilinogen deaminase.* Scandinavian Journal of Clinical & Laboratory Investigation, 2005. **65**(2): p. 93-105.

65. Pluta, P., et al., *Structural basis of pyrrole polymerization in human porphobilinogen deaminase.* Biochim Biophys Acta Gen Subj, 2018. **1862**(9): p. 1948-1955.

66. Jordan, P.M. and M.J. Warren, *Evidence for a dipyrromethane cofactor at the catalytic site of E. coli porphobilinogen deaminase.* FEBS Lett, 1987. **225**(1-2): p. 87-92.

67. Song, G., et al., *Structural insight into acute intermittent porphyria.* FASEB J, 2009. **23**(2): p. 396-404.

68. Miller, A.D., et al., *Evidence that the pyrromethane cofactor of hydroxymethylbilane synthase (porphobilinogen deaminase) is bound to the protein through the sulphur atom of cysteine-242.* Biochem J, 1988. **254**(3): p. 915-8.

69. Louie, G.V., et al., *Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site.* Nature, 1992. **359**(6390): p. 33-9.

70. Frydman, B., et al., *Biosynthesis of uroporphyrinogens from porphobilinogen: mechanism and the nature of the process.* Philos Trans R Soc Lond B Biol Sci, 1976. **273**(924): p. 137-60.

71. Jordan, P.M., *The biosynthesis of uroporphyrinogen III: mechanism of action of porphobilinogen deaminase.* Ciba Found Symp, 1994. **180**: p. 70-89; discussion 89-96.

72. Bustad, H.J., et al., *Characterization of porphobilinogen deaminase mutants reveals that arginine-173 is crucial for polypyrrole elongation mechanism.* iScience, 2021. **24**(3): p. 102152.

73. Lee, J.S., et al., *Genetic heterogeneity of the porphobilinogen deaminase gene in Swedish families with acute intermittent porphyria.* Hum Genet, 1991. **87**(4): p. 484-8.

74. Andersson, C., et al., *Diagnosis of acute intermittent porphyria in northern Sweden: an evaluation of mutation analysis and biochemical methods.* J Intern Med, 1995. **237**(3): p. 301-8.

75. Tollali, G., E.W. Nielsen, and O.L. Brekke, *[Acute intermittent porphyria].* Tidsskr Nor Laegeforen, 2002. **122**(11): p. 1102-5.

76. Cooper, D.N. and H. Youssoufian, *The CpG dinucleotide and human genetic disease.* Hum Genet, 1988. **78**(2): p. 151-5.

77. Floderus, Y., P.M. Shoolingin-Jordan, and P. Harper, *Acute intermittent porphyria in Sweden. Molecular, functional and clinical consequences of some new mutations found in the porphobilinogen deaminase gene.* Clin Genet, 2002. **62**(4): p. 288-97.

78. Schneider-Yin, X., et al., *Mutation hotspots in the human porphobilinogen deaminase gene: recurrent mutations G111R and R173Q occurring at CpG motifs.* J Inherit Metab Dis, 2004. **27**(5): p. 625-31.

79. Gill, R., et al., *Structure of human porphobilinogen deaminase at 2.8 A: the molecular basis of acute intermittent porphyria.* Biochem J, 2009. **420**(1): p. 17-25.

80. Llewellyn, D.H., S. Whatley, and G.H. Elder, *Acute Intermittent Porphyria Caused by an Argine to Histidine Substitution (R26h) in the Cofactor-Binding Cleft of Porphobilinogen Deaminase.* Human Molecular Genetics, 1993. **2**(8): p. 1315-1316.

81. Ong, P.M., et al., *Acute intermittent porphyria: the in vitro expression of mutant hydroxymethylbilane synthase.* Mol Cell Probes, 1997. **11**(4): p. 293-6.

82. Younger, D.S. and K. Tanji, *Demyelinating neuropathy in genetically confirmed acute intermittent porphyria.* Muscle Nerve, 2015. **52**(5): p. 916-7.

83. Kauppinen, R., et al., *Acute intermittent porphyria in Finland: 19 mutations in the porphobilinogen deaminase gene.* Hum Mol Genet, 1995. **4**(2): p. 215-22.

84. De Siervi, A., et al., *Acute intermittent porphyria: characterization of two novel mutations in the porphobilinogen deaminase gene, one amino acid deletion (453-455delAGC) and one splicing aceptor site mutation (IVS8-1G>T).* Hum Mutat, 1999. **14**(4): p. 355.

85. Hrdinka, M., H. Puy, and P. Martasek, *May 2006 update in porphobilinogen deaminase gene polymorphisms and mutations causing acute intermittent porphyria: comparison with the situation in Slavic population.* Physiol Res, 2006. **55 Suppl 2**: p. S119-36.

86. To-Figueras, J., et al., *Genetic and biochemical characterization of 16 acute intermittent porphyria cases with a high prevalence of the R173W mutation.* J Inherit Metab Dis, 2006. **29**(4): p. 580-5.

87. Jordan, P.M. and S.C. Woodcock, *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, 1991. **280 (Pt 2)**: p. 445-9.

88. Bung, N., et al., *Human hydroxymethylbilane synthase: Molecular dynamics of the pyrrole chain elongation identifies step-specific residues that cause AIP.* Proceedings of the National Academy of Sciences, 2018. **115**(17): p. E4071-E4080.

89. Bung, N., et al., *Computational modeling of the catalytic mechanism of hydroxymethylbilane synthase.* Phys Chem Chem Phys, 2019. **21**(15): p. 7932-7940.

90. Whatley, S.D., et al., *Diagnostic strategies for autosomal dominant acute porphyrias: retrospective analysis of 467 unrelated patients referred for mutational analysis of the HMBS, CPOX, or PPOX gene.* Clin Chem, 2009. **55**(7): p. 1406-14.

91. Ulbrichova, D., et al., *Gene symbol: HMBS. Disease: Porphyria, acute intermittent.* Hum Genet, 2008. **124**(3): p. 315.

92. Kim, M.K. and Y.K. Kang, *Positional preference of proline in alpha-helices.* Protein Sci, 1999. **8**(7): p. 1492-9.

93. Martinez di Montemuros, F., et al., *Acute intermittent porphyria: heterogeneity of mutations in the hydroxymethylbilane synthase gene in Italy.* Blood Cells Mol Dis, 2001. **27**(6): p. 961-70.

94. Fu, Y., et al., *Systematically Analyzing the Pathogenic Variations for Acute Intermittent Porphyria.* Front Pharmacol, 2019. **10**: p. 1018.

95. Shoolingin-Jordan, P.M., et al., *Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 731-5.

96.   Guo, J., et al., *Structural studies of domain movement in active-site mutants of porphobilinogen deaminase from Bacillus megaterium.* Acta Crystallogr F Struct Biol Commun, 2017. **73**(Pt 11): p. 612-620.

97.   Brownlie, P.D., et al., *The three-dimensional structures of mutants of porphobilinogen deaminase: toward an understanding of the structural basis of acute intermittent porphyria.* Protein Sci, 1994. **3**(10): p. 1644-50.

98.   Delfau, M.H., et al., *Molecular heterogeneity of acute intermittent porphyria: identification of four additional mutations resulting in the CRIM-negative subtype of the disease.* Am J Hum Genet, 1991. **49**(2): p. 421-8.

99.   Gu, X.F., et al., *Detection of eleven mutations causing acute intermittent porphyria using denaturing gradient gel electrophoresis.* Hum Genet, 1994. **93**(1): p. 47-52.

100.   Yang, C.C., et al., *HMBS mutations in Chinese patients with acute intermittent porphyria.* Ann Hum Genet, 2008. **72**(Pt 5): p. 683-6.

101.   Lander, M., et al., *Studies on the mechanism of hydroxymethylbilane synthase concerning the role of arginine residues in substrate binding.* Biochem J, 1991. **275 (Pt 2)**: p. 447-52.

102.   Bung, N., et al., *Human hydroxymethylbilane synthase: Molecular dynamics of the pyrrole chain elongation identifies step-specific residues that cause AIP.* Proc Natl Acad Sci U S A, 2018. **115**(17): p. E4071-E4080.

103.   Delfau, M.H., et al., *Two different point G to A mutations in exon 10 of the porphobilinogen deaminase gene are responsible for acute intermittent porphyria.* J Clin Invest, 1990. **86**(5): p. 1511-6.

104.   Llewellyn, D.H., et al., *Homozygous acute intermittent porphyria: compound heterozygosity for adjacent base transitions in the same codon of the porphobilinogen deaminase gene.* Hum Genet, 1992. **89**(1): p. 97-8.

105.   Chen, C.H., et al., *Acute intermittent porphyria: identification and expression of exonic mutations in the hydroxymethylbilane synthase gene. An initiation codon missense mutation in the housekeeping transcript causes "variant acute intermittent porphyria" with normal expression of the erythroid-specific enzyme.* J Clin Invest, 1994. **94**(5): p. 1927-37.

106.   Andersson, C., et al., *The W198X and R173W mutations in the porphobilinogen deaminase gene in acute intermittent porphyria have higher clinical penetrance than R167W. A population-based study.* Scand J Clin Lab Invest, 2000. **60**(7): p. 643-8.

107.   Solis, C., et al., *Acute intermittent porphyria: studies of the severe homozygous dominant disease provides insights into the neurologic attacks in acute porphyrias.* Arch Neurol, 2004. **61**(11): p. 1764-70.

108.   Bustad, H.J., et al., *Conformational stability and activity analysis of two hydroxymethylbilane synthase mutants, K132N and V215E, with different phenotypic association with acute intermittent porphyria.* Biosci Rep, 2013. **33**(4).

109.   Ulbrichova, D., et al., *Correlation between biochemical findings, structural and enzymatic abnormalities in mutated HMBS identified in six Israeli families with acute intermittent porphyria.* Blood Cells Mol Dis, 2009. **42**(2): p. 167-73.

110.   Fortgens, P., et al., *Molecular characterisation of acute intermittent porphyria in a cohort of South African patients and kinetic analysis of two expressed mutants.* J Clin Pathol, 2017. **70**(6): p. 515-520.

