UNIVERSITY OF SOUTHAMPTON

STRUCTURAL CHARACTERISATION OF THE ACTIVE SITE OF *ESCHERICHIA COLI* 5-AMINOLAEVULINIC ACID DEHYDRATASE

by

Edwin Bernard Norton

A thesis submitted for the degree of

Doctor of Philosophy

Division of Biochemistry and Molecular Biology

School of Biological Sciences

Faculty of Science

September 1999

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE
DIVISION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Doctor of Philosophy

STRUCTURAL CHARACTERISATION OF THE ACTIVE SITE OF *ESCHERICHIA COLI* 5-AMINOLAEVULINIC ACID DEHYDRATASE

by Edwin Bernard Norton

5-Aminolaevulinic acid dehydratase (ALAD) is an early enzyme in the biosynthesis pathway of tetrapyrrole compounds, catalysing the condensation between two 5-aminolaevulinic acid (ALA) molecules to form porphobilinogen (PBG), the monopyrrole precursor of all tetrapyrroles. Within the active site of *E. coli* ALAD there are two discrete substrate binding sites. The first substrate molecule to bind to the enzyme forms the propionate half of PBG whilst the second ALA molecule forms the acetate half, leading to the terminology of "P" and "A" ALA binding sites.

This thesis describes the co-crystallisation and structural determination of *E. coli* ALAD with ALA present at both the P- and A-sites allowing extensive characterisation of substrate binding residues. These observations are confirmed by the crystal structure of ALAD in complex with the product, PBG. Site directed mutagenesis of residues implicated in substrate binding and catalysis by these structural studies has been carried out, and the contribution of each residue to PBG formation determined. Two invariant P-site lysine residues, K247 and K195, were the subject of further study. Mutants K247A and K247C proved to be completely inactive whereas mutants K195A and K195C retained < 0.1% of wild type activity. Activity could be rescued by modification of K247C and K195C to lysine analogues with bromoethylamine. Low levels of activity shown by the mutant K247N are explained by its crystal structure in complex with laevulinic acid. The pK_a of Lys 247 is shown to be suppressed to 7.5 by Lys 195, allowing it to form a Schiff base with substrate ALA at physiological pH values. *E. coli* ALAD in complex with the inhibitor molecule 5-chlorolaevulinic acid has been crystallised and the structure determined, revealing alkylation of Lys 195. Conclusions are drawn regarding the contributions of active site residues to the mechanism of PBG formation.

Acknowledgements

I would like to thank my supervisor, Professor Peter Shoolingin-Jordan for his time spent assisting and guiding me through this work.

I am also very grateful to:

Dr. Muhammad Sawar for preparation of the mutants studied in this thesis and general insights into molecular biology and the workings of the C.I.A.

Dr. Jack Cheung for advice on organic chemistry and synthesis of the 5-chlorolaevulinic acid inhibitor molecule.

Dr. Iain Campuzano for his help with mass spectrometric analysis of proteins.

Dr. Adrian Lloyd for tips on protein purification and general biochemistry

Dr. Darren Thompson, Dr. Gareth Lewis and Dr. Jon Cooper for assistance with protein crystallisation, data collection and processing.

Everybody in the lab, namely Neil, Mark, Shu-Fen, Ipsita, Abeer, Sooad, Will, Nicky, Graham, Jeremy, Rob, Luke, Jo, Julie, and Loraine.

My girlfriend Jo, for keeping me sane and happy.

My mother and sister for their moral (and financial) support.

Chapter 1. Introduction

Chapter 2: Materials and methods

List of figures

List of schemes

List of tables

Abbreviations used in this thesis

ALA	5-aminolaevulinic acid
ALAD	ALA dehydratase
ALAS	ALA synthase
APS	ammonium persulphate
ATP	adenosine 5'-triphosphate
CLA	chlorolaevulinic acid
CoA	Coenzyme A
DEAE	diethylaminoethyl
EDTA	ethylene diamine tetraacetic acid
ESMS	electrospray mass spectrometry
EXAFS	extended X-ray absorption fine structure
GABA	γ-amino butyric acid
kDa	kilodalton
LB	Luria broth
MER	modified Ehrlich's reagent
NEM	N-ethylmaleimide
n.m.r.	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
PBG	porphobilinogen
PBGD	PBG deaminase
PBGS	PBG synthase

1.1 The tetrapyrrole pathway

The tetrapyrrole biosynthesis pathway can be considered to be broadly similar through all living organisms with the first common intermediate being 5-aminolaevulinic acid (ALA) which is produced either by the C_5 pathway from glutamate or an alternate pathway utilising glycine and succinyl-CoA. Two molecules of ALA are then condensed by 5-aminolaevulinic acid dehydratase to form porphobilinogen (PBG). Four molecules of PBG are next polymerised by porphobilinogen deaminase (PBGD) to give preuroporphyrinogen which is cyclised and one of the pyrrole rings is rearranged to give uroporphyrinogen III. Uroporphyrinogen III is the universal precursor from which haems, chlorophylls and other biologically important molecules are biosynthesised (scheme 1.1).

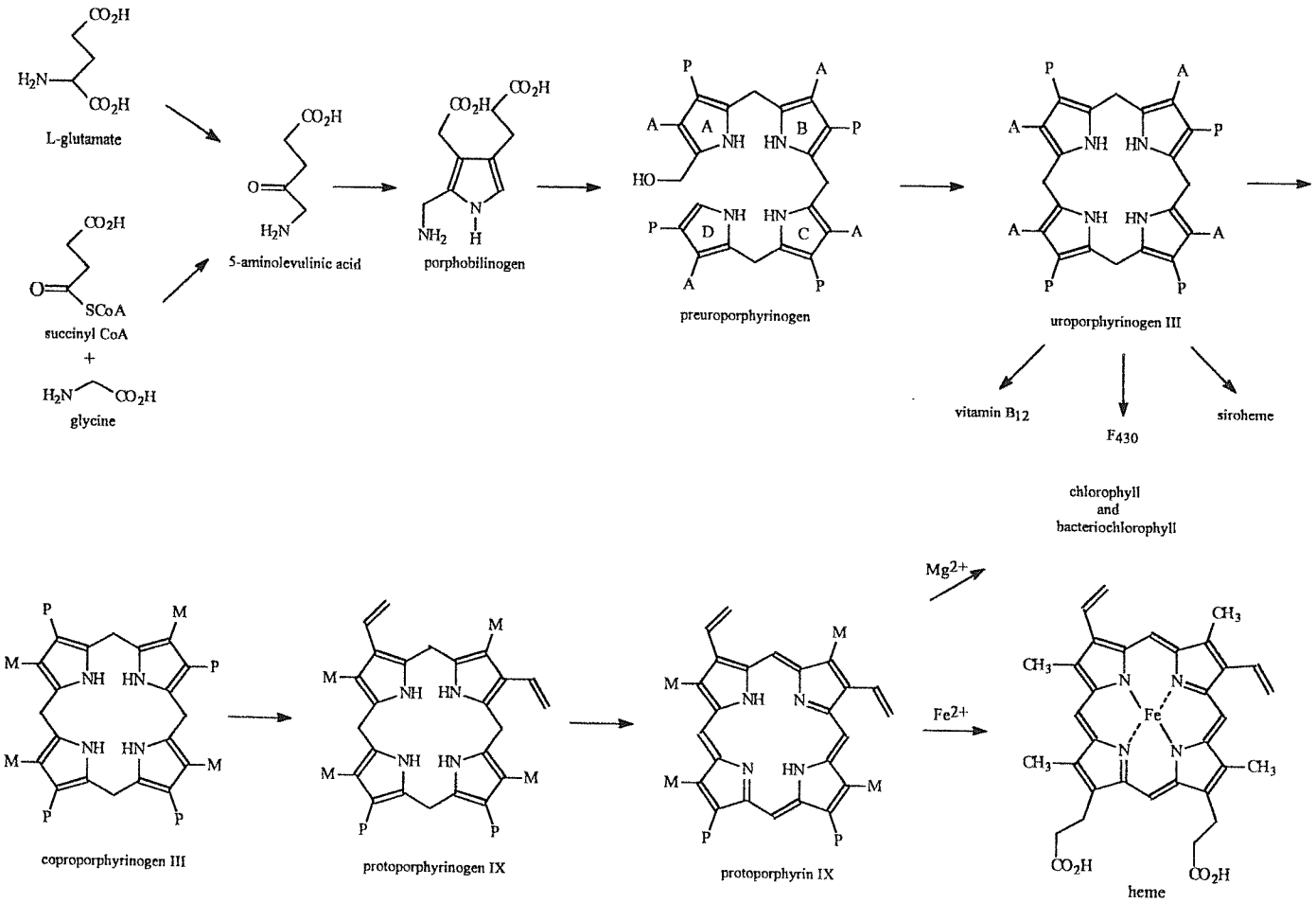
1.2 The biosynthesis of ALA

ALA is the first intermediate specifically committed to tetrapyrrole biosynthesis. In higher plants, algae and some bacteria ALA formation occurs *via* the C_5 pathway (scheme 1.3,a). In animals, yeast and other bacteria ALA is formed from succinyl-CoA and glycine by ALA synthase (scheme 1.3,b). Both are discussed below.

1.2.1 The C_5 pathway

The C_5 pathway, discovered in the mid 1970s (1) is now considered to be a more ubiquitous path for ALA formation than the ALA synthase route and has been shown to involve three soluble enzymes (2) and the involvement of $tRNA^{glu}$ as an essential component of the pathway. The first enzyme in the pathway is glutamayl-$tRNA^{glu}$ synthase. Work carried out on this enzyme isolated from barley chloroplasts (3) suggests its role is to couple glutamate to $tRNA^{glu}$ in the presence of ATP and magnesium ions. The second enzyme, glutamyl-$tRNA^{glu}$ reductase catalyses the reduction of glutamyl-$tRNA^{glu}$ to the aldehyde oxidation level in a NADPH requiring reaction, forming glutamate 1-semialdehyde (GSA) (4). The conversion of GSA to ALA is mediated by GSA aminotransferase. This reaction involves the formation of a Schiff base between the C-1 aldehyde group of GSA with the amino group of an enzyme bound pyridoxamine 5'- phosphate.

Scheme 1.1 The tetrapyrrole pathway



The amino group from the pyridoxamine 5'-phosphate is transferred to the substrate to yield a 4,5-diaminovaleric acid intermediate that remains enzyme bound. The transfer of the amino group at the C-4 position to the pyridoxal 5'-phosphate thus regenerates the pyridoxamine 5'-phosphate and yields the product ALA molecule.

1.2.2 ALA synthase

ALA synthase operates in all mammalian cells, yeast, photosynthetic bacteria, and some algae. In all eukaryotic cells ALA synthase is located in the mitochondria but is encoded by a nuclear gene, with the pre-enzyme having a targetting sequence that is removed after import into the mitochondria.(5).

All ALA synthases exist as homodimers with subunit Mrs of 40k-60k with none accepting any amino acid substrate other than glycine. Glycine analogues are poor inhibitors (6), however, specificity for the acyl-CoA substrate is less rigid, with analogues of the natural substrate being accepted by the enzyme (7). Steady state kinetic studies performed on ALA synthase from *R. sphaeroides Y* have revealed a reaction course with an ordered binding of glycine followed by succinyl-CoA. The release of CoASH from the enzyme occurs before the release of the product, ALA, (8) as shown in scheme 1.2.