111.   Tjensvoll, K., et al., *Haplotype analysis of Norwegian and Swedish patients with acute intermittent porphyria (AIP): Extreme haplotype heterogeneity for the mutation R116W.* Dis Markers, 2003. **19**(1): p. 41-6.

112.   Yarra, P., et al., *Benefits of prophylactic heme therapy in severe acute intermittent porphyria.* Mol Genet Metab Rep, 2019. **19**: p. 100450.

113.	Lundin, G., et al., *Four mutations in the porphobilinogen deaminase gene in patients with acute intermittent porphyria.* J Med Genet, 1995. **32**(12): p. 979-81.

114.	Nissen, H., et al., *Diagnostic strategy, genetic diagnosis and identification of new mutations in intermittent porphyria by denaturing gradient gel electrophoresis.* Hum Mutat, 1997. **9**(2): p. 122-30.

115.	Xian, F., C.L. Hendrickson, and A.G. Marshall, *High resolution mass spectrometry.* Anal Chem, 2012. **84**(2): p. 708-19.

116.	Ho, C.S., et al., *Electrospray ionisation mass spectrometry: principles and clinical applications.* Clin Biochem Rev, 2003. **24**(1): p. 3-12.

117.	Svasti, J. and B. Panijpan, *SDS-polyacrylamide gel electrophoresis. A simple explanation of why it works.* Journal of Chemical Education, 1977. **54**(9): p. 560.

118.	Smith, B.J., *SDS Polyacrylamide Gel Electrophoresis of Proteins.* Methods Mol Biol, 1984. **1**: p. 41-55.

119.	Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nat Protoc, 2006. **1**(6): p. 2876-90.

120.	Micsonai, A., et al., *Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy.* Proc Natl Acad Sci U S A, 2015. **112**(24): p. E3095-103.

121.	Miles, A.J., R.W. Janes, and B.A. Wallace, *Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review.* Chem Soc Rev, 2021. **50**(15): p. 8400-8413.

122.	Vedadi, M., et al., *Biophysical characterization of recombinant proteins: a key to higher structural genomics success.* J Struct Biol, 2010. **172**(1): p. 107-19.

123.	Martin-Malpartida, P., et al., *HTSDSF Explorer, A Novel Tool to Analyze High-throughput DSF Screenings.* Journal of Molecular Biology, 2021: p. 167372.

124.	Gao, K., R. Oerlemans, and M.R. Groves, *Theory and applications of differential scanning fluorimetry in early-stage drug discovery.* Biophys Rev, 2020. **12**(1): p. 85-104.

125.	Chayen, N.E. and E. Saridakis, *Protein crystallization: from purified protein to diffraction-quality crystal.* Nat Methods, 2008. **5**(2): p. 147-53.

126.	McPherson, A. and J.A. Gavira, *Introduction to protein crystallization.* Acta Crystallogr F Struct Biol Commun, 2014. **70**(Pt 1): p. 2-20.

127.	Støve SI, F.M., Hausvik E, Underhaug J, and Martinez A, *Differential scanning fluorimetry in the screening and validation of pharmacological chaperones for soluble and membrane proteins, Chapter 15.* 2020: Walthum: Academic Press.

128.	Sato, H., et al., *Crystal structures of hydroxymethylbilane synthase complexed with a substrate analog: a single substrate-binding site for four consecutive condensation steps.* Biochemical Journal, 2021. **478**(5): p. 1023-1042.

129.	Awan, S.J., et al., *Reconstitution of the holoenzyme form of Escherichia coli porphobilinogen deaminase from apoenzyme with porphobilinogen and preuroporphyrinogen: a study using circular dichroism spectroscopy.* Biochemistry, 1997. **36**(30): p. 9273-82.

130.	Jarzab, A., et al., *Meltome atlas-thermal proteome stability across the tree of life.* Nat Methods, 2020. **17**(5): p. 495-503.

131.	Shoolingin-Jordan, P.M., M.J. Warren, and S.J. Awan, *Discovery that the assembly of the dipyrromethane cofactor of porphobilinogen deaminase holoenzyme proceeds initially by the reaction of preuroporphyrinogen with the apoenzyme.* Biochem J, 1996. **316 (Pt 2)**: p. 373-6.

132.	Li, N., et al., *Functional studies of rat hydroxymethylbilane synthase.* Bioorg Chem, 2008. **36**(5): p. 241-51.

133. Ulbrichova, D., et al., *Acute intermittent porphyria--impact of mutations found in the hydroxymethylbilane synthase gene on biochemical and enzymatic protein properties.* FEBS J, 2009. **276**(7): p. 2106-15.

134. Poli, A., et al., *Givosiran in acute intermittent porphyria: A personalized medicine approach.* Mol Genet Metab, 2022.

135. Ricci, A. and P. Ventura, *Givosiran for the treatment of acute hepatic porphyria.* Expert Rev Clin Pharmacol, 2022: p. 1-11.

136. Medlock, A.E., et al., *Identification of the Mitochondrial Heme Metabolism Complex.* PLoS One, 2015. **10**(8): p. e0135896.

10 Appendix I

Manuscript submitted to FEBS OpenBio based on results obtained during this master's project

Characterisation of a common hotspot mutation in acute intermittent porphyria sheds light on hydroxymethylbilane synthase mechanism

| Abstract: | Hydroxymethylbilane synthase (HMBS), the third enzyme in the haem biosynthesis, catalyses the formation of tetrapyrrole 1-hydroxymethylbilane (HMB) by binding four consecutive substrate molecules, creating the enzyme intermediate complexes ES, ES_2, ES_3 and ES_4. Mutations in HMBS are associated with the dominantly inherited disorder acute intermittent porphyria. In this study, we have characterised the p.R26H mutant to shed light on the role of Arg26 in the elongation mechanism of HMBS and to provide insight into its effect on the enzyme. With selected biophysical methods, we have been able to show that p.R26H forms a single enzyme intermediate complex in the ES_2-state. We were also able to demonstrate that p.R26H causes an inactive enzyme, unable to produce the HMB product. |

1   **Characterisation of a common hotspot mutation in acute intermittent porphyria sheds**

2   **light on hydroxymethylbilane synthase mechanism**

3   Marthe Christie Sæter[1], Mikko Laitaoja[2], Aasne K. Aarsand[3,4], Juha P. Kallio[1], and

4   Helene J. Bustad[3,†]

5   [1]Department of Biomedicine, University of Bergen, 5020 Bergen, Norway
6   [2]Department of Chemistry, University of Eastern Finland, 80130 Joensuu, Finland
7   [3]Norwegian Porphyria Centre (NAPOS), Department for Medical Biochemistry and
8   Pharmacology, Haukeland University Hospital, 5021 Bergen, Norway
9   [4]Norwegian Organization for Quality Improvement of Laboratory Examinations, Haraldsplass
10 Deaconess Hospital, 5009 Bergen, Norway

11 [†]Corresponding author: Helene J. Bustad, Norwegian Porphyria Centre (NAPOS),
12 Department for Medical Biochemistry and Pharmacology, Haukeland University Hospital,
13 5021 Bergen, Norway, Helene.bustad.johannessen@helse-bergen.no

1   Abstract

15 Hydroxymethylbilane synthase (HMBS) is the third enzyme in the haem biosynthesis, where

16 it catalyses the formation of tetrapyrrole 1-hydroxymethylbilane (HMB). In this process,

17 HMBS binds four consecutive substrate molecules, creating the enzyme intermediate

18 complexes ES, ES_2, ES_3 and ES_4. Mutations in the *HMBS* gene are associated with the

19 dominantly inherited disorder acute intermittent porphyria (AIP). In this study, we have

20 characterised the p.R26H mutant to shed light on the role of Arg26 in the elongation

21 mechanism of HMBS and to provide insight into its effect on the enzyme. With selected

22 biophysical methods, we have been able to show that p.R26H forms a single enzyme

23 intermediate complex in the ES_2-state. We were also able to demonstrate that p.R26H causes

24 an inactive enzyme, unable to produce the HMB product.

25 Keywords:

26 Acute intermittent porphyria, hydroxymethylbilane synthase, porphobilinogen deaminase,

27 haem biosynthesis, recombinant proteins, protein mutation, enzyme-intermediates, Fourier

28 transform ion cyclotron resonance mass spectrometry

29 Abbreviations:

30 AIP, Acute intermittent porphyria; ALA, δ-aminolaevulinic acid; CD, circular dichroism;

31 DSF, differential scanning fluorimetry; FT-ICR MS, Fourier transform ion cyclotron

32 resonance mass spectrometry; HMB, 1-hydroxymethylbilane; HMBS, hydroxymethylbilane

33 synthase; MD, molecular dynamics; PBG, porphobilinogen

2   Introduction

Hydroxymethylbilane synthase (HMBS or porphobilinogen deaminase (PBGD)) is the third enzyme in the haem biosynthetic pathway, where it catalyses the formation of the linear tetrapyrrole 1-hydroxymethylbilane (HMB) [1]. In this process, four consecutive PBG molecules are bound, creating the enzyme intermediate complexes ES, ES_2, ES_3, and ES_4 respectively, as reviewed in Phillips *et al.* [2]. In humans, there are two tissue-specific isoforms of HMBS derived from alternative splicing: the housekeeping (HMBS1) and the erythroid-specific (HMBS2) enzyme, with 361 and 344 amino acids, respectively. The 17 extra amino acids at the N-terminus of HMBS1 have no known function, however, it has been suggested they have a regulatory function or a role in e.g., protein trafficking {Brons-Poulsen, 2005 #90;Pluta, 2018 #59}. Mutations in the *HMBS* gene are associated with the low penetrant autosomal dominant disorder acute intermittent porphyria (AIP). Both HMBS isoforms are affected if the mutation appears in the shared part of *HMBS* because of alternative splicing [5]. In *HMBS* mutation carriers, wild-type (wt) enzyme expresses only from one allele and the loss in total HMBS enzymatic activity can therefore be up to 50 % [6]. However, most *HMBS* carriers will never develop symptomatic disease, which is characterised by potentially life-threatening neurovisceral acute attacks, as the associated loss of HMBS activity is not sufficient for clinical expression, and the prevalence of likely pathogenic *HMBS* variants observed in the general population greatly outnumbers symptomatic disease prevalence [7]. In symptomatic AIP patients, clinical disease is associated with situations that increase the demand for haem, leading to an upregulation of δ-aminolaevulinic acid (ALA) synthase 1, the first enzyme in the haem biosynthesis, whereupon the affected HMBS enzyme becomes rate-limiting [2]. In this setting, porphyrin precursors upstream from HMBS, ALA, and porphobilinogen (PBG), are produced in excess in the liver [8]. The overproduction and consequently accumulation of neurotoxic precursors is considered to be the cause of acute attacks, and increased levels of ALA and PBG in urine are used as diagnostic markers of AIP, and other acute porphyrias in which ALA and PBG accumulate. AIP is associated with a varying natural history, from never-symptomatic carriers to patients with recurrent severe acute attacks. There is also a high phenotypic heterogeneity, even within the same family. Clinically, acute attacks are characterised by automatic dysfunction including severe abdominal pain, tachycardia, hypertension, and a wide range of neurological and psychiatric symptoms [8-10]. Acute attacks are triggered by various factors, most importantly drug-induced hepatic haem depletion via induction or suicidal inactivation of cytochrome P450 enzymes (CYPs) [11], hormonal changes associated with the menstrual

68 cycle [12], low-calorie intake [13, 14], infection and stress [15, 16], though often a clear

69 triggering factor cannot be established. Additionally, *HMBS* mutation carriers are at risk of

70 long-term complications including primary liver cancer, hypertension, and kidney failure [17-

71 20]. Sporadic acute attacks are treated by supportive measures and haem infusions, where the

72 latter by negative feedback downregulates the hepatic ALAS1 activity. Currently, the only

73 cure for AIP is liver transplantation [21], however, a novel and recently introduced RNA

74 interference therapy, Givosiran, has been shown to be effective in reducing attacks in patients

75 with recurrent acute attacks [22, 23]. So, further research is still needed to explore alternative

76 and more efficient treatment options.

77   More than 500 different HMBS mutations, associated with clinical disease, have been

78 published [22]. However, for many variants, the effect on the HMBS enzyme has not been

79 clearly described and the knowledge of the HMBS mechanism at the molecular level is still

80 limited. Even with the numerous HMBS crystal structures and molecular dynamic (MD)

81 simulations published in the last decades, the exact mechanisms of polypyrrole elongation are

82 not yet completed through to the stage of Michaelis complex EP [24]. From the earliest

83 studies it was recognised that the enzyme in the crystal had to be with the cofactor in its active

84 form [25]. Arg26 has been suggested to have a crucial role in the elongation process. Its

85 location in hydrogen-bonding distance to the proposed substrate entry site (yellow, Figure

86 1A), close to the available substrate and "growing end" of the intermediate complex where it

87 can interact with the acetate of PBG, indicates a role in both positioning and possibly

88 deaminating the incoming substrates [26, 27] The substitution of arginine 26 to histidine

89 (Figure 1B) results in a disease-associated mutant (c.77G>A; p.R26H), which has been

90 described in several different AIP families [28, 29]. Two other missense mutations at the

91 same location, p.R26C and p.R26L, have also been reported [30-37]. In this work, we have

92 sought to characterise the p.R26H mutant to shed light on Arg26 and its role in the elongation

93 mechanism of HMBS. Using a combination of anion-exchange chromatography, native

94 PAGE, circular dichroism, differential scanning fluorimetry, high-resolution mass

95 spectrometry, we demonstrate how the mutant is trapped in the ES_2-state, and that it is

96 inactive without product turnover. Characterising HMBS mutations in depth can provide an

97 insight into their effect on the enzyme's structure and function, thereby providing more

98 detailed information on the HMBS mechanism, allowing for a better understanding of

99 genotype-phenotype correlations, and potentially providing useful information for future

100 enzyme specific treatment options.

3   Materials and Methods

Human wt-HMBS and mutant were expressed in *Escherichia coli* BL21(DE3) strains as a

His-tagged fusion proteins with a tobacco etched virus (TEV) protease cleavage site

(NLYFQ/G) in a pET-28a(+)-TEV vector. In this study the erythroid-specific (HMBS2)

enzyme was used but the numbering was kept corresponding to the housekeeping isoform

(HMBS1) to match the reports better from previous studies where p.R9H (c.26G>A) in the

presented sequence equals to p.R26H (c.77G>A). HMBS expression constructs

were purchased from GenScript Biotech (Piscataway, New Jersey, USA), with the following

sequence for the mutant:

```
(G)MRVIRVGTHKSQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSKIGEKSLF
TKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPHDAVVFHPKFVGKTLETLPEK
SVVGTSSLRRAAQLQRKFPHLEFRSIRGNLNTRLRKLDEQQEFSAIILATAGLQRMGWHNRVG
QILHPEECMYAVGQGALGVEVRAKDQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVA
VHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITARNIPRGPQL
AAQNLGISLANLLLSKGAKNILDVARQLNDAH
```

Theoretical masses: $HMBS_{apo}$: 37735.94 Da; ES_2 38573.22 Da; ES_3: 38782.29 Da

Protein expression and purification

Wt-HMBS and mutant p.R26H were expressed in *E. coli* BL21 (DE3). 50 mL LB broth

(EZmix™) starter cultures supplemented with 50 µg/mL kanamycin were grown overnight in

a shaking incubator at 200 RPM and 30 °C. The starter cultures were diluted in 1 L LB broth

supplemented with kanamycin and incubated further at 200 RPM and 37 °C until OD_{600} of

0.5–0.6 before induction with 0.5 mM IPTG and further incubated overnight at 200 RPM and

27 °C. Cells were harvested at 6000 × g for 30 minutes and pellets were resuspended in

15 mL wash buffer (400 mM NaCl, 25 mM Tris at pH 8, 20 mM imidazole at pH 7.5, 0.5 mM

tris (2-carboxyethyl) phosphine (TCEP) added protease inhibitors (1 tablet cOmplete™

(Roche) w/EDTA, 10 mM benzamidine and 0.2 mM PMSF). Resuspended pellets were frozen

at −20 °C until use.

The resuspended pellet was thawed on ice and lysis buffer was added to 40 mL before

sonication on ice for 3 × 2 minute at 25 W output with 10 sec pulses and pauses on ice, then

centrifuged at 20 000 × g for 45 minutes. The supernatant was loaded onto a nickel-charged

affinity resin (Ni-NTA Agarose, Qiagen) equilibrated in wash buffer, using a gravity flow

purification column for His-tag affinity chromatography. The resin was washed 3 × CV with

wash buffer before eluting the protein with 400 mM NaCl, 25 mM Tris at pH 8, 400 mM

imidazole at pH 7.5 and 0.5 mM TCEP.

Eluted His-tagged HMBS was added 1:100 TEV protease to cleave off the fusion tag,

during dialysis with 400 mM NaCl, 25 mM Tris pH 8 and 0.5 mM TCEP in a cellulose

137   membrane dialysis tubing with MW cut-off at 14 000 ON at 4 °C. Cleaved protein was run

138   through Ni-NTA pre-equilibrated with 150 mM NaCl, 25 mM Tris pH 8 and 0.5 mM TCEP,

139   collected and concentrated using Amicon Ultra centrifugal 30 kDa cut-off filters (Millipore)

140   at 4000 × g. NanoDrop Spectrophotometer (Thermo Fisher Scientific) and sequence-derived

141   extinction coefficients were used to determine protein concentrations.

142   Concentrated wt-HMBS and p.R26H were further purified by size exclusion

143   chromatography (SEC) on a GE HiLoad Superdex 75 16/60 PG column connected to an Äkta

144   Pure Protein Purification System (Cytiva Europe GmbH, Freiburg, Germany) at 4 °C, using

145   30 mM NaCl, 25 mM Tris pH 8 (buffer A) at flow rate 1.000 ml/min.

146   Anion-exchange chromatography

147   Anion-exchange chromatography using a Mono Q 4.6/100 column (Cytiva) connected to an

148   Äkta Pure (Cytiva) was used to separate the enzyme intermediate complexes. The Mono Q

149   column was equilibrated in buffer A at 4 °C. The protein was eluted in a gradient from 0 % to

150   65 % using 400 mM NaCl, 25 mM Tris pH 8 (buffer B) at 4 °C and peak fractions were

151   collected. The isolated intermediates were concentrated using Amicon Ultra centrifugal

152   30 kDa cut-off filters (Millipore) for further analyses.

153   Polyacrylamide Gel Electrophoresis

154   Proteins were analysed by native polyacrylamide gel electrophoresis (native-PAGE) with a

155   loading buffer of 125 mM Tris-HCl, 40 % v/v glycerol, 0.002 % bromophenol blue. The

156   samples were loaded on a 10 % Mini-PROTEAN® TGX^{TM} Precast Protein Gel (Bio-Rad

157   Laboratories, USA) and ran for 2 ½ hrs at 140 V and 4 °C. The gel was stained with

158   Coomassie blue staining and visualised by Molecular Imager ChemiDoc XRS+ with

159   ImageLab software (Bio-Rad).

160   Circular Dichroism spectroscopy

161   Far-UV spectra were obtained using a J-810 Jasco spectropolarimeter with a CDF-426S

162   Peltier element for temperature control and a 300 µL quartz cuvette with a path length of

163   1 mm. Wt-HMBS and p.R26H mutant at 5 µM were prepared in 10 mM K_2HPO_4, 100 mM

164   NaF. Three scans were obtained for each spectrum and the buffer scans were subtracted using

165   the integrated Jasco software. BeStSel web server was used for secondary structure prediction

166   [38], and statistics analysed using unpaired t-test (GraphPad Prism version 9.2.0 for MacOS,

167   GraphPad Software, San Diego, California USA, www.graphpad.com).

168   Enzymatic activity

169 The enzymatic activity was assayed by adding 5 μg enzyme to 50 mM Na-Hepes at pH 8.2

170 and pre-incubating the sample for 3 minutes at 37 °C. To start the enzyme reaction 100 μM

171 PBG was added, and the reaction was stopped by adding 5 M HCl and 0.1 % benzoquinone

172 after 4 minutes. Samples were incubated on ice for 30 min before the absorbance was

173 measured at 405 nm by a NanoDrop Spectrophotometer. Michaelis-Menten kinetics was used

174 to determine the kinetic parameters; K_M and V_{max} was obtained using a range of PBG

175 concentrations from 3.125–1000 μM and analysed using the Michaelis-Menten enzyme

176 kinetics model in GraphPad Prism.