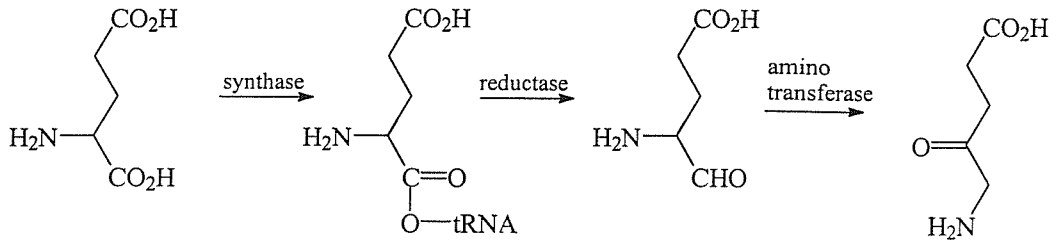
Scheme 1.2. Ordered binding of substrates to ALA-synthase

Glycine	Succinyl-CoA	Co-ASH	CO_2	ALA
↓	↓	↑	↑	↑

All ALA-synthases require a loosely bound pyridoxal 5'-phosphate molecule for activity. Glycine binds to this pyridoxal 5'-phosphate-enzyme complex in the initial event of ALA formation , *via* a Schiff base. The reaction then proceeds by formation of a stabilised carbanion (or its equivalent) at the glycine α-carbon atom, which reacts with succinyl-CoA to form ALA.

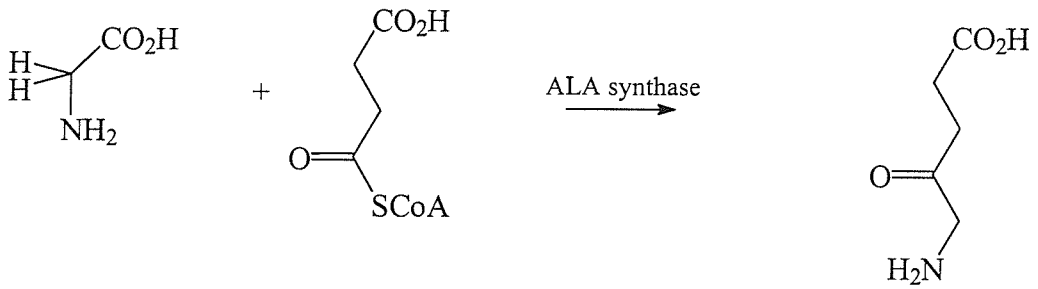
Scheme 1.3

A. ALA formation *via* the C_5 pathway



L-glutamate glutamyl-$tRNA^{glu}$ glutamate 1-semialdehyde 5-aminolaevulinic acid

B. ALA synthase catalysed ALA formation.



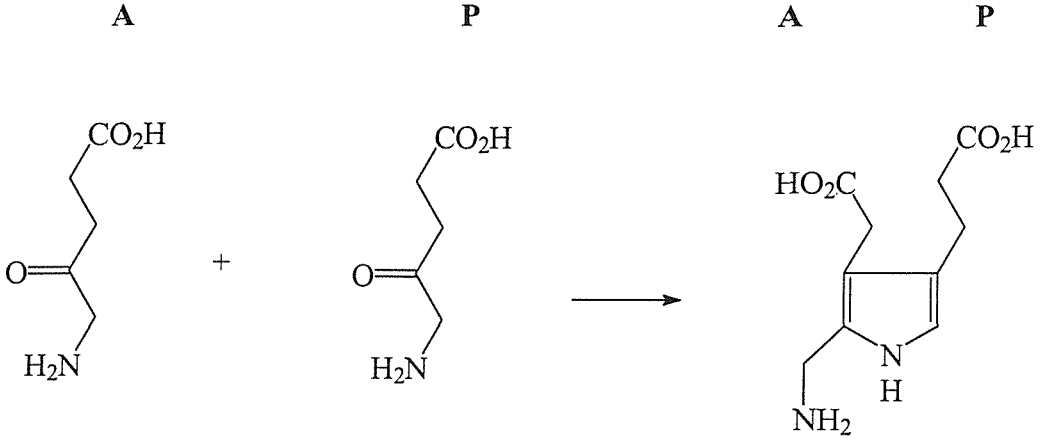
glycine succinyl-CoA 5-aminolaevulinic acid

1.3. 5-Aminolaevulinic acid dehydratase

5-Aminolaevulinic acid dehydratase (ALAD) is a metalloprotein that catalyses the asymmetric formation of porphobilinogen from two molecules of 5-aminolaevulinic acid (ALA) scheme 4. This reaction is the first step common in the biosynthesis of all tetrapyrroles which, in turn, are precursors for vitamin B_{12}, chlorophyll, haem, and many other biologically important molecules. ALAD is a highly conserved homo-octameric protein, with eight active sites per molecule. The mechanism of PBG formation is thought to be identical in all organisms.

ALAD was first isolated in 1955 from ox liver (9) and avian erythrocytes (10). It has now been shown to be present in virtually all living organisms and has been isolated from such widely different sources as human erythrocytes (11), spinach (12) and *E. coli* (13).

Scheme 4. ALAD catalysed PBG formation



5-Aminolaevulinic acid Porphobilinogen

1.3.1 Metal requirements of ALAD

All known ALADs show a requirement for bivalent metal ions and are inhibited by the metal chelator EDTA. Mammalian and some bacterial ALADs show a requirement for Zn^{2+}, whereas plant and photosynthetic bacterial ALADs exhibit a

requirement for Mg^{2+}. The zinc dependent ALADs contain four oxygen-sensitive cysteines that have been considered important for interacting with the zinc ion. Oxidation of these cysteine residues, to form di-sulphide bridges, results in inactivation (13). Magnesium-dependent enzymes, however, contain aspartates at some of the equivalent positions. Substitution, *via* site directed mutagenesis, of aspartate residues in Mg^{2+} dependent enzymes by cysteines has shown a change in requirement to Zn^{2+} .

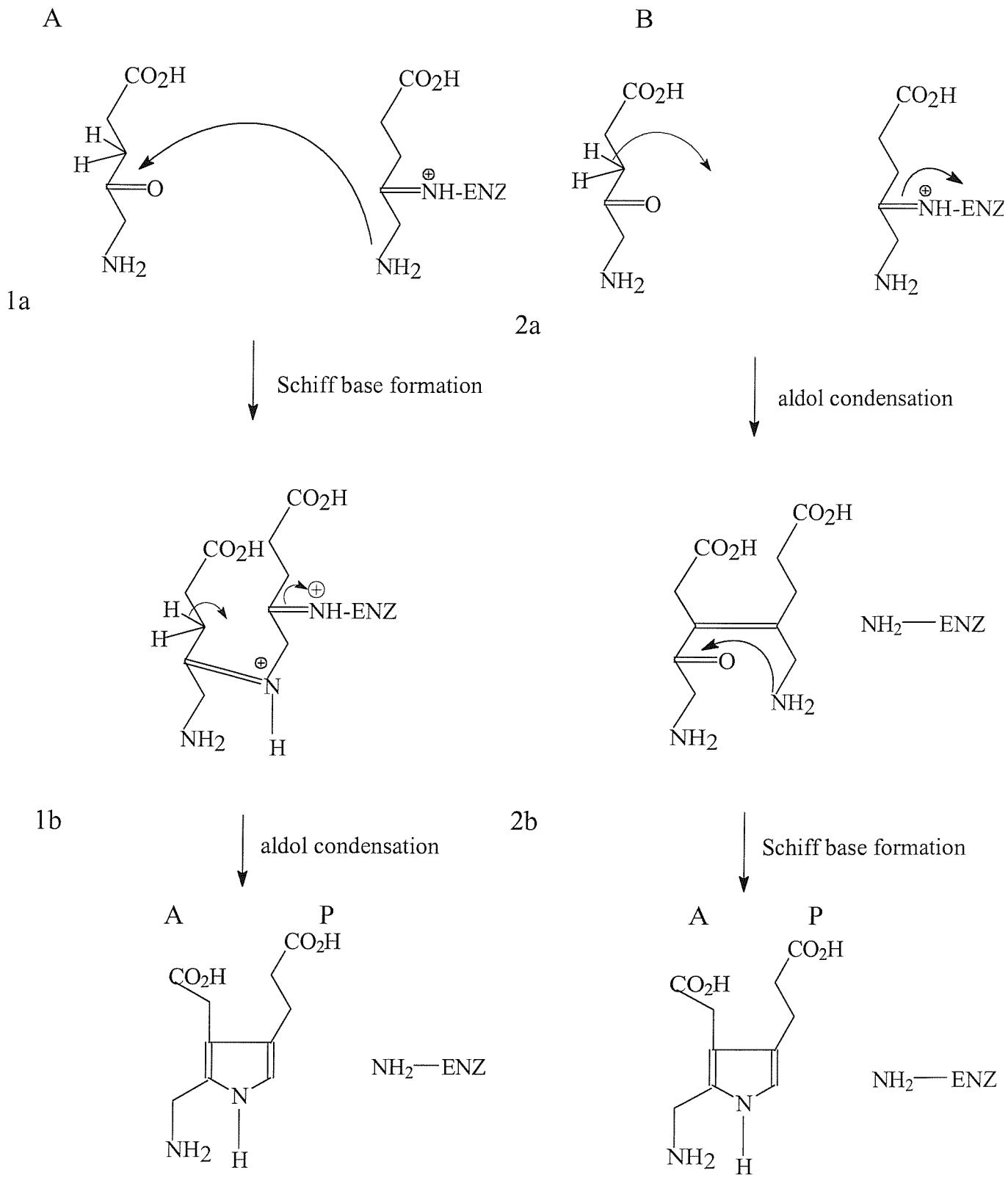
ALADs that require Zn^{2+} are highly oxygen sensitive, whereas ALADs of plant origin (Mg^{2+} dependent) are not. It is believed that the aspartate residues found in plant ALADs give rise to this oxygen insensitivity and protect the plant ALAD in the more oxidising environment of a photosynthesising plant cell.

Another metal ion demonstrated to have an effect on the activity of ALAD is lead. Studies have shown that the activities of human (14) and rat erythrocyte ALADs are inhibited by lead both *in vivo* and *in vitro*, however zinc is able to overcome this inhibition. One direct effect of erythrocyte ALAD inhibition by lead is the build up of ALA in the blood, which is one of the major manifestations of acute lead poisoning. This in turn leads to severe neurological symptoms as ALA may act as a 4-aminobutyric acid agonist and has been found to inhibit the activity of mammalian spinal motor neurones possibly *via* 4-aminobutyric acid synapses (15). Anaemia caused by lead poisoning may also stem from ALAD inhibition.

1.3.2 Mechanism of PBG formation

ALAD (also known as porphobilinogen synthase) catalyses the formation of PBG, the monopyrrole precursor of all tetrapyrroles, *via* the condensation of two ALA molecules. One ALA molecule forms the P (propionic acid) side of PBG, with the amino nitrogen being incorporated into the pyrrole ring of PBG. This ALA molecule has been shown to bind at the "P-site" of the active site *via* a Schiff base linkage with an invariant lysine. The second ALA molecule involved in the reaction forms the A (acetic acid) site of PBG, retaining its primary amino group (scheme 1.4). The binding of this second substrate molecule at the "A side" of the active site is reliant on a divalent ion. Chaudhry *et al* (16) demonstrated that by removing the metal ion

Scheme 1.5 Two possible broad mechanistic pathways for PBG formation.



enzymatic activity is prevented but ALA binding at the P-site through the Schiff base is unaffected. ALA is a highly reactive molecule and, unless controlled enzymatically by ALAD, it will dimerize to form a symmetrical pyrazine.