177   Mass spectrometry

178 Prior to mass spectrometric experiments, the enzyme samples were buffer-exchanged to

179 20 mM NH4OAc pH 6.8 using PD Miditrap G-25 column (Cytiva). Their concentrations were

180 determined by UV absorbance at A_{280} using their sequence-derived extinction coefficient

181 ($15470 M^{-1}cm^{-1}$). Further dilutions were made using HPLC quality acetonitrile, water, and

182 acetic acid. Mass spectra were measured using a Bruker 12-T solariX XR FT-ICR mass

183 spectrometer (Bruker Daltonik GmbH, Bremen, Germany). All samples were measured by

184 direct infusion, at 2 μL/min, using standard electrospray source. The instrument was

185 calibrated using NaPFHA clusters before measurements. The mass spectrometer was operated

186 using ftmsControl 2.2 software, and the data were processed using the Bruker

187 DataAnalysis 5.1 software. Deconvolution (i.e., zero-charge) spectra were calculated using a

188 Maximum Entropy (MaxEnt) deconvolution algorithm. All masses are reported as most

189 abundant isotopic masses.

190   Differential Scanning Fluorimetry

191 A Light Cycler 480 Real-Time PCR (Roche) was used to obtain thermal denaturing

192 scans. 5 μM protein was prepared in a phosphate buffer with 5× SYPRO Orange (Agilent

193 technologies, USA) and analysed in 384-well plates, as previously described [39]. Unfolding

194 curves were recorded at a scan rate of 2 °C/min with 0.2 °C intervals from 20 °C to 99 °C.

195 Data were analysed using HTSDSF Explorer [40].

196   4   Results

197 In this work, we have utilised the recombinant HMBS wt and missense mutant p.R9H, both in

198 the human HMBS2 isoform (NM_001024382.2:c.26G>A), which corresponds to p.R26H in

199 the HMBS1 isoform (NM_000190.4:c.77G>A). There are minimal structural and no known

200 functional differences between these two isoforms [4]. To keep consistency with the

201 literature, we will refer to the constructs in this work as wt-HMBS and the mutant as p.R26H.

202 The recombinant p.R26H was successfully purified, in amounts similar to the wt-

203 HMBS. To confirm whether the mutant was properly folded, circular dichroism (CD)

204 spectroscopy was performed prior to further experiments. A comparison of wt-HMBS and

205 p.R26H (Figure 1) indicates correct folding of the mutant, with a 4.7% decrease in α helix

206 content and 5.4% increase in β sheet parallels (Table 1), however, not significantly different

207 from wt-HMBS.

208 Purified HMBS can be further separated into its enzyme intermediates' complexes,

209 using anion-exchange chromatography with buffer gradient [4, 39, 41]. Comparing results for

210 wt-HMBS and p.R26H, it is evident that p.R26H mainly elutes as a single intermediate at

211 ~26% of gradient buffer B (Figure 2B), corresponding to ES_2 in wt-HMBS. Native PAGE has

212 commonly been used to demonstrate the enzyme-intermediate complexes, because of the

213 difference in net charge between them [4, 39, 42]. As seen in Figure 2B, lane 1, all the

214 expected bands from wt-HMBS were observed. On the other hand, p.R26H migrated as one

215 major band (Figure 2B, lane 2), with minor bands corresponding to the smaller peaks in the

216 chromatogram (Figure 2A). The isolated main peak of p.R26H is shown in Figure 2B, lane 3.

217 The migration pattern of p.R26H was inconclusive, as the major peak migrated further than

218 the ES_2 of wt-HMBS. The isolated intermediate complexes from both wt- and mutant HMBS

219 were analysed using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR

220 MS), as previously reported [43]. As expected, we determined that the peaks from anion-

221 exchange chromatography of wt-HMBS corresponded with E_{holo}, ES, ES_2 and ES_3 (data not

222 shown). Furthermore, FT-ICR MS measurements in denaturing conditions verified the

223 findings for p.R26H, being mainly present in the intermediate ES_2-state (Figure 1C). The

224 tailing of the ES_2 peak is most likely caused by incomplete desalting. Uroporphyrinogen was

225 also observed (Figure S1A), indicating that this cyclic product has affinity for and can be co-

226 purified with the enzyme [44]. Uroporphyrinogen I or III can be derived from either

227 recombinant HMBS or from endogenous *E. coli* HMBS enzyme activity, respectively. The

228 native-like conditions confirmed that the mutation did not cause any unfolding of the enzyme

229 (Figure S1B).

230 The steady-state kinetic characterisation was done in conditions as reported in [39].

231 Wt-HMBS demonstrated a specific activity of 757.70±67.67 $nmol \times mg^{-1} \times h^{-1}$, however,

232 p.R26H was completely inactive and incapable of producing the HMB product. Hence, the

233 kinetic parameters of the mutant could not be determined because it lacked enzymatic activity

234   (Figure S2). Addition of $10\times$ PBG (50 µM to 5 µM enzyme) did not change the distribution as

235   seen by FT-ICR MS (Figure S1C), further corroborating that the p.R26H is completely

236   inactive without turnover.

237   We have previously demonstrated that the thermostability of HMBS decreases with

238   the number of bound PBG substrates [39]. We wanted to investigate whether the mutation had

239   an impact on the structural stability, or if the mutational effect could solely be attributed to the

240   change in amino acid. Therefore, we analysed the thermostability of the isolated wt and

241   mutant enzyme-intermediate complexes, using differential scanning fluorimetry (DSF). The

242   wt-HMBS intermediates demonstrated a loss in the half-denaturing temperature (T_m)

243   represented as maxima of the first derivative (dF/dT) from 78.62 ± 0.07 °C to 71.02 ± 0.15 °C

244   for E_{holo} and ES_3, respectively (Figure 1D). The Tm-value of the isolated intermediate in

245   p.R26H mutant was 72.56 ± 0.10 °C, corresponding to the ES_2 of wt-HMBS, which was

246   72.78 ± 0.10 °C.

5   Discussion

248   Arg26 is one of many conserved arginine residues across species and has been implicated to

249   be critical for HMBS catalysis [45, 46]. In crystal structures PDB IDs 1PDA and 3ECR,

250   Arg26 has been observed in hydrogen-bonding distance to an acetate or sulphate ion (derived

251   from solutions used in the crystallisation process) located in the predicted substrate-binding

252   site [26, 47]. Other crystal structures such as PDB IDs 5M6R and 7AAK also indicate several

253   intradomain interactions for Arg26, such as Pi-cation interaction with Phe77, hydrophobic

254   alkyl interaction with Ala31 and Leu81; hydrogen-bond to Thr25 and electrostatic interactions

255   with Asp99 [4, 43], indicating its important location in the active site, and putative role in

256   both enzyme activity and structural stability. Sato *et al.* demonstrated that Arg26, Arg167 and

257   Arg173 all interact with the substrate analogue positioned in the substrate-binding site [41].

258   Interestingly, previously described mutants R167W and R173W, both interacting with the

259   propionate side chain of PBG, exhibit different effects, where the first can form all enzyme

260   intermediate complexes, whereas the latter creates a catalytic blockage and a subsequent

261   accumulation of a single intermediate [43]. Our results demonstrate that p.R26H also results

262   in a mutant that is catalytically blocked with ES_2 as the accumulated intermediate, however,

263   no major structural changes are indicated based on secondary structure content and

264   thermostability analysis.

265   Arg26 in humans corresponds to Arg11 in *E. coli* HMBS. The p.R11H- and p.R11L

266   mutants expressed in *E. coli* were reported as correctly folded but unable to bind

267   substrate [46, 48]. Interestingly, two previous reports on the human p.R26H mutant have

268   incongruent observations, with one suggesting altered interaction with the cofactor and the

269   other no pyrrole chain elongation [35, 49]. However, we do not have direct evidence of the

270   formation of ES_2 from ES and cannot rule out that endogenous HMBS produces HMB

271   product, which is subsequently inserted in apo-mutant without involving Arg26 proton

272   donation.

273   Recent advances in the field include the crystal structure of wt-HMBS in its ES_2-state

274   (PDB ID: 5M6R), demonstrating how the cofactor-binding loop with the Cys261 residue that

275   anchors the cofactor, is pulled backwards with the two incoming PBG substrates [4]. The two

276   new substrate molecules replace the cofactor in their positions and interactions [4, 43].

277   Furthermore, our own report on the p.R173W-mutant trapped in the ES_2-state (PDB ID:

278   7AAK) revealed that Arg173 is an essential residue for the enzyme to move from ES_2 to ES_3

279   in the elongation process [43]. Despite providing new structural information, neither of these

280   reports, naturally, could offer any light on details beyond the ES_2-state, or the role of Arg26 in

281   the elongation process. Nevertheless, the critical function of Arg26 in enzymatic catalysis has

282   been described based on docking and MD simulations, which suggested Arg26 as the most

283   likely proton donor in the forming of ES and ES_2, but too far away for the next step from ES_2

284   to ES_3 and beyond [45]. However, it was assumed that the elongation process extended from

285   the cofactor without involving rearrangement of cofactor-binding loop [45], a movement that

286   has later been demonstrated in other reports [4, 43]. Our current results support that Arg26 is

287   essential for the formation of the intermediates up to the ES_2 state, possibly in concert with

288   Arg173, where Arg173 likely acts as the coordinator for the bound PBG also stabilising the

289   structure, and where Arg26 has the role as proton donor for upcoming substrates. Both

290   mutants p.R173W and p.R26H are trapped in the ES_2-state, however, neither have, in our

291   studies, responded to the addition of excess substrate. Where p.R173W displays some

292   structural instability as a function of temperature [39], p.R26H resembles wt-HMBS in the

293   ES_2-state in its reaction to temperature. This could be explained by the direct interaction

294   between Arg173 and the bound PBG, where the Arg→Trp substitution disturbs the stabilising

295   hydrogen-bond network, which connects the structural domains via the bound PBG, and

296   yields a lowered thermostability, whereas the interactions that are lost when Arg26 is replaced

297   by His has little or no impact on the structural stability. In addition, the similar secondary

298   structure content between wt-HMBS and p.R26H indicate that the overall mutant structure is

299   not affected by the introduction of the histidine. A modelling of the mutant based on existing

300   high resolution ES$_2$-structure (PDB ID: 7AAK) was therefore seen as sufficient to illustrate

301   the p.R26H structure (Figure 1).