Single turnover experiments designed to study the order of substrate binding were carried out by Jordan and Seehra (17). Stoichiometric equivalents of [5-^{14}C]ALA and ALAD were rapidly mixed, with large amounts of unlabeled substrate being added after 100ms. Using this method it was possible to isolate [^{14}C]-PBG using 8 moles of labelled substrate per mole of enzyme (the enzyme being an octamer). The position of the radioactive label in the PBG was determined by degradation. The results of this analysis showed that only 17% of radioactivity could be found at the C-11 position using the human enzyme, whilst the figure for the bovine enzyme being slightly lower at 14%. These results provided strong evidence that ALA binds first at the P-site of ALAD, forming a Schiff base, and only then can the A site be filled with the second ALA molecule.

In order to investigate the proposed formation of the Schiff base, Jordan demonstrated that on reduction by $NaBH_4$ it is possible to trap the Schiff base, proving that it is a viable intermediate in enzymatic PBG formation.

Further evidence of this ordered binding of ALA to ALAD was provided by Spencer and Jordan (18) through investigations of the structural requirements of the A and P sites using a range of compounds related to ALA. The failure of 5-aminovaleric acid to inhibit ALAD and to compete with [^{14}C]ALAD for the A-site suggested an important role for the carbonyl group of ALA when binding to both the A and P sites. The lack of enzyme inhibition in the presence of 3-oxo-1-hexylamine also points to a critical role in A- and P-site binding for the carboxylic acid group of ALA. The inability of laevulinic acid (LA) to displace [^{14}C]ALA bound at the A-site shows LA to be a specific P site inhibitor and demonstrates an important role for the 5-amino group of ALA in binding to the A site. Further studies of the inhibition of ALAD by analogues of ALA and PBG include those carried out by Appleton (19), Cheung (20) and Senior (21)

There are two mechanistic schemes possible after the initial binding of ALA at the P site (scheme 1.5). The first is an aldol condensation to form a new C-C bond followed by the formation of a C-N bond (scheme 1.5b), to form the five membered ring. The second mechanism proceeds *via* the prior formation of the C-N bond (scheme1.5a), with the condensation occurring second.

1.4 TIM barrel folding motif

The α/β barrel is the most frequently occurring folding motif in protein structures. There are three main classes of α/β, the Rossman fold, the horseshoe fold and the TIM barrel. The TIM barrel, so called because it was first observed in the structure of triose phosphate isomerase (22), consists of a core of twisted parallel β-strands arranged close together with the surface α helices that connect the β-strands on the outside of the barrel.

The catalytic and substrate binding sites of TIM barrel proteins occur in loop regions that connect the carboxyl termini of the β-strands with the amino termini of the adjacent α helices. To date, the structures of over 20 proteins of separate catalytic function with the TIM barrel motif have been determined. In all structures there is a framework of over ~150 residues that give the barrel structure. Variations in primary structure generally occur *via* extended polypeptide termini and elongated loop regions.

In addition to the main barrel, some known TIM barrel structures have an amino terminal domain which contribute to the structural or functional properties of the enzyme. An example of this can be found in the structure of yeast flavocytochrome b_2, the extra domain of which incorporates a cytochrome (23).

The enzyme PRA-isomerase : IGP-synthase catalyses two reactions in the synthesis of tryptophan. The structure of this enzyme from *E. coli* reveals two barrels with one barrel displaying IGP-synthase activity and the other PRA-isomerase activity. Interestingly the same two reactions are catalysed by two separate enzymes in *Bacillus subtilis*. The two *Bacillus subtilis* enzymes have a high degree of homology to the corresponding *E. coli* barrels. It is thought that the *E. coli* enzyme

is an example of gene fusion although the obvious functional advantage to this structure is unknown (24).

1.5 X-Ray structure of yeast ALAD

The first X-ray structure of an ALAD to be solved was that of yeast (25). The structure was solved to a resolution of 2.3Å by a MAD experiment using a selenomethionyl substituted enzyme. The crystals grew in space group I422 with cell dimensions of a = b = 102.8Å, c = 167.9Å with one monomer per asymmetric unit and with an estimated solvent content of 59%.

1.5.1 Yeast ALAD monomer structure

The tertiary structure of the yeast ALAD monomer reveals an $(\alpha/\beta)_8$ or TIM barrel. The barrel is formed by an eight membered β-sheet surrounded by eight α-helices, with the loop regions between the α and β segments being modified to include extra secondary structure elements. A good example of these modifications is in the active site area of the enzyme, which can be found within elaborated loops connecting the C-terminal ends of the barrel. There is also an extended arm region at the N-terminal formed by residues 1-39 that is highly important in quaternary interactions.

1.5.2 Yeast ALAD dimer structure

The strongest quaternary contacts in the ALAD octamer are those between monomers interacting to form dimers. The monomers associate about a two-fold axis with their arms round each other to form dimers. The $(\alpha/\beta)_8$ domains of each subunit interact extensively around the two-fold axis positioning the barrel axis at right angles to each other.

1.5.3 Octamer structure

The dimers in the octamer interact identically with each other, with each active site orientated directly towards the solvent region. The dimers can be viewed as being roughly ellipsoid with their long axes slightly tilted with respect to the central 4-fold

surface of each dimer interacting with the N-terminal end of a β-barrel in a neighbouring dimer, with each dimer interacting identically with the others. The final octamer has dimensions of 104Å x 104Å x 94Å with a large solvent filled channel running through the centre.

1.5.4 The active site of yeast ALAD

The active site of yeast ALAD is dominated by lysine residues 210 and 263 which are located on two adjacent β-strands of the TIM barrel. Lys 263 is known to form a Schiff base with the P-site ALA and is found in a more hydrophobic region than Lys 210. Also found at the active site is a zinc ion, co-ordinated with tetrahedral geometry to three highly conserved cysteine residues (133,135,143) with a hydroxyl ligand position pointed towards the active site lysines. This zinc ion has been shown to be essential in A-site substrate binding and may be involved in co-ordinating the 4-keto group of A-site ALA.

1.6 X-Ray structure of *E. coli* ALA dehydratase

Once the structure of yeast ALAD had been solved the structure of *E .coli* ALAD in complex with LA was quickly determined by molecular replacement (26). The *E. coli* structure was also solved in space group I422 with unit cell dimensions of a = b = 126.7 Å c = 141.6 Å.

1.6.1 *E. coli* ALAD monomer structure

The *E. coli* monomer consists of 11 right handed α-helical segments and 11 right handed β-strands. There is an extended arm like structure pointing away from the compact $(\alpha/\beta)_8$ region or TIM-barrel. This arm is formed by the first 30 residues in the protein and is structurally important in forming quaternary contacts.

The TIM-barrel itself is also formed by an 8-membered cylindrical β-sheet surrounded by 8 α-helices. The *E. coli* ALAD TIM barrel is formed by the following α-helices and β-sheets (N- to C- terminus): β1, α2, β4,α3, β5,α4, β6,α5, β7,α6, β8,α7, β9,α9, β10,α10. Like the yeast ALAD, the active site is located in an opening formed by the loops connecting the C-terminal ends of the parallel β-strands in the barrel with their ensuing α-helical segments. At the exposed end of the β-

barrel, the loop regions between the α and β segments are elaborated extensively with the insertion of several extra secondary structure elements. In one such example a loop of residues (197-220) covering the active site, includes a region of α-helix involving residues 202-208. There is also an α-helix situated between β9 and α10 which is important for quaternary interactions.

The active site itself is formed mainly by the loop regions at the base of the β-barrel. The residues that comprise the active site are highly conserved among ALADs as are those that form the active site loop residues (197-220) (fig. 1.1).

1.6.2 *E. coli* ALAD dimer structure

The dimer is the most closely associated structure within in the octamer. In order to form the dimer unit, the extended N-terminal arm region formed by the first 30 residues of the monomer partly wraps around the TIM-barrel domain of another monomer with the inner surfaces of the arm interacting with helices α4 and α5 in the opposing monomer, in this respect there are strong similarities with the yeast ALAD. The helical regions (between residues 8-12 and 15-21) of the arm interact with helix α6 of the adjacent subunit in the dimer. Several salt bridges are made between arm and barrel of respective monomers e.g. Arg 12 with Glu 232 (in α6). There is also an interaction between Arg 18 and Lys 217 of both monomers which is mediated by a sulphate ion. Lys 217 is positioned close to Arg 216 which is thought to be involved in interactions with the carboxyl group of A-site ALA, suggesting an active site stabilising role for dibasic anions such as sulphate and phosphate. The two monomers are related by an intervening two-fold axis that positions the active sites roughly perpendicular to each other.

As well as barrel-arm contacts within the dimer there are several barrel-barrel interactions, with the principal ones involving helical segments α7 (252-263), α8 (273-284) and α9 (290-304). Helix α8 is roughly perpendicular to the 2-fold axis of the dimer and pairs in an antiparallel manner with the equivalent helix in the neighbouring monomer, as does helix α7, residues from which form several inter and intra-molecular electrostatic interactions (fig. 1.2).

1.6.3 *E. coli* ALAD octamer structure

Each of the four dimers that comprise the final octamer of the complete ALAD molecule can be roughly viewed as being ellipsoid in shape with their long axes inclined slightly with respect to the 4-fold symmetry axis. The dimers within the octamer interact with each other in an identical manner, principally being mediated by the arm regions on the surface of each dimer. In the dimer-dimer contacts, the end of one β-barrel is effectively capped by the helical segments (residues 8-12 and 15-21) of the arm region of a neighbouring subunit. This leaves the active site, contained at the other end of the barrel, exposed on the outside of the protein. These barrel capping interactions involve the formation of two salt bridges with the neighbouring subunit and also a sulphate anion positioned between Lys 13 of the N-terminal arm and Lys 107 at the base of the adjacent barrel (fig. 1.3).

1.6.4 The active site of *E. coli* ALAD and the binding of LA

As previously mentioned, the active site of ALAD is located at the end of the β-barrel that is exposed on the surface of the octamer. The loops at the end of this barrel form a cavity with the two active site lysines (195 and 247) at its base. Lys 247 is followed by an invariant proline (248) which has been shown to adopt the *cis*-conformation in the yeast ALAD structures. There is a hydrophobic pocket formed at the base of the cavity by the side chains of tyrosines 192, 201, 270

Fig 1.1 *E. coli* ALAD monomer. The active site of ALAD is positioned in a pronounced cavity located in the TIM barrel segment of the molecule. Key active site residues Lys 195 and 247 are highlighted. The green spheres indicated the locations of the zinc binding sites.