302   Bung et al. also predict, *via* MD simulations, that Arg26 interacts with both Thr58 and

303   Asp61 [45]. Asp61 is located in the active-site loop and is a highly conserved residue,

304   predicted to have important interactions contributing to the stability of the structure [50].

305   Interestingly, Arg173 is also in the vicinity of Asp61 considering the flexible nature of the

306   active-site loop, and it could be speculated that these three residues act in concert. Based on

307   structural evidence, the interactions between Arg26 and the active-site loop (residues Ser57—

308   Lys74) can only occur when the is loop completely closed as shown in the HMBS structure

309   from *Arabidopsis thaliana* (PDB ID: 4HTG) [51]. In this structure the co-factor is in its

310   oxidised form. Thus, we cannot conclude that the closing of the active-site loop occurs as a

311   part of the elongation process. However, p.D61N, p.D61H, and p.D61Y are known disease-

312   causing mutations of Asp61 [52-54] indicating the importance of the loop and these residues

313   for the function of the enzyme.

314   The p.R26H mutation has been reported in several different families with AIP,

315   originating in different countries and spanning several continents. In AIP, there is a large

316   molecular heterogeneity and most mutations are restricted to single or just a few families, but

317   the p.R26H occurs at a known hotspot involving CpG dinucleotides [55] and has also been

318   described to occur *de novo* [28, 29]. Most publications on p.R26H have diagnostic data, but

319   there is relatively little clinical data included and thus any assessment as to severity such as

320   early symptom onset or clinical penetrance is therefore not possible. However, generally, few

321   genotype-phenotype correlations have yet been established in AIP and only null-allele

322   mutations have been described to be associated with a more severe phenotype and higher

323   penetrance [49].

324   Overall, characterising a mutant alone does not explain a correlation between genotype

325   and phenotype, and the clinical penetrance of *HMBS* mutations is most likely modulated by

326   other genetic and environmental factors [49]. Nevertheless, our report, together with previous

327   mutant characterisation at the protein level, demonstrates how the effect of missense

328   mutations can be grouped into at least two groups where i) the enzyme has residual activity

329   yielding some turnover of protein, and ii) the enzyme is blocked for turnover in an

330   intermediate state. These findings can be important in future work exploring alternative

331   treatment options for patients suffering from AIP.

6 Acknowledgement

7 Conflict of interest

AKA, JPK and HJB have filed a patent application for the potential of compound 5-(2chlorophenyl)methyl]-2-hydroxy-3-nitrobenzaldehyde and 4-chloro-3-nitrophenyl(phenyl)methanone for development of a treatment for acute intermittent porphyria. United Kingdom Patent Application No. 1,916,983.8.

8 Author contribution

AKA, JPK and HJB conceived the study; JPK and HJB designed experiments and supervised the study; MSC and ML performed experiments; MSC, ML, JPK and HJB analysed data; MSC wrote the first manuscript draft; AKA, JPK and HJB made revisions and finalised the manuscript.

9 References

1. Wang, B. (2021) The acute hepatic porphyrias, *Transl Gastroenterol Hepatol.* **6**, 24.
2. Phillips, J. D. (2019) Heme biosynthesis and the porphyrias, *Mol Genet Metab.* **128**, 164-177.
3. Brons-Poulsen, J., Christiansen, L., Petersen, N. E., Horder, M. & Kristiansen, K. (2005) Characterization of two isoalleles and three mutations in both isoforms of purified recombinant human porphobilinogen deaminase, *Scand J Clin Lab Invest.* **65**, 93-105.
4. Pluta, P., Roversi, P., Bernardo-Seisdedos, G., Rojas, A. L., Cooper, J. B., Gu, S., Pickersgill, R. W. & Millet, O. (2018) Structural basis of pyrrole polymerization in human porphobilinogen deaminase, *Biochim Biophys Acta Gen Subj.* **1862**, 1948-1955.
5. Grandchamp, B., De Verneuil, H., Beaumont, C., Chretien, S., Walter, O. & Nordmann, Y. (1987) Tissue-specific expression of porphobilinogen deaminase. Two isoenzymes from a single gene, *Eur J Biochem.* **162**, 105-10.
6. Badminton, M. N. & Elder, G. H. (2005) Molecular mechanisms of dominant expression in porphyria, *J Inherit Metab Dis.* **28**, 277-86.
7. Chen, B., Solis-Villa, C., Hakenberg, J., Qiao, W., Srinivasan, R. R., Yasuda, M., Balwani, M., Doheny, D., Peter, I., Chen, R. & Desnick, R. J. (2016) Acute Intermittent Porphyria: Predicted Pathogenicity of HMBS Variants Indicates Extremely Low Penetrance of the Autosomal Dominant Disease, *Hum Mutat.* **37**, 1215-1222.
8. Bissell, D. M., Anderson, K. E. & Bonkovsky, H. L. (2017) Porphyria, *N Engl J Med.* **377**, 862-872.
9. Bonkovsky, H. L., Dixon, N. & Rudnick, S. (2019) Pathogenesis and clinical features of the acute hepatic porphyrias (AHPs), *Mol Genet Metab.* **128**, 213-218.
10. Stein, P. E., Badminton, M. N. & Rees, D. C. (2017) Update review of the acute porphyrias, *Br J Haematol.* **176**, 527-538.
11. Correia, M. A., Sinclair, P. R. & De Matteis, F. (2011) Cytochrome P450 regulation: the interplay between its heme and apoprotein moieties in synthesis, assembly, repair, and disposal, *Drug Metab Rev.* **43**, 1-26.
12. Innala, E., Backstrom, T., Bixo, M. & Andersson, C. (2010) Evaluation of gonadotropin-releasing hormone agonist treatment for prevention of menstrual-related attacks in acute porphyria, *Acta Obstet Gynecol Scand.* **89**, 95-100.
13. Delaby, C., To-Figueras, J., Deybach, J. C., Casamitjana, R., Puy, H. & Herrero, C. (2009) Role of two nutritional hepatic markers (insulin-like growth factor 1 and transthyretin) in the clinical assessment and follow-up of acute intermittent porphyria patients, *J Intern Med.* **266**, 277-85.
14. Handschin, C., Lin, J., Rhee, J., Peyer, A. K., Chin, S., Wu, P. H., Meyer, U. A. & Spiegelman, B. M. (2005) Nutritional regulation of hepatic heme biosynthesis and porphyria through PGC-1alpha, *Cell.* **122**, 505-15.
15. Bylesjo, I., Wikberg, A. & Andersson, C. (2009) Clinical aspects of acute intermittent porphyria in northern Sweden: a population-based study, *Scand J Clin Lab Invest.* **69**, 612-8.
16. Eales, L. (1979) Porphyria and the dangerous life-threatening drugs, *S Afr Med J.* **56**, 914-7.
17. Pallet, N., Mami, I., Schmitt, C., Karim, Z., Francois, A., Rabant, M., Nochy, D., Gouya, L., Deybach, J. C., Xu-Dubois, Y., Thervet, E., Puy, H. & Karras, A. (2015) High prevalence of and potential mechanisms for chronic kidney disease in patients with acute intermittent porphyria, *Kidney Int.* **88**, 386-95.
18. Baravelli, C. M., Sandberg, S., Aarsand, A. K., Nilsen, R. M. & Tollanes, M. C. (2017) Acute hepatic porphyria and cancer risk: a nationwide cohort study, *J Intern Med.* **282**, 229-240.
19. Saberi, B., Naik, H., Overbey, J. R., Erwin, A. L., Anderson, K. E., Bissell, D. M., Bonkovsky, H. L., Phillips, J. D., Wang, B., A, K. S., B, M. M., Desnick, R. J. & Balwani, M. (2021) Hepatocellular Carcinoma in Acute Hepatic Porphyrias: Results from the Longitudinal Study of the U.S. Porphyrias Consortium, *Hepatology.* **73**, 1736-1746.
20. Lissing, M., Vassiliou, D., Floderus, Y., Harper, P., Bottai, M., Kotopouli, M., Hagstrom, H., Sardh, E. & Wahlin, S. (2022) Risk of primary liver cancer in acute hepatic porphyria patients: A matched cohort study of 1,244 individuals, *J Intern Med.*
21. Lissing, M., Nowak, G., Adam, R., Karam, V., Boyd, A., Gouya, L., Meersseman, W., Melum, E., Oldakowska-Jedynak, U., Reiter, F. P., Colmenero, J., Sanchez, R., Herden, U., Langendonk, J., Ventura, P., Isoniemi, H., Boillot, O., Braun, F., Perrodin, S., Mowlem, E., Wahlin, S., European, L. & Intestine Transplant, A. (2021) Liver Transplantation for Acute Intermittent Porphyria, *Liver Transpl.* **27**, 491-501.
22. Bustad, H. J., Kallio, J. P., Vorland, M., Fiorentino, V., Sandberg, S., Schmitt, C., Aarsand, A. K. & Martinez, A. (2021) Acute Intermittent Porphyria: An Overview of Therapy Developments and Future Perspectives Focusing on Stabilisation of HMBS and Proteostasis Regulators, *Int J Mol Sci.* **22**.
23. Wang, B. (2021) Novel treatment options for acute hepatic porphyrias, *Curr Opin Gastroenterol.* **37**, 194-199.
24. Helliwell, J. R. (2021) The crystal structures of the enzyme hydroxymethylbilane synthase, also known as porphobilinogen deaminase, *Acta Crystallogr F Struct Biol Commun.* **77**, 388-398.

415 25. Hadener, A., Matzinger, P. K., Battersby, A. R., McSweeney, S., Thompson, A. W., Hammersley, A. P.,
416 Harrop, S. J., Cassetta, A., Deacon, A., Hunter, W. N., Nieh, Y. P., Raftery, J., Hunter, N. & Helliwell, J. R.
417 (1999) Determination of the structure of seleno-methionine-labelled hydroxymethylbilane synthase in its active
418 form by multi-wavelength anomalous dispersion, *Acta Crystallogr D Biol Crystallogr.* **55**, 631-43.
419 26. Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Warren, M. J.,
420 Woodcock, S. C. & Jordan, P. M. (1992) Structure of porphobilinogen deaminase reveals a flexible multidomain
421 polymerase with a single catalytic site, *Nature.* **359**, 33-9.
422 27. Song, G., Li, Y., Cheng, C., Zhao, Y., Gao, A., Zhang, R., Joachimiak, A., Shaw, N. & Liu, Z. J. (2009)
423 Structural insight into acute intermittent porphyria, *FASEB J.* **23**, 396-404.
424 28. Younger, D. S. & Tanji, K. (2015) Demyelinating neuropathy in genetically confirmed acute intermittent
425 porphyria, *Muscle Nerve.* **52**, 916-7.
426 29. Puy, H., Deybach, J. C., Lamoril, J., Robreau, A. M., Da Silva, V., Gouya, L., Grandchamp, B. &
427 Nordmann, Y. (1997) Molecular epidemiology and diagnosis of PBG deaminase gene defects in acute
428 intermittent porphyria, *Am J Hum Genet.* **60**, 1373-83.
429 30. Goncharova, M., Pshenichnikova, O., Luchinina, Y., Pustovoit, Y., Karpova, I. & Surin, V. (2019)
430 Molecular genetic study of acute intermittent porphyria in Russia: HMBS gene mutation spectrum and problem
431 of penetrance, *Clin Genet.* **96**, 91-97.
432 31. To-Figueras, J., Badenas, C., Carrera, C., Munoz, C., Mila, M., Lecha, M. & Herrero, C. (2006) Genetic and
433 biochemical characterization of 16 acute intermittent porphyria cases with a high prevalence of the R173W
434 mutation, *J Inherit Metab Dis.* **29**, 580-5.
435 32. Kauppinen, R., Mustajoki, S., Pihlaja, H., Peltonen, L. & Mustajoki, P. (1995) Acute intermittent porphyria
436 in Finland: 19 mutations in the porphobilinogen deaminase gene, *Hum Mol Genet.* **4**, 215-22.
437 33. Moran-Jimenez, M. J., Borrero-Corte, M. J., Jara-Rubio, F., Garcia-Pastor, I., Diaz-Diaz, S., Castelbon-
438 Fernandez, F. J., Enriquez-de-Salamanca, R. & Mendez, M. (2020) Molecular Analysis of 55 Spanish Patients
439 with Acute Intermittent Porphyria, *Genes (Basel).* **11**.
440 34. Mendez, M., Moran-Jimenez, M. J., Gomez-Abecia, S., Garcia-Bravo, M., Garrido-Astray, M. C.,
441 Fontanellas, A., Poblete-Gutierrez, P., Frank, J. & Enriquez de Salamanca, R. (2009) Identification and
442 characterization of HMBS gene mutations in Spanish patients with acute intermittent porphyria, *Cell Mol Biol*
443 *(Noisy-le-grand).* **55**, 55-63.
444 35. Ulbrichova, D., Hrdinka, M., Saudek, V. & Martasek, P. (2009) Acute intermittent porphyria--impact of
445 mutations found in the hydroxymethylbilane synthase gene on biochemical and enzymatic protein properties,
446 *FEBS J.* **276**, 2106-15.
447 36. Floderus, Y., Shoolingin-Jordan, P. M. & Harper, P. (2002) Acute intermittent porphyria in Sweden.
448 Molecular, functional and clinical consequences of some new mutations found in the porphobilinogen deaminase
449 gene, *Clin Genet.* **62**, 288-97.
450 37. Yang, C. C., Kuo, H. C., You, H. L., Wang, J., Huang, C. C., Liu, C. Y., Lan, M. Y., Stephenson, D. A. &
451 Lee, M. J. (2008) HMBS mutations in Chinese patients with acute intermittent porphyria, *Ann Hum Genet.* **72**,
452 683-6.
453 38. Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.-H., Goto, Y., Réfrégiers, M. & Kardos,
454 J. (2018) BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the
455 circular dichroism spectra, *Nucleic Acids Research.* **46**, W315-W322.
456 39. Bustad, H. J., Vorland, M., Ronneseth, E., Sandberg, S., Martinez, A. & Toska, K. (2013) Conformational
457 stability and activity analysis of two hydroxymethylbilane synthase mutants, K132N and V215E, with different
458 phenotypic association with acute intermittent porphyria, *Biosci Rep.* **33**.
459 40. Martin-Malpartida, P., Hausvik, E., Underhaug, J., Torner, C., Martinez, A. & Macias, M. J. (2021)
460 HTSDSF Explorer, A Novel Tool to Analyze High-throughput DSF Screenings.
461 41. Sato, H., Sugishima, M., Tsukaguchi, M., Masuko, T., Iijima, M., Takano, M., Omata, Y., Hirabayashi, K.,
462 Wada, K., Hisaeda, Y. & Yamamoto, K. (2021) Crystal structures of hydroxymethylbilane synthase complexed
463 with a substrate analog: a single substrate-binding site for four consecutive condensation steps, *Biochem J.* **478**,
464 1023-1042.
465 42. Li, N., Chu, X., Wu, L., Liu, X. & Li, D. (2008) Functional studies of rat hydroxymethylbilane synthase,
466 *Bioorg Chem.* **36**, 241-51.
467 43. Bustad, H. J., Kallio, J. P., Laitaoja, M., Toska, K., Kursula, I., Martinez, A. & Jänis, J. (2021)
468 Characterization of porphobilinogen deaminase mutants reveals that arginine-173 is crucial for polypyrrole
469 elongation mechanism, *iScience.* **24**, 102152.
470 44. Shoolingin-Jordan, P. M., Warren, M. J. & Awan, S. J. (1996) Discovery that the assembly of the
471 dipyrromethane cofactor of porphobilinogen deaminase holoenzyme proceeds initially by the reaction of
472 preuroporphyrinogen with the apoenzyme, *Biochem J.* **316 (Pt 2)**, 373-6.