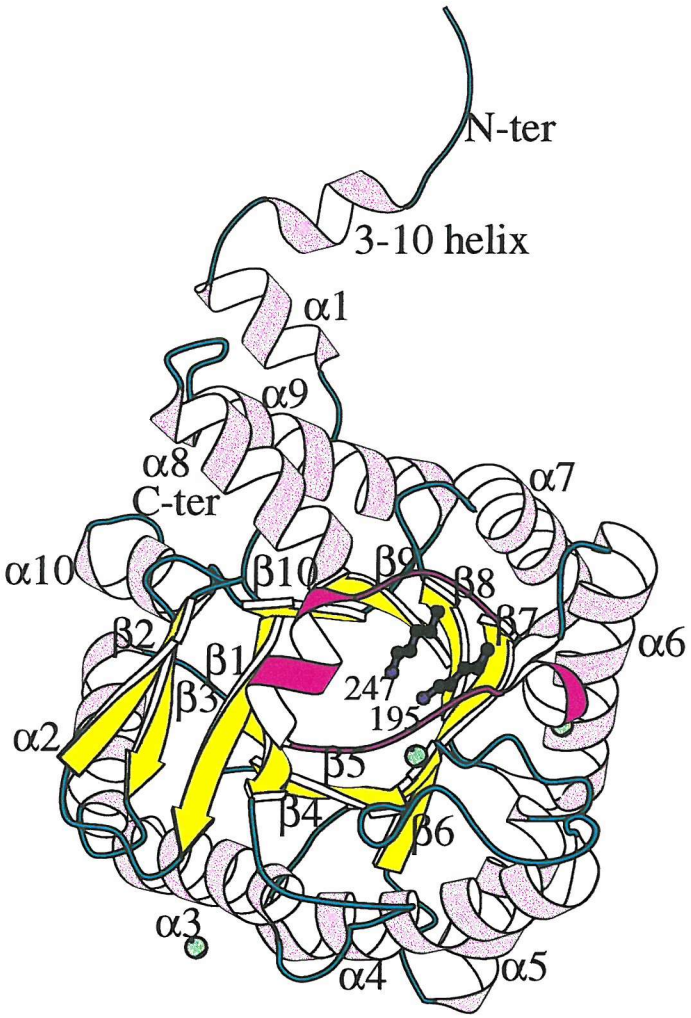
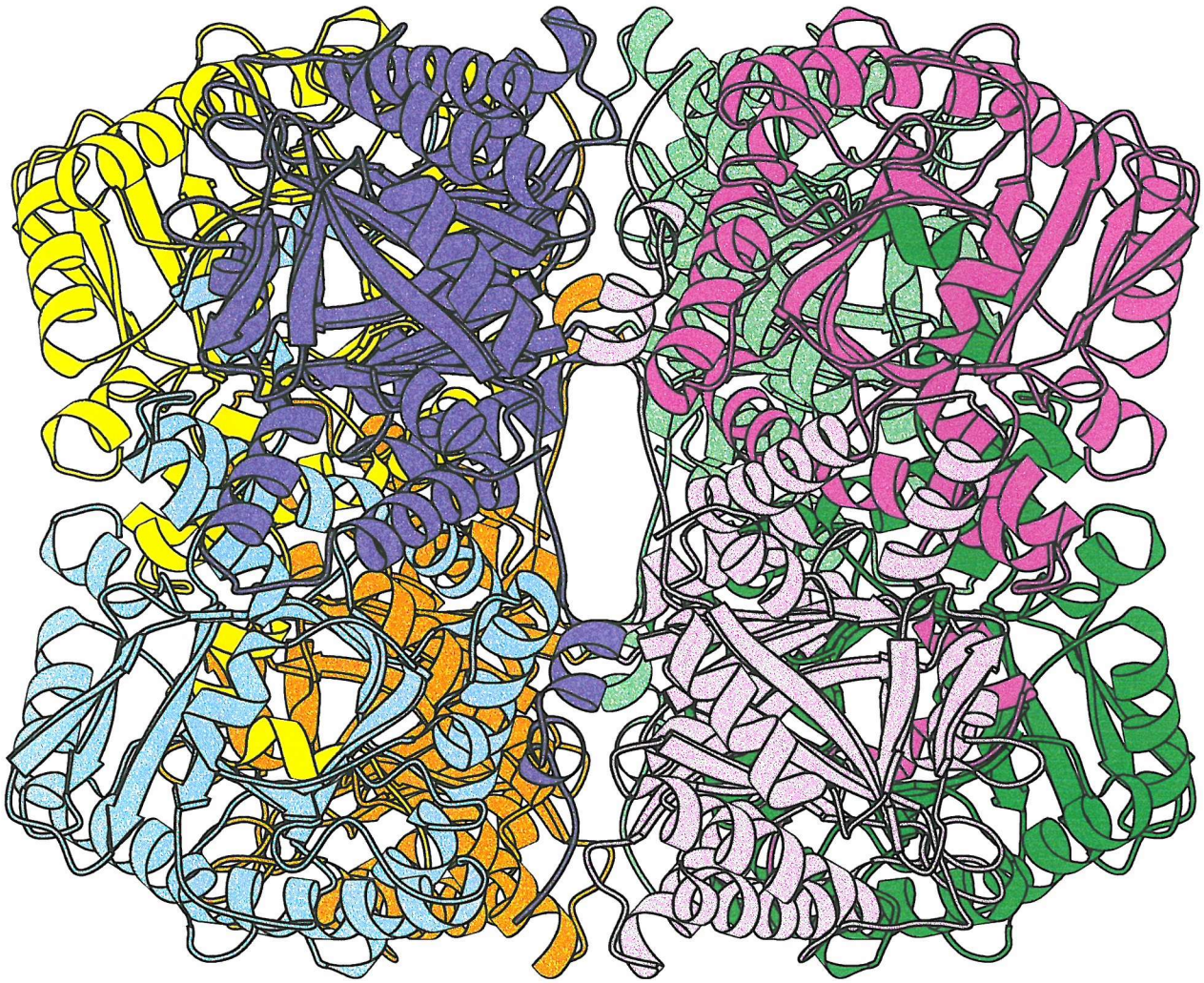


Fig 1.2 *E. coli* ALAD dimer. The dimer is the most closely associated subunit in the octamer. The N-terminal arm of each monomer in the dimer wraps around the barrel domain of its neighbouring monomer.

Fig 1.3 *E.coli* ALAD octamer. Four dimers associate along their long axis to form the octamer all interacting with each other in an identical manner.. The active site of each monomer is positioned at the surface of the octamer.



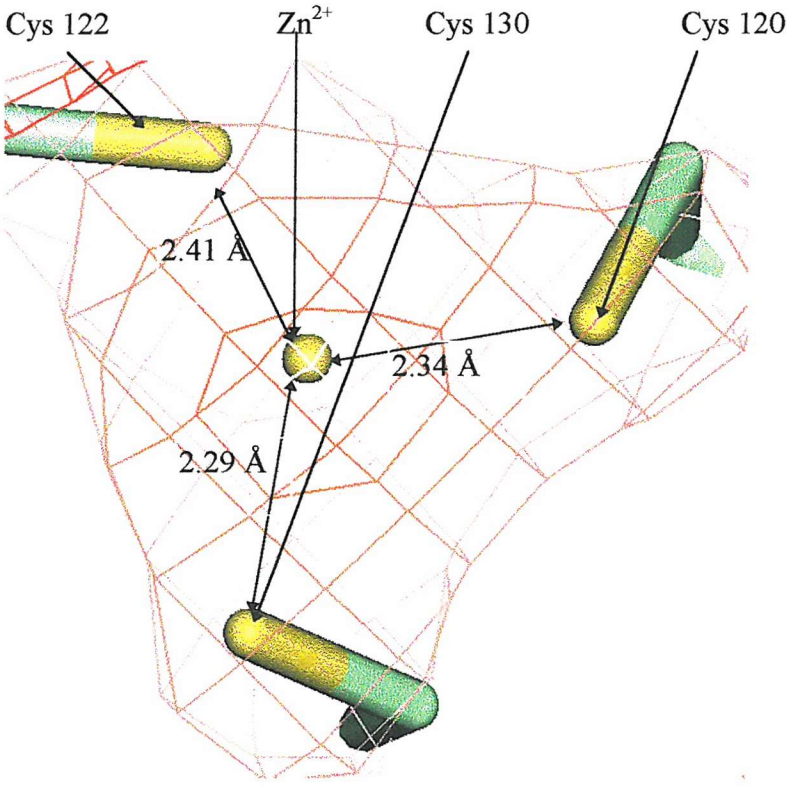
and other conserved hydrophobic residues such as Val 272 and Gln 220 with Lys 247 being in a more hydrophobic environment than nearby Lys 195. The proximity of the two amino groups at the end of the Lys side chains suggests that they may be sharing a positive charge, with Lys 195 more likely to be protonated as its environment is more polar. The effect of a positive charge on Lys 195 would be to lower the pK_a of Lys 247 which would also increase its ability to act as a nucleophile when forming a Schiff base with substrate P-site ALA.

Laevulinic acid (LA) is clearly visible at the active site of the *E.coli* ALAD structure and is bound to the P-site of the enzyme *via* a Schiff base formed with Lys 247. The carboxyl group of the LA forms hydrogen bonds with the side chains of invariant residues, Tyr 312 and Ser 273. The more hydrophobic methylene chain of the LA molecule interacts with several conserved residues in the active site especially Phe 204 which is part of the loop that covers the active site. Significantly, the active site loop is visible in the *E. coli* ALAD-LA complex but is poorly defined in the native yeast structure. This is thought to be due to the strong interaction between Phe 204 and the P-site LA molecule.

1.6.5 Zinc binding sites of *E. coli* ALAD

There are two zinc binding sites present in the *E. coli* ALAD monomer. The main catalytic zinc is bound by three cysteine residues (120, 122 and 130) located on the loop connecting β5 and α4. These residues are highly conserved among zinc-dependent enzymes whereas magnesium dependent enzymes utilise aspartates to co-ordinate their metal ligand. The three cysteine side chains co-ordinate the zinc with tetrahedral geometry, with a solvent ligand forming the fourth ligand orientated towards the two active site lysines. In close proximity to this zinc ion is the side chain of the invariant serine residue Ser 165 which is itself hydrogen bonded to a conserved aspartate residue. These two residues interact with the zinc ion through hydrogen bonded water molecules and are thought to be involved in substrate binding or catalysis (fig 1.4), as discussed later.

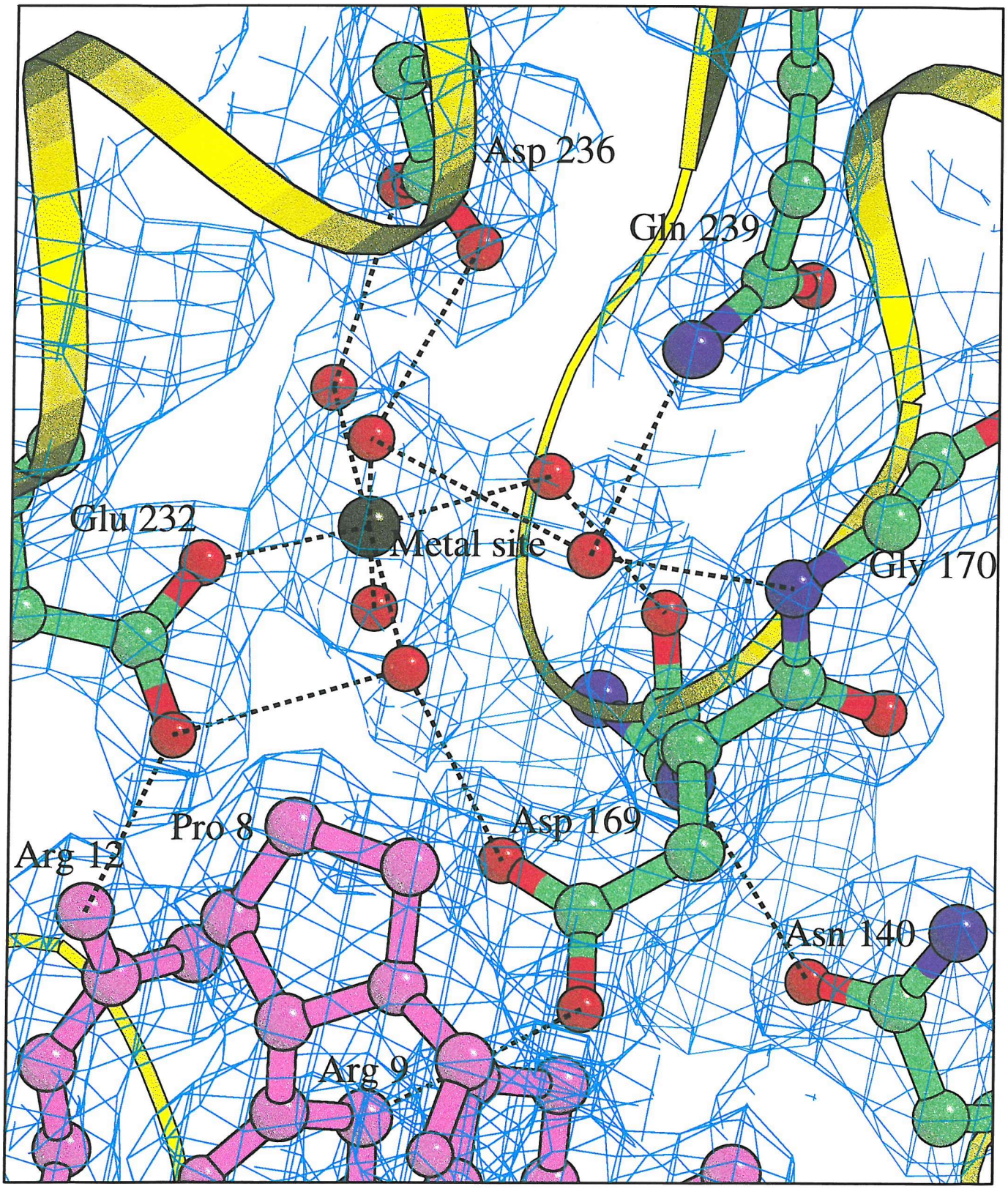
Fig 1.4. Active site zinc of *E. coli* ALAD showing the catalytic zinc ion co-cordinated by 3 conserved Cys residues.



1.6.6 Secondary zinc binding site of *E. coli* ALAD

In addition to the zinc ion bound at the catalytic site, there is a second zinc ion that seems to play an important structural role in the quaternary structure of the protein. It is located in a water filled pocket which lies between the TIM barrel and the N-terminal arm of the neighbouring subunit in the dimer. The ion is co-ordinated octahedrally by one side chain oxygen of Glu 232 and five solvent molecules which are in turn hydrogen bonded to surrounding side chain atoms of Asp 169, Gln 171 and Asp 236 (fig. 1.5). The presence of a second metal binding side is in agreement with previous experimental data (13).

Fig 1.5. The second zinc binding site of *E. coli* ALAD



1.7 Structure of the Mg^{2+} dependant ALAD from *Pseudomonas aeruginosa*

The structure of Mg^{2+} dependant ALAD from *Pseudomonas aeruginosa* has recently been solved to a resolution of 1.67 Å (27). This enzyme is of great interest as it is the first reported structure of a Mg^{2+} dependent ALAD and also shows evidence of half site activity.

The enzyme was co-crystallised with LA giving crystals growing in space group $P42_12$ with a = b = 128.3Å and c = 86.2Å and with a solvent content of 50%. The structure was solved by molecular replacement using the native yeast ALAD as the search model.

The structure of *Pseudomonas aeruginosa* ALAD reveals a homo-octamer, similar to that of other ALAD structures. However, in contrast to previously solved structures, there are two monomers in the asymmetric unit. The monomer itself consists of a TIM barrel structure with extended N-terminal arm, with barrel dimensions of 45 Å x 50 Å x 55 Å. The active site is contained in the barrel of the monomer as with yeast and *E. coli* ALADs, with the N-terminus arm involved in secondary and quaternary structure interactions.

The two monomers in the asymmetric unit (A and B) form a dimer, four of which assemble in the crystallographic 4-fold axis to form a cylindrical D_4 octamer with dimensions 133 Å x 97 Å. The active sites of all monomers in the octamer are located at the surface of the molecule.

Although there is a high degree of overall structural similarity between *Pseudomonas aeruginosa* ALAD and other ALAD structures there are some notable differences. Monomers A and B show different structural properties, most notably an increased flexibility in the active site and differences in metal ion binding.

The active site of the monomer is located in a cleft 20 Å deep with dimensions of 10 x 10 x 5 $Å^3$. In monomer A, the active site is shielded by a lid of two short α

helices, with the active site in monomer B being much wider and more accessible to solvent due to flexibility in the lid region (indicated by lack of density). The lid region of the molecule has been shown to be catalytically essential and contains many conserved residues. This leads to monomer A being designated the "closed monomer" and B the "open monomer."

The active site can be further analysed when considering the position of the LA inhibitor molecule in the active site. LA binds *via* a Schiff base to an invariant lysine at the P-side of the active site in an identical manner to that seen in the yeast and *E. coli* structures. A magnesium ion cannot be seen at the predicted metal binding position at the active site which is comprised of two aspartate residues and it is thought the Mg^{2+} is only bound at the active site upon binding of substrate ALA.

A second Mg^{2+} binding site, away from the actual active site, is more defined and is of great interest. In monomer A the site is occupied and located in a pocket at the surface of the TIM barrel. The Mg^{2+} ion interacts with numerous conserved amino acids on the surface of the barrel and the N-terminal arm of monomer B. The surrounding area of the metal binding site shows a high level of negatively charged amino acids that are involved in salt bridge interactions with oppositely charged residues in the N-terminal arm of the other monomer in the dimer.

At the Mg^{2+} binding site of monomer B, no metal ion is present, leading to the conclusion that the occupancy of this site has an affect on whether the monomer is open or closed. Notable conformation differences between monomers in the asymmetric unit can be found in the region between the active site and this second metal binding site. In the closed monomer conserved arginine residue 181 is involved in interactions with the bound Mg^{2+}. However, in molecule B where Mg^{2+} is not present, the residue adopts a different conformation. This causes the residue to interact *via* a salt bridge with the side chain of aspartate 139 pulling its side chain away from the active site. Asp 139 is thought to be involved in Mg^{2+} binding at the active site. In the closed A monomer, the side chain of Asp 139 is pointing towards the A-side of the active site and interacts with invariant lysine residue 229 belonging to the ordered loop region.

From this evidence it is suggested that the side chain of Arg 181 is able to switch the monomer on or off by either decoupling from the Mg^{2+} binding site and disrupting the active site, or interacting with the Mg^{2+} on the outer metal site, leaving the active site and lid in a well ordered and stable conformation. The importance of the residues proposed to take part in this switch mechanism have been demonstrated by site directed mutagenesis.(27)

1.8 ALAD also functions as a proteasome inhibitor

Recent research (28) revealed that the human proteasome inhibitor (CF-2) is identical to human ALAD.. The conclusion that the two proteins are identical was reached after the following observations:

1) Both proteins share a common sequence for the first 14 N-terminal amino acids.
2) The proteins showed identical migration on native and SDS polyacrylamide gel electrophoresis.
3) The proteins share identical isoelectric points of pH 7.1 .
4) Cross reactivity of specific polyclonal antibodies was observed for both proteins..
5) Either proteins showed similar activity when substituted to perform the function of the other.
6) Purified recombinant ALAD was able to perform both enzymes' functions.

1.8.1 The proteosome

The proteosome is the central centre of protein degradation in both the cytosol and the nucleus. Its main biological functions are to remove abnormal, misfolded or improperly assembled proteins, transcription factors and oncoproteins. It also has an important role in the cellular immune response, its function being to generate antigenic peptides presented by major histocompatibility complex-1 molecules.

The proteosome core is a barrel shaped 20S (Mr 700,000) particle also known as CF-3 or the 20S proteosome. In eukaryotes a regulatory cap is complexed at each end of the barrel. Each cap is a 19S particle consisting of at least 18 proteins. When a 20S

core particle is associated with two 19S caps the complex is known as a 26S proteosome. Both the 20S and 26S proteosome have catalytic activity. The 20S proteosome is able to degrade unfolded proteins in a non-ATP dependant reaction, however when incorporated into the 26S proteosome the 20S particle forms the catalytic core of a complex able to perform the functions outlined above.

1.8.2 Role of ALAD/CF-2 as the proteosome inhibitor within the 26S proteosome

The CF-2 factor has been found to exist in the 26S proteosome bound to ubiquitin (a highly conserved ubiquitous protein) as a conjugate (29). The ubiquitinylated ALAD/CF-2 component is thought to regulate the selective proteolysis of ubiquitinated proteins performed by the 26S proteosome. Etlinger suggests that when the component is ubiquitinylated it activates the proteasome to allow it actively to cleave substrate (28). Notably, ALAD does not have to be catalytically active to carry out its role as a CF-2 protein.

This very different function for ALAD is an example of gene sharing, where the product of one gene fulfils two distinct roles.

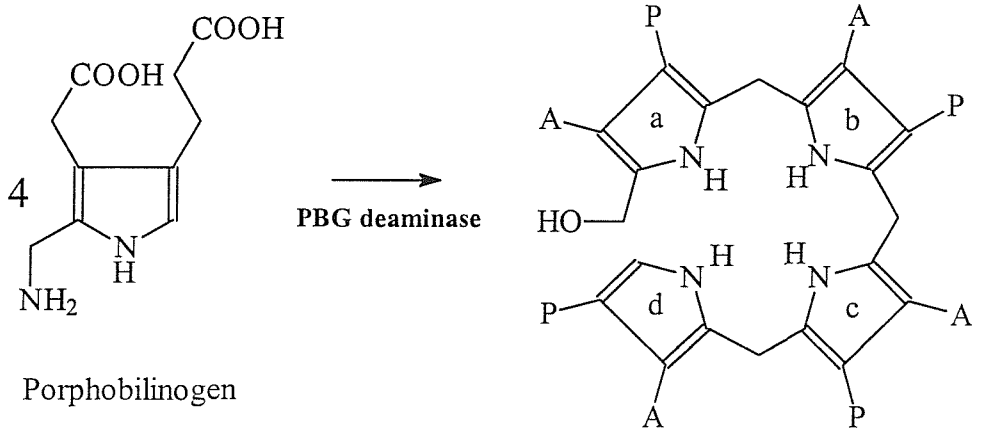
1.9 Human ALAD mutations

The human ALAD gene is localised on chromosome 9q34 with two common alleles (ALAD1 and ALAD2). It has been demonstrated that there are several naturally occurring mutations which occur in human ALAD that result in acute porphyria, also known as Doss porphyria (30). A less harmful mutation (K95N) has also been described that is present in over 10% of the human population. Inspection of the yeast crystal structure suggests that this surface mutation would not unduly affect enzyme function or folding. Other harmful mutations include G133R, R240W, A247T and V275M all of which are associated with Doss porphyria. With the exception of G133R, which probably affects zinc binding at the active site, these mutations appear to have structural effects on the ALAD molecule. The result of these mutations is to produce a dysfunctional enzyme which cannot synthesise PBG at the normal rate and which causes the accumulation of ALA. ALA is a close analogue of the neurotransmitter GABA and is thought to cause neurological problems associated with this type of porphyria.

1.10 Porphobilinogen deaminase

Porphobilinogen deaminase catalyses the polymerisation of four molecules of PBG to yield a linear chain called preuroporphyrinogen (scheme 1.6), an extremely unstable 1-hydroxy-methylbilane molecule with a half life of 4min at pH 8.4 (34). The four rings of preuroporphyrinogen are assembled in a stepwise fashion while covalently attached to an enzyme bound dipyrromethane cofactor, itself derived from two molecules of PBG, which is self assembled by the apoprotien and is attached *via* a covalent thioester linkage to an invariant cysteine at the active site (31). Once assembled, the preuroporphyrinogen is released by hydration, leaving the cofactor still attached to the enzyme.