473   45. Bung, N., Roy, A., Chen, B., Das, D., Pradhan, M., Yasuda, M., New, M. I., Desnick, R. J. & Bulusu, G.
474   (2018) Human hydroxymethylbilane synthase: Molecular dynamics of the pyrrole chain elongation identifies
475   step-specific residues that cause AIP, *Proc Natl Acad Sci U S A.* **115**, E4071-E4080.
476   46. Jordan, P. M. & Woodcock, S. C. (1991) Mutagenesis of arginine residues in the catalytic cleft of
477   Escherichia coli porphobilinogen deaminase that affects dipyrromethane cofactor assembly and tetrapyrrole
478   chain initiation and elongation, *Biochem J.* **280 (Pt 2)**, 445-9.
479   47. Gill, R., Kolstoe, S. E., Mohammed, F., Al, D. B. A., Mosely, J. E., Sarwar, M., Cooper, J. B., Wood, S. P.
480   & Shoolingin-Jordan, P. M. (2009) Structure of human porphobilinogen deaminase at 2.8 A: the molecular basis
481   of acute intermittent porphyria, *Biochem J.* **420**, 17-25.
482   48. Lander, M., Pitt, A. R., Alefounder, P. R., Bardy, D., Abell, C. & Battersby, A. R. (1991) Studies on the
483   mechanism of hydroxymethylbilane synthase concerning the role of arginine residues in substrate binding,
484   *Biochem J.* **275 (Pt 2)**, 447-52.
485   49. Lenglet, H., Schmitt, C., Grange, T., Manceau, H., Karboul, N., Bouchet-Crivat, F., Robreau, A. M.,
486   Nicolas, G., Lamoril, J., Simonin, S., Mirmiran, A., Karim, Z., Casalino, E., Deybach, J. C., Puy, H., Peoc'h, K.
487   & Gouya, L. (2018) From a dominant to an oligogenic model of inheritance with environmental modifiers in
488   acute intermittent porphyria, *Hum Mol Genet.* **27**, 1164-1173.
489   50. Brownlie, P. D., Lambert, R., Louie, G. V., Jordan, P. M., Blundell, T. L., Warren, M. J., Cooper, J. B. &
490   Wood, S. P. (1994) The three-dimensional structures of mutants of porphobilinogen deaminase: toward an
491   understanding of the structural basis of acute intermittent porphyria, *Protein Sci.* **3**, 1644-50.
492   51. Roberts, A., Gill, R., Hussey, R. J., Mikolajek, H., Erskine, P. T., Cooper, J. B., Wood, S. P., Chrystal, E. J.
493   & Shoolingin-Jordan, P. M. (2013) Insights into the mechanism of pyrrole polymerization catalysed by
494   porphobilinogen deaminase: high-resolution X-ray studies of the Arabidopsis thaliana enzyme, *Acta Crystallogr
495   D Biol Crystallogr.* **69**, 471-85.
496   52. von Brasch, L., Zang, C., Haverkamp, T., Schlechte, H., Heckers, H. & Petrides, P. E. (2004) Molecular
497   analysis of acute intermittent porphyria: mutation screening in 20 patients in Germany reveals 11 novel
498   mutations, *Blood Cells Mol Dis.* **32**, 309-14.
499   53. Gregor, A., Schneider-Yin, X., Szlendak, U., Wettstein, A., Lipniacka, A., Rufenacht, U. B. & Minder, E. I.
500   (2002) Molecular study of the hydroxymethylbilane synthase gene (HMBS) among Polish patients with acute
501   intermittent porphyria, *Hum Mutat.* **19**, 310.
502   54. Whatley, S. D., Woolf, J. R. & Elder, G. H. (1999) Comparison of complementary and genomic DNA
503   sequencing for the detection of mutations in the HMBS gene in British patients with acute intermittent
504   porphyria: identification of 25 novel mutations, *Hum Genet.* **104**, 505-10.
505   55. Cooper, D. N. & Youssoufian, H. (1988) The CpG dinucleotide and human genetic disease, *Hum Genet.* **78**,
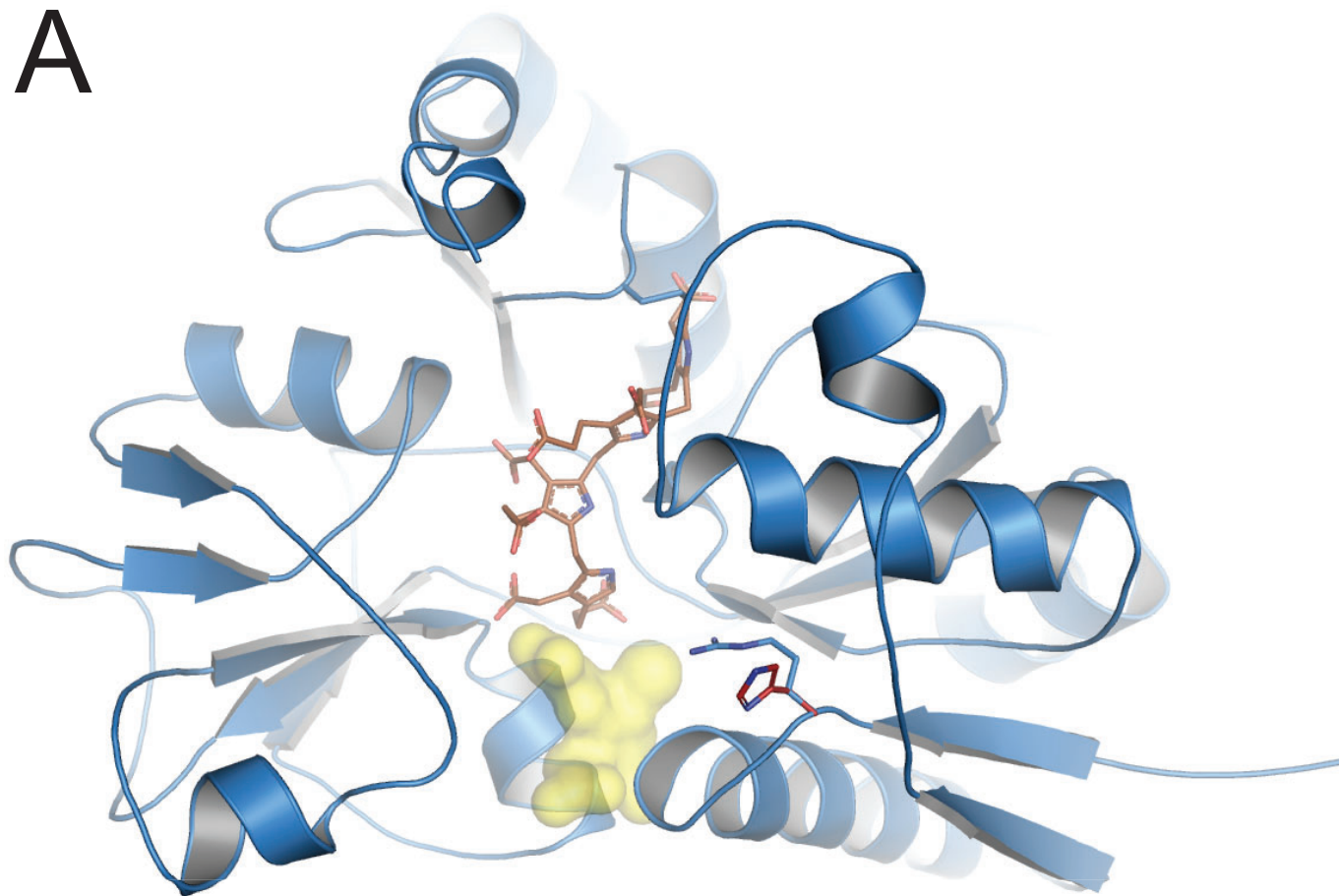506   151-5.
507

508

509   # 10 Tables

510   **Table 1. Prediction of secondary structure content.** The secondary structure content based

511   on far-UV CD spectrum recorded at 190–250 nm was estimated in percentage using the

512   BeStSel software derived from DSSP [38]. Multiple unpaired t-test showed no significant

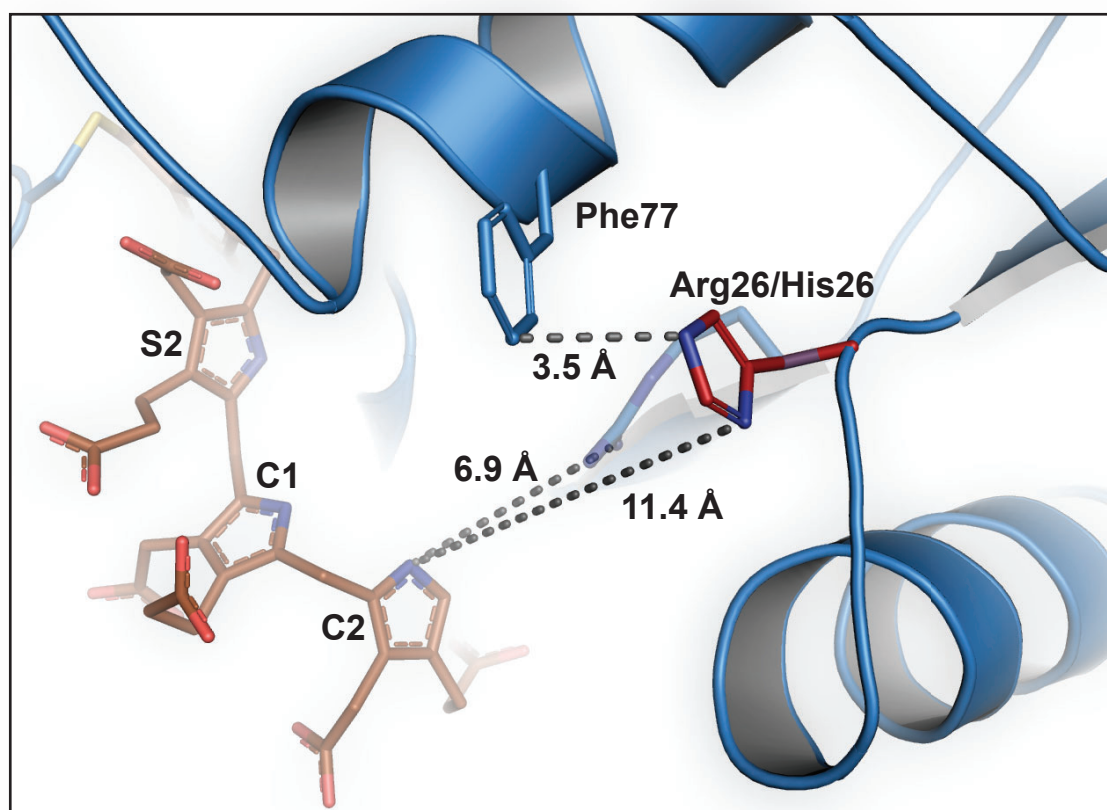513   difference between wt-HMBS and p.R26H.

	Wt-HMBS (%)	p.R26H (%)
Helix	18.63 ± 1.70	15.53 ± 3.45
Antiparallel	25.50 ± 2.36	23.43 ± 3.43
Parallel	2.73 ± 1.70	7.20 ± 2.71
Turn	12.23 ± 0.84	11.93 ± 0.85
Others	40.93 ± 0.80	41.90 ± 0.87

514

A

B

S2

C1

C2

Phe77

Arg26/His26

3.5 Å

6.9 Å

11.4 Å

A

B

C

ES$_2$ (38573.16 Da)

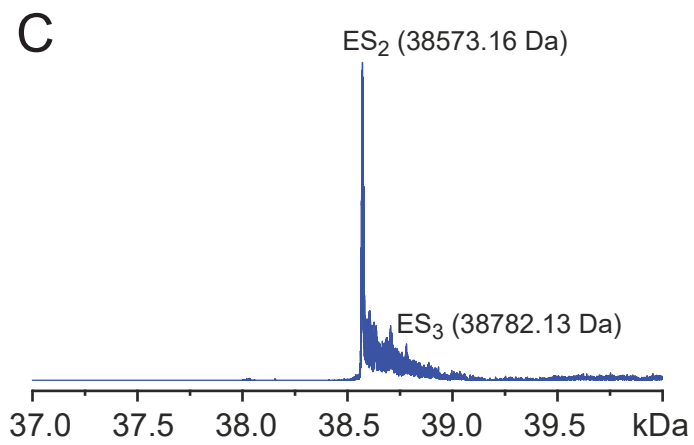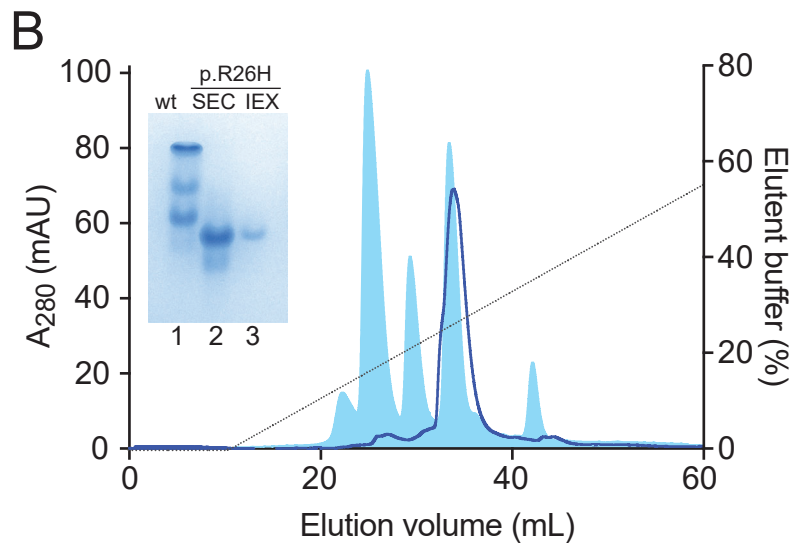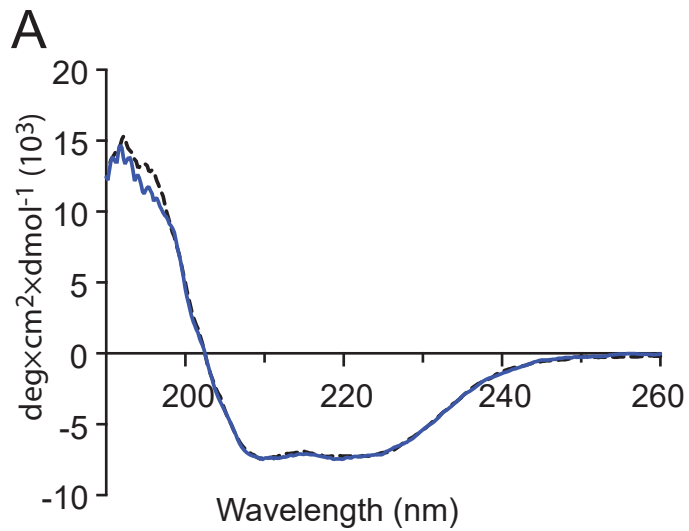ES$_3$ (38782.13 Da)