Scheme 1.6. Formation of preuroporphyrinogen



Porphobilinogen

Preuroporphyrinogen

1.10.1 Structure of *E. coli* PBG deaminase

The 3 dimensional structure of porphobilinogen deaminase from *E. coli* has been solved to a resolution of 1.7 Å (32). The dipyrromethane cofactor is attached to cysteine 242 on a loop of domain III and positioned at the mouth of a deep cavity formed between domains I and II. The cavity is lined with small amino acid side chains, providing space for the polymerisation reaction to proceed protected from solvent. The A and P side chains of each of the cofactor rings interact electrostatically with arginine, threonine and serine residues in the cavity, with both

pyrrole nitrogens from the cofactor being hydrogen bonded to aspartate 84 from domain I, an invariant residue throughout all deaminase sequences (fig 1.6).

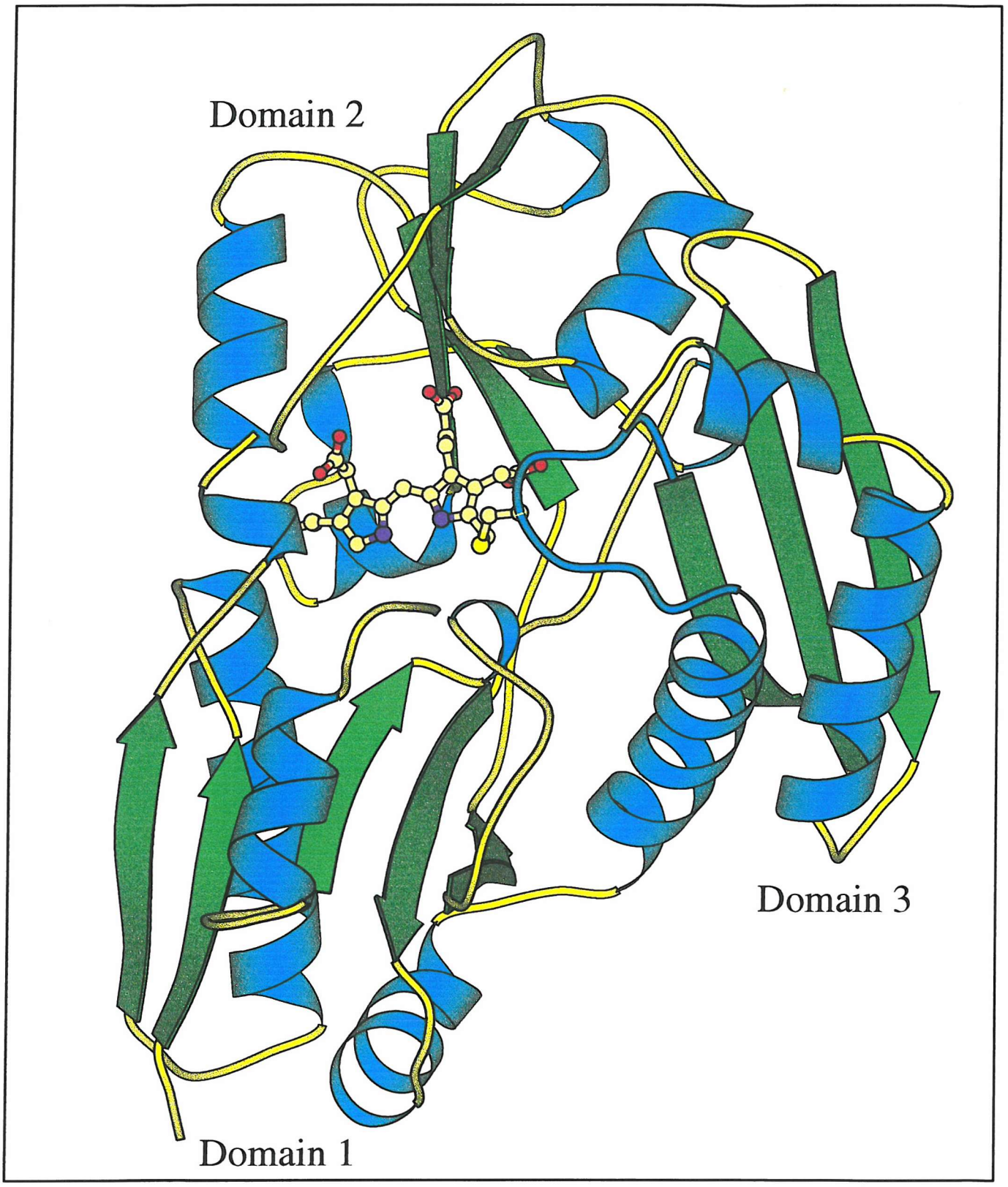
Although it is still not known exactly how the enzyme constructs the preuroporphyrinogen molecules, it is thought that domain adjustments may be involved, a theory currently under investigation by crystallisation of ES complexes from mutants (33). Other attempts to study the mechanism of the reaction have used substrate analogues (34).

The elucidation of the *E. coli* PBGD structure and the high degree of sequence homology with respect to the human enzyme has allowed a reliable model of the human structure to be constructed and an insight into the molecular basis of mutations that cause acute intermittent porphyria to be investigated (35).

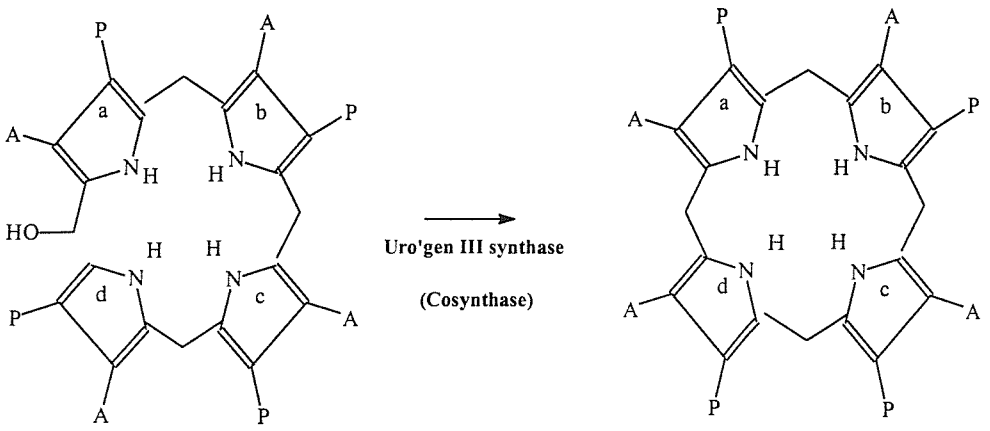
1.11 Biosynthesis of uroporphyrinogen III from preuroporphyrinogen

Uroporphyrinogen III is the universal cyclic tetrapyrrole from which all linear and cyclic tetrapyrroles, such as haem and chlorophyll are derived. The synthesis of uroporphyrinogen III from PBG requires the participation of both PBG deaminase and uroporphyrinogen III synthase. In the absence of uroporphyrinogen III synthase the unstable preuroporphyrinogen produced by PBG deaminase cyclises to give uroporphyrinogen I. However the enzymatic cyclisation catalysed by uroporphyrinogen III synthase causes the fourth pyrrole ring of preuroporphrinogen to rearrange, the molecule then cyclising to give uroporphyrinogen III (scheme 1.7).

Fig 1.6 Structure of *E. coli* porphobilinogen deaminase. The three domains comprising the enzyme have been highlighted, as has the dipyrromethane cofactor.



scheme 1.7. Uroporphyrinogen III formation.

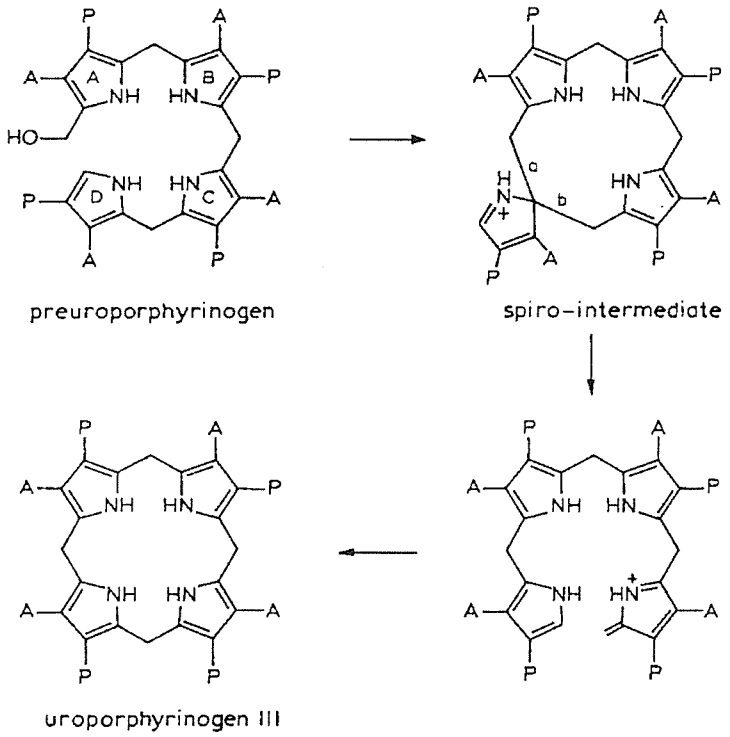


Preuroporphyrinogen

Uroporphyrinogen III

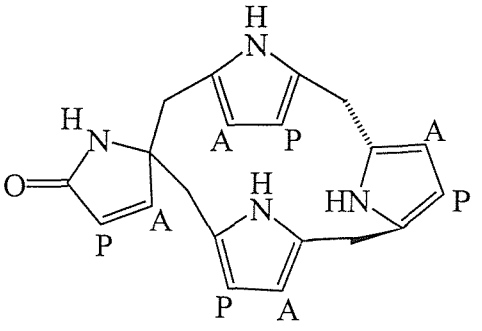
The mechanism most favoured for the for the formation of uroporphrinogen III from its preuroporphyrinogen precursor proceeds *via* a spiro-intermediate (scheme 1.8) in order to facilitate the inversion of ring D (36).

Scheme 1.8. Mechanism for the biosynthesis of uroporphyrinogen III from preuroporphyrinogen. The mechanism proceeds *via* a spiro intermediate to invert the D pyrrole ring of the forming molecule.



In order to investigate whether the proposed spiro intermediate would be too strained to be viable, analogues of the intermediate containing the same macrocyclic ring were prepared (37). Once the synthesis of the macrocyclic ring had been achieved inhibitor molecules with the natural acetate and propionate side chains were prepared as potential enzyme inhibitors. The most potent inhibitor of uroporphrinogen III synthase (scheme 1.9) bound strongly to the enzyme with an inhibition constant (K_i) of 1-2 μM suggesting that the spiro-intermediate is formed during the synthesis of uroporphrinogen III.

Scheme 1.9 Synthesised spiro intermediate analogue. The analogue binds strongly to uroporphrinogen synthase with a K_i of 1-2 μM.



E. coli uroporphyrinogen III synthase has been isolated and been shown to have a Mr. of 29k (38). Sequence comparison between human and *E. coli* proteins show low levels of homology.