D

E$_{holo}$

ES

ES$_2$

ES$_3$
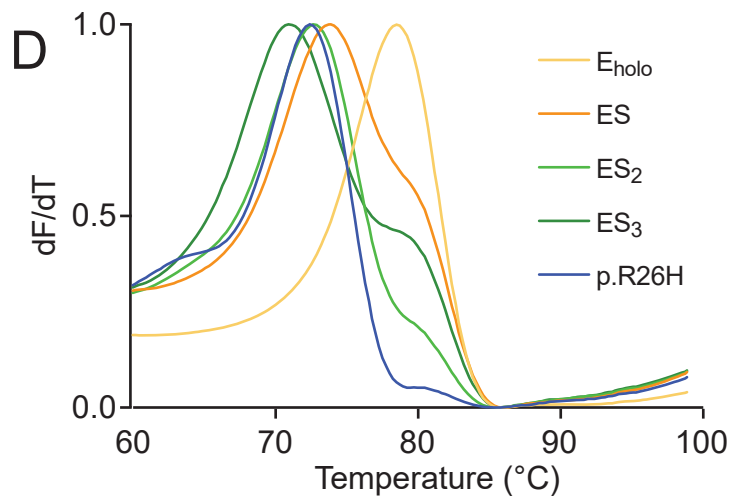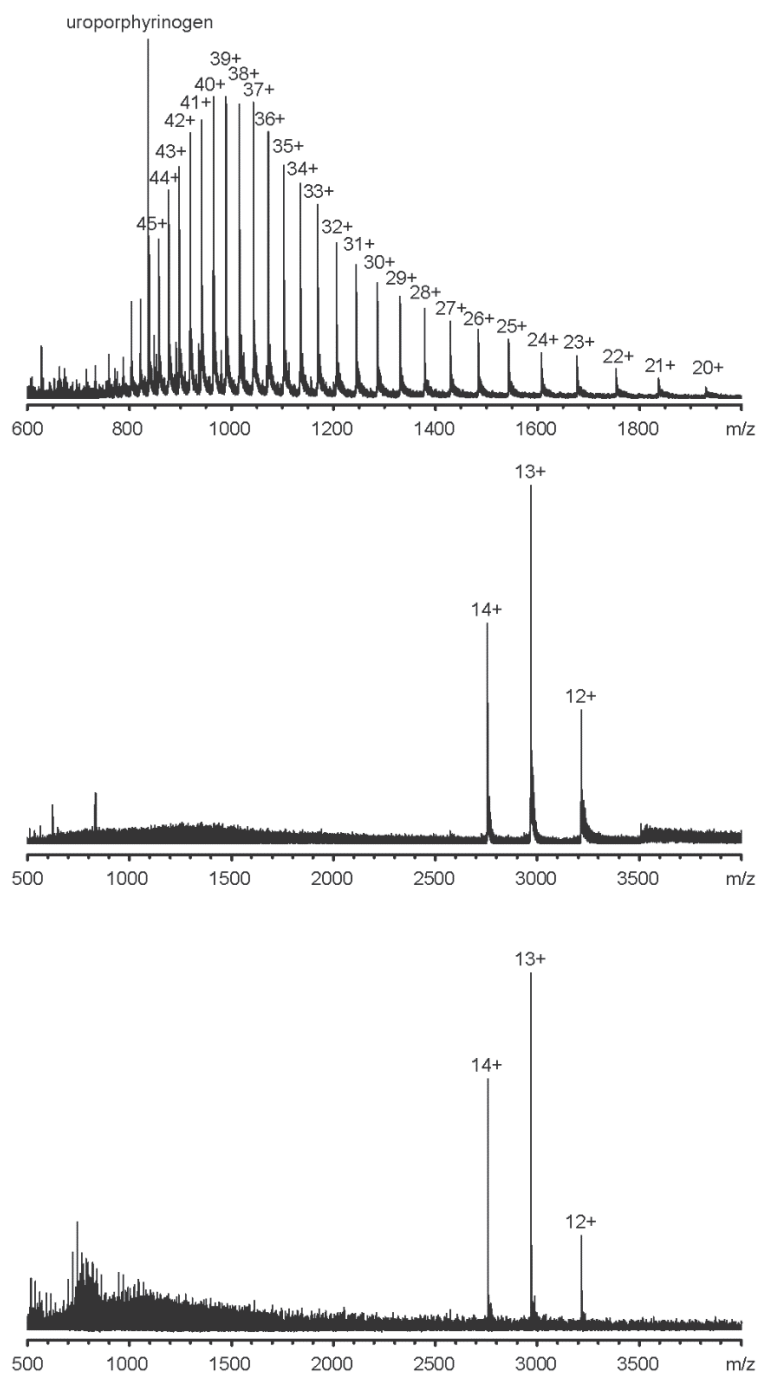
p.R26H

Supporting information

Figure S1: High-resolution mass spectra of HMBS-mutant p.R26H. (A) p.R26H measured at denaturing conditions with at 1 µM protein. Numbers denote different protein ion-charge states as $[M+nH]^{n+}$. The high charges and wide charge state distribution indicates that the protein is fully unfolded. Uroporphyrinogen is also detected in denaturing conditions. (B) p.R26H in native-like conditions at 5 µM concentration. Low charges and narrow charge state distribution shows that the protein is folded, and the mutation does not cause unfolding of the enzyme. (C) p.R26H in native-like conditions measured in the presence of 10× PBG (50 µM). No change in the enzyme intermediates was observed, indicating that in the p.R26H mutant reaction is halted at the ES_2 intermediate.
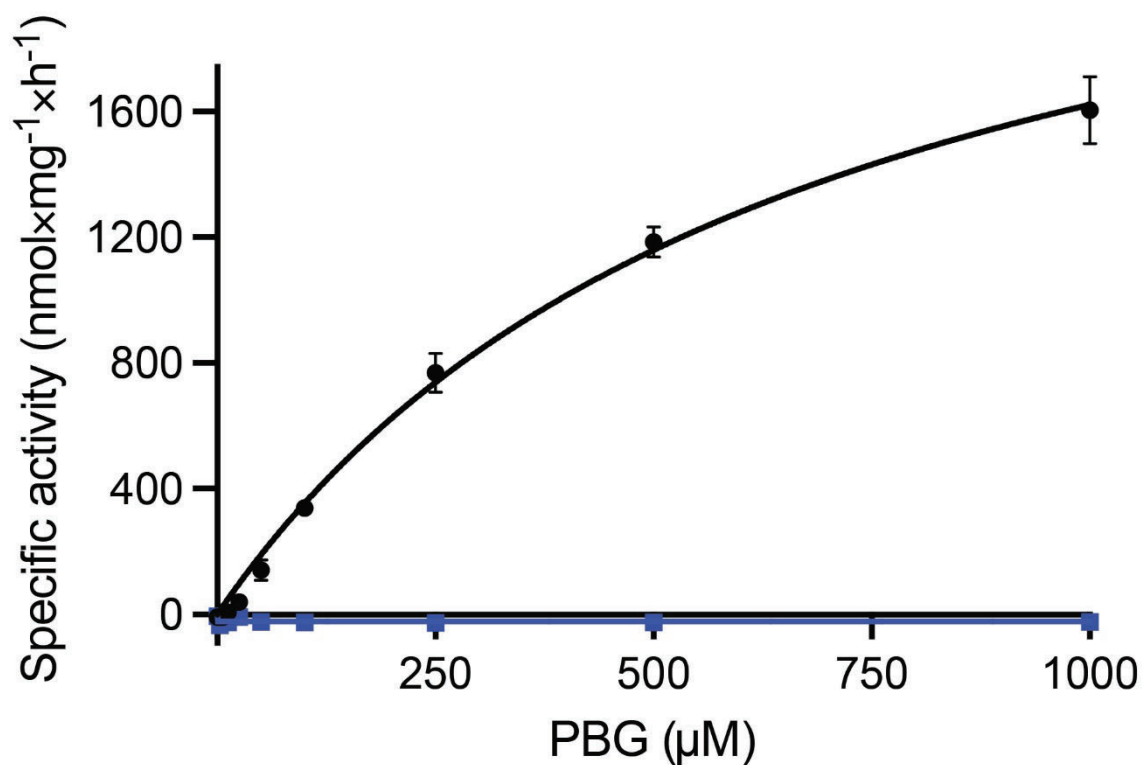
Figure S2: The catalytic activity of wt-HMBS and p.R26H as a function of substrate (PBG) concentration. Wt-HMBS and p.R26H measured at standard conditions with 5 µg protein and varying PBG concentrations (0–1000 µM) for a reaction time of 4 min at 37 °C. The specific activity of HMBS was defined as nmol of uroporphyrinogen I/h per mg of enzyme under the given assay conditions. The data was fitted to Michaelis–Menten kinetics.