1.12 Biosynthesis of haem

Once formed, uroporphyrinogen III acts as the precursor for a wide range of tetrapyrrole containing molecules. The initial step in haem biosynthesis from uroporphyrinogen III involves the decarboxylation of the acetic acid side chains to methyl groups to give coproporphyrinogen III by uroporphyrinogen decarboxylase. The three dimensional structure of uroporphyrinogen decarboxylase has recently been solved (39)

1.12.1 X-Ray structure of uroporphyrinogen decarboxylase

The three dimensional X-ray structure of recombinant human uroporphyrinogen decarboxylase has been solved to a resolution of 1.6 Å (39). The structure reveals that the 40.8k protein forms a single domain containing an $(\alpha/\beta)_8$ barrel. Loops at the end of C-terminal barrel strands form a deep cleft containing the active site. The cleft contains many of the residues that are conserved across species. The dimensions of the cleft are about 15 Å x 15 Å x 7 Å which is suitable for insertion of a porphyrinogen into the active site with most of the substrate shielded from solvent. The cleft contains a high proportion of positively charged side chains which is consistent with the high negative charge on the substrate. These residues include Arg37, Arg41 and His339 which are thought to be involved in substrate binding. Other residues, such as Asp86, Tyr164 and Ser219 may have binding or catalytic roles. Interestingly, mutation of His339 to Asn results in full activity with the initial 8-COOH substrate, but is only weakly active for additional decarboxylations of 7- , 6-, and 5-COOH intermediates. It is thought that the position of this residue at the opening of the active site cleft may allow it to function in the orientation of the partially decarboxylated substrates in the active sites.

In solution uroporphyrinogen decarboxylase is a dimer which is also formed in the crystal by operation of a crystallographic two fold axis on the single molecule of the asymmetric unit.. The active site cleft of one monomer is adjacent to that of its neighbour in the dimer. This creates a single extended cleft which is large enough to accommodate two substrate molecules in close proximity, or to allow intermediates possibly to shuttle between monomers.

1.13 Coproporphyrinogen III oxidase

Coproporphyrinogen III oxidase catalyses the oxidative decarboxylation of the propionic acid side chains of rings A and B of coproporphyrinogen III to the vinyl groups of protoporphyrinogen IX. Two classes of coproporphyrinogen III oxidase exist, with one class being found in aerobic organisms and another in anaerobic, with some organisms having both types.

1.13.1 Anaerobic coproporphyrinogen III oxidase

The first evidence for anaerobic coproporphyrinogen III oxidase came from studies on *Rhodobacter spharoides*. A coproporphyrinogen III oxidase was isolated from this organism which catalysed the transformation of coproporphyrinogen III into protoporphyrinogen IX in the presence of NADP, ATP and L-methionine but did not require oxygen (40). Although progress on this form of the enzyme has proved difficult, other organisms have been shown to posses the anaerobic form e.g. *Salmonella Typhimurium* (41).

1.13.2 Aerobic coproporphyrinogen III oxidase

Aerobic coproporphyrinogen III oxidase has been purified from a variety of sources such as rat and bovine liver and yeast and appears to be a dimer. The enzyme is oxygen dependent and has no requirement for a chromophore or metal ion and exists as a soluble mitochondrial protein. (42).

1.14 Protoporphyrinogen IX oxidase

Protoporphyrinogen IX oxidase catalyses the oxidation of protoporphyrinogen IX to protoporphrin IX. This process establishes the planar porphyrin ring that is able to chelate iron in the final step of haem synthesis.

Protoporphyrinogen IX oxidase was first isolated in yeast (43) with protoporphyrinogen IX being the only major substrate. In eukaryotes, the enzyme is located in the mitochondria with a K_m for the substrate of 5-10 μM. Enzyme purified from a bovine source was shown to be a monomer with an Mr. of 65k.

1.15 Ferrochelatase

The final step in the haem biosynthesis pathway, the insertion of ferrous iron into protoporphyrin IX, is catalysed by ferrochelatase. Ferrochelatase has been isolated from many sources including rat liver (44) yeast (45) and bacterial sources (46).

Ferrochelatases are monomeric proteins (Mr=36k-40k) with similar catalytic properties in different organisms, but with varying cellular locations. In mammalian

and yeast cells it is located in the inner mitochondrial membrane while in bacteria it is found in the cytoplasm or associated with the cytoplasmic membrane. It is thought that the membrane association might have evolved due to the low solubility of substrate and product of the reaction in aqueous solutions at neutral pH.

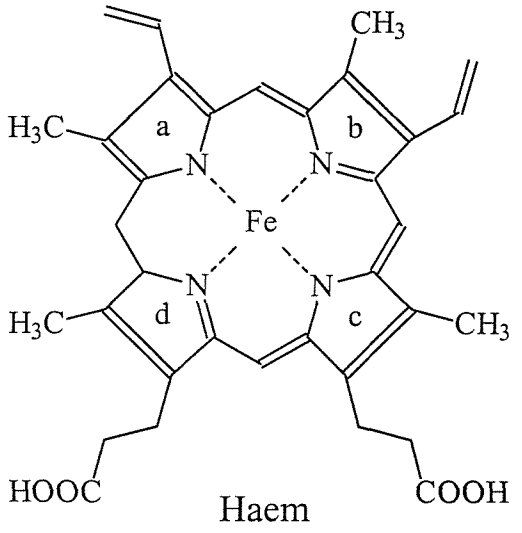
The overall sequence identity between ferrochelatases is around 12% with ~20 invariant residues, some of which have been shown to be catalytically important by site-directed mutagenesis.

1.15.1 X-Ray structure of ferrochelatase

The three dimensional structure of *Bacilius subtillis* ferrochelatase at 1.9 Å resolution, solved by multiple isomorphous replacement, has recently been determined (47). The structure reveals the 310 amino acid protein is folded into two similar domains, each with a four stranded parallel β-sheet flanked by α-helices. Structural elements from both domains form a cleft containing several amino acid residues that are conserved between organisms. It is thought that an invariant histidine in this cleft is involved in ferrous iron binding and that the porphyrin binds in this cleft to give the final haem molecule (scheme 1.10).

Haem can then be incorporated as a prosthetic group into a wide range of proteins that have important roles in areas such as oxygen binding, oxygen metabolism and electron transfer.

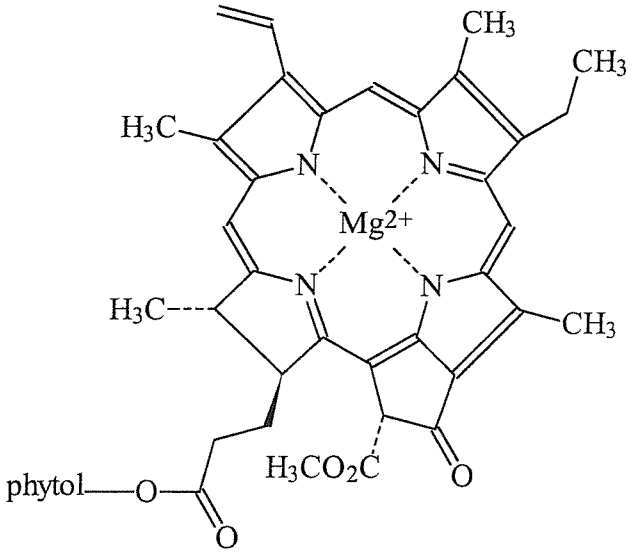
Scheme 1.10. Haem molecule



Haem

Other biologically important modified tetrapyrroles include sirohaem (the prosthetic group of sulphite and nitrite reductases), chlorophyll (scheme 1.11) and vitamin B_{12}.

Scheme 1.11 Chlorophyll molecule



Chlorophyll-*a*

Chapter 2: Materials and Methods

2.1 Chemicals

All chemicals were purchased from Sigma chemical company (Poole, UK) except for growth media (Difco, UK), and bacto agar (lab M, UK).

2.2. Buffers and solutions used in this study

2.2.1 Luria Broth (LB) medium for growth of *E. coli*

For 1 litre -

tryptone	10 g
yeast extract	5 g
NaCl	5 g

this was sterilised by autoclaving before use.

2.2.2 Modified Ehrlichs reagent (MER)

4-dimethylaminobenzaldehyde	1 g
70 % perchloric acid	8 ml
glacial acetic acid	42 ml

Once made, this is stored at 4°C and replaced daily.

2.2.3 Haemin

Solid hemin (30 mg) was dissolved in 0.2 M KOH (400 µl) by vortexing. AnalaR water (600 µl) and 1M Tris/HCl pH 7.8, (100 µl) were added, followed by ethylene glycol (4.8 ml). The mix was vortexed and centrifuged at 5000xg for 5 minutes at 4°C, and the insoluble pellet was discarded. The solution, at a final concentration of 7 mM haemin, was filtered through a 0.2 µm filter (Schleicher and Schuell, FP 030/3) and stored at 4°C until use

2.2.4 Polyacrylamide gel electrophoresis (PAGE)

Acrylamide stock:

acrylamide	38 g
N,N-methylene*bis*acrylamide	2 g

Water was added to a final volume of 10ml and the solution was stirred gently for 30 minutes. The solution was then filtered through a whatman 0.2 μm nylon membrane filter and stored at 4°C. Care was taken when handling this solution..

2.2.5 12% SDS-PAGE gel

Main gel:

Acrylamide stock	3.2 ml
10% (w/v) Ammonium persulphate (APS)	100 μl
1.5 M Tris/HCl pH 8.6	1.0 ml
N,N,N',N' Tetramethylethylendiamine (TEMED)	20 μl
Water	2.8 ml
10% (w/v) Sodium Dodecyl Sulphate (SDS)	80 μl

2.2.6 Stacking gel :

Acrylamide stock	1.0 ml
10%(w/v) APS	100 μl
0.5M Tris/HCl pH 6.8	1.0 ml
TEMED	20 μl
Water	3.0 ml
10%(w/v) SDS	50 μl

To prepare a gel, all the ingredients for the main gel, except 10%(w/v) APS and TEMED, were mixed before the gel plate apparatus was assembled. When the apparatus was ready, the APS and TEMED were added and the mix was cast into the gel plates. After at least ten minutes, when the main gel had set, the stacking gel ingredients were mixed without the APS and TEMED. These were added together and the stacking gel was cast with a comb to form the sample wells. When this also had set, the comb was removed and the gel was placed in the buffer reservoir, which was then filled with running buffer. Samples were supplemented by a 3:5 volume of 1.6 x disruption buffer and placed in boiling water for 1 minute. Protein (15μg) was applied to the gel and electrophoresis was carried out at 30 mA until the bromophenol blue dye front had reached the bottom of the gel. The gel was then

stained and the protein bands were visualised by destaining according to the experiment being performed.

2.2.7 Destain for SDS gels

Glacial acetic acid	70 ml
Methanol Add 0.25% brilliant blue R to make stain solution.	400 ml
Distilled water	530 ml

2.2.8 10 x Disruption buffer

SDS	2 ml of a 10% (W/V) solution
1M Tris/Hcl pH6.8	0.5 ml
Glycerol	0.6 ml
Bromophenol blue	0.01 g
Distilled water	6.4 ml

2.2.9 5 x Running buffer:

Tris base	9 g
Glycine	43.2 g
Sodium dodecyl sulphate	3g

The ingredients were dissolved in 600 ml distilled water, then stored at 4°C until used. To use, 60 ml was diluted to 300 ml with distilled water.

2.3 Preparation of *E. coli* ALAD mutants

E. coli ALAD mutants were kindly prepared by Dr. M. Sawar by the PCR method described by Innis *et al.*, (48). Mutant PCR DNA was restricted with *Eco*RI and *Bam*HI enzymes and ligated into *Eco*RI and *Bam*HI restricted pUC 19 vector.

2.3.1 Preparation of competent cells for transformation

A single colony of *E. coli* strain TB1 *hem* B^- was grown overnight at 37°C in 20ml LB medium with continuous shaking, with haemin added if necessary. LB medium (15ml) was inoculated with 150µl of this overnight culture and incubation

was continued until the OD_{600} = 0.3. Cells were harvested by centrifugation at 3000rpm at 4°C for 20 minutes, then resuspended in 10ml ice-cold 50mM calcium chloride, freshly diluted from a 1M stock. These cells were left on ice for one hour, then harvested as before, and suspended gently in 1ml ice-cold 50mM calcium chloride. Cells were either used immediately or a glycerol stock made from these and used within one week.

2.3.2 Transformation of competent cells with plasmid DNA

A 40 µl volume of pUC 19 plasmid DNA, was added to competent cells (200µl) and the mixture stored at 4°C for 5 minutes. The cells were heat-shocked for 2 minutes at 42°C, then incubated at 4°C for 5 minutes once again. LB medium (800µl), with hemin (7µM) if required, was added and the mix incubated with inversion at 37°C for 1 hour, before 100µl of the mixture was spread on LB agar plates supplemented with ampicillin and these were left to grow overnight. Plasmids held the gene for ampicillin resistance and thus cells which had taken up plasmid DNA could grow on the ampicillin plates. A control plate with untransformed wild type bacteria was also set up to check for contamination.

2.3.3 Bulk growth of transformed cells

Colonies of transformed bacteria were then selected individually and grown overnight in 10 ml of LB (containing ampicillin). These cultures were added to baffled flasks of sterile LB (500 ml) and grown overnight with shaking rate of 160 rpm in a New brunswick scientific innova 4330 incubator/shaker.

2.3.4 Harvesting of bulk growth

The culture, (4L in total) was then centrifuged at 5000 x g in a Beckman centrifuge fitted with a J2-20 rotor for 30 minutes and the resulting pellet cells frozen in 5 g portions for future use.

2.4 *E. coli* ALAD purification

2.4.1 Sonication.

Once thawed, the cells (6g wet weight) were resuspended in 5 ml of 50 mM potassium phosphate buffer, pH 6.0, containing 50 μM $ZnCl_2$ and 10 mM β–mercaptoethanol to which was added 0.2 mM PMSF to reduce any digestion of ALAD by proteases. The cells were then sonicated for fifteen 30 second bursts, with 60 second intervening cooling periods. At all times during sonication, the cells were jacketed in an ice water bath to keep the sonicate cool. The sonicated cells were then centrifuged for 20 minutes at 12000 rpm in a Beckman centrifuge using a J2-20 rotor. The supernatent was then removed and a 10 μl aliquot added to 990 μl of sonication buffer, detailed above, and the absorbance at 280 nm (A^{280}) was measured. The supernatant was then completely decanted and diluted with sonication buffer until the A^{280} of the above assay was 0.5.

2.4.2 Ammonium sulphate precipitation.

Ammonium sulphate was added to the enzyme solution to give 33% saturation (196g/litre). After allowing the protein to precipitate for 15 minutes, the precipitate was removed by centrifugation as above, and the pellet discarded. The supernatant subjected to a further ammonium sulphate fractionation (33% - 40% saturation) and was again centrifuged. The resulting pellet, containing the ALAD activity, was then resuspended in 1ml of 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol.

2.4.3 Gel filtration using Sephacryl HR 300

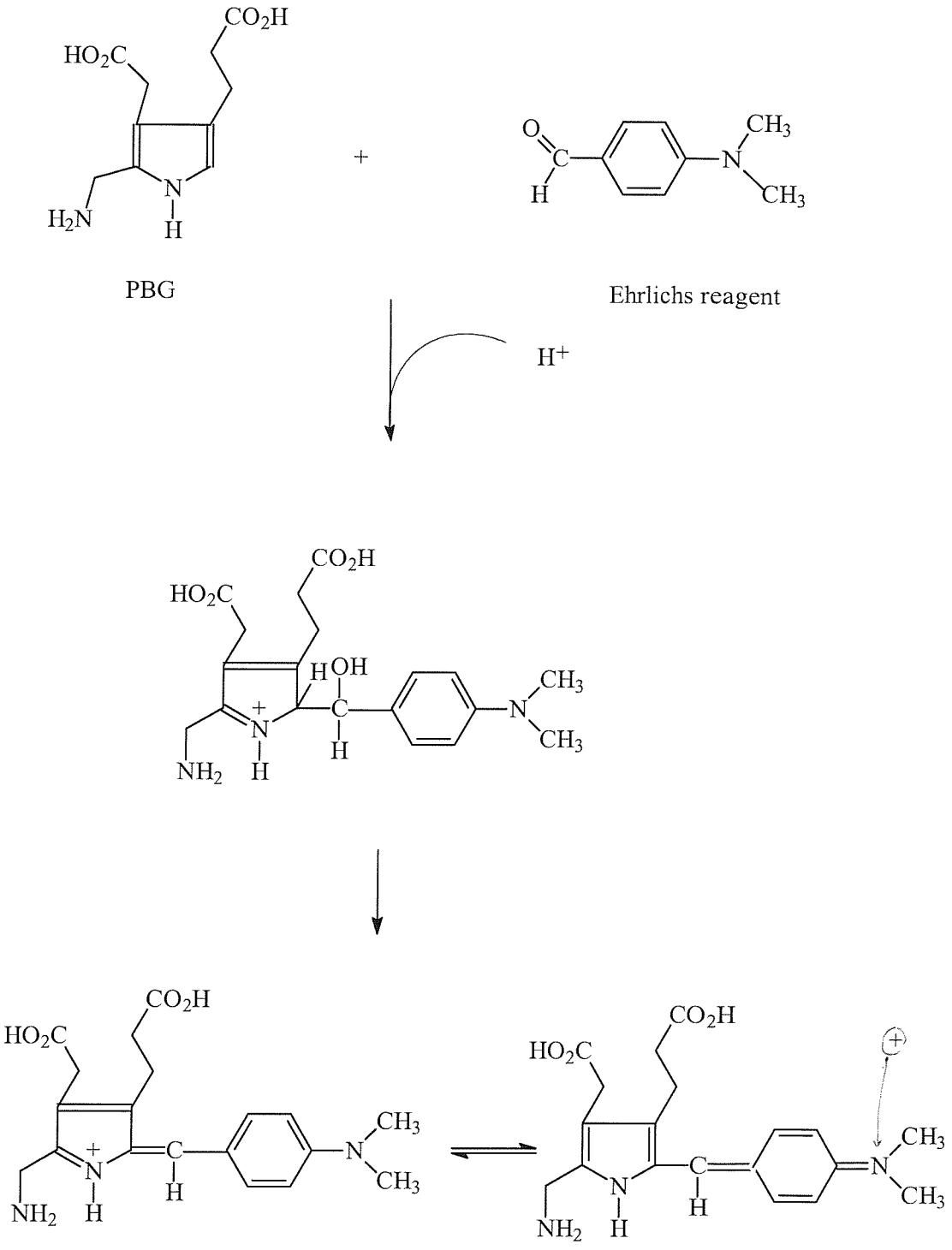
A 4cm x 50cm, 500 ml column (Sephacryl HR 300) was equilibrated with 5 column volumes of 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol at a flow rate of 20 ml/min. The resuspended pellet (1ml) was applied to the top of this column. Fractions of 5ml were collected after 200 ml (void volume of column) automatically ready for later assay. The enzyme eluted in fractions 30-40 (approximately).

2.4.4 ALAD assay

E. coli ALAD was assayed in 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol in a total volume of 500 μl. The reaction was initiated by the addition of ALA to give a final concentration of 5 mM. After incubation at 37°C for 3 minutes, the reaction was terminated by the addition of an equal volume of 10% trichloroacetic acid containing 0.1M $HgCl_2$ to precipitate protein and thiol. The solution was centrifuged, and an aliquot was removed and mixed with an equal volume of modified Ehrlichs reagent (49) (scheme 2.1). After 15 min. the absorbance at 555 nm was determined (ε_{555} = 60200 litres/mol^{-1}/cm^{-1}.). Using this information, and by measuring protein concentration at 280 nm, the specific activity of the ALAD can be obtained. SDS-PAGE gels were also run with all fraction samples to check the purity of the preparation. The specific activity of the ALAD obtained was 22.8 μmoles PBG produced/hour/mg (table 2.1) ALAD and the fractions estimated to be 90-95% pure. Rates for kinetic analysis were determined from three independant trials using three substrate concentrations higher and three lower than the K_m. The rates for each substrate concentration were determined in duplicate and data were analysed with EZ-Fit (Perrella scientific, Inc.) and a P.C.S. 166 computer to generate values of K_m and k_{cat}.

It is important to note that during kinetic analysis of *E. coli* ALAD mutants the time period of the assays and protein concentrations used were increased in order to compensate for low enzyme activity.

Scheme 2.1 Assay for PBG using modified Ehrlichs reagent

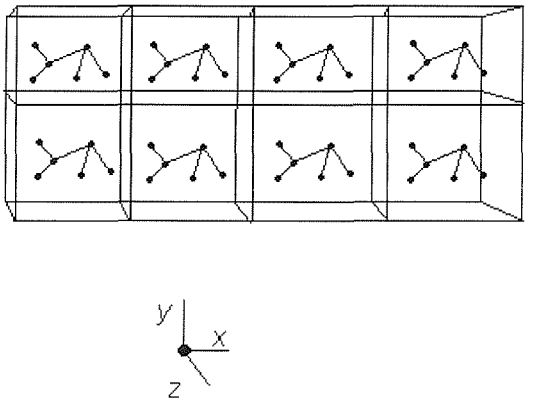


$A^{555\,nm}$

2.5 Protein crystallisation

Under certain circumstances, many molecular substances e.g. proteins, can solidify to form crystals. A crystal is an orderly three dimensional array of molecules held together by noncovalent interactions. (fig 2.1).

fig. 2.1 Eight unit cells in a crystalline lattice.



The figure has been divided into eight identical unit cells, the smallest and simplest volume element that is completely representative of the whole crystal. When the unit cell is repeated by translations that are parallel to its edges in three directions without rotating the unit, it makes the crystal lattice. The three directions in which the crystal is translated define the three crystal axes.

Elucidation of a protein structure by X-ray diffraction analysis requires three dimensional, large, individual, well ordered crystals. In the case of *E. coli* ALAD this was achieved using the vapour diffusion method of crystallisation *via* the "hanging drop" technique (fig 2.2). This method uses vapour diffusion between protein sample and a reservoir of crystallisation solution with a concentration of precipitant higher than that needed to achieve nucleation. Nucleation can be described as the point at which the solubility of the protein has been sufficiently

decreased so as to reach a supersaturated level , at this point protein molecules come out of solution in an ordered formation, to form the first semblance of a crystal lattice. Crystallisation trays containing 24 wells were used, allowing slightly varied crystallisation conditions to be tried in each well (table 2.1). The crystallisation solution is placed in the well and a small drop of protein solution is mixed with crystallisation solution, in this case 4µl of each, is placed on a siliconised coverslip. The coverslip is then placed over the pre-greased well, thus creating a closed environment. This allows vapour diffusion between reservoir and drop to take place, causing the concentration of precipitant in the drop to increase. This lessens the interaction between protein molecules and solvent water which forces protein molecules to form interactions with each other. These interactions decrease the solubility of the protein allowing nucleation to occur.

Although the crystallisation conditions, i.e. the contents of the crystallisation solution, for *E. coli* ALAD are known, there are many unknown factors about the mechanism of protein crystallisation in general. This means that obtaining the correct conditions to crystallise an new protein can be extremely laborious and may take years before useful crystals can be obtained. This has led to protein crystallisation being seen as the rate-limiting step in obtaining the 3-dimensional structure of a molecule. Specific crystallisation screens containing a range of different crystallisation buffers are now available. The use of these screens can make finding ideal crystallisation conditions for a new protein a much faster process.

2.5.1 Crystallisation of *E. coli* ALAD

Once purified, native ALAD underwent buffer exchange into 10 mM Tris/HCl pH 7.4, containing 500 µM $ZnCl_2$ and 10 mM β-mercaptoethanol and was concentrated to 8 mg/ml.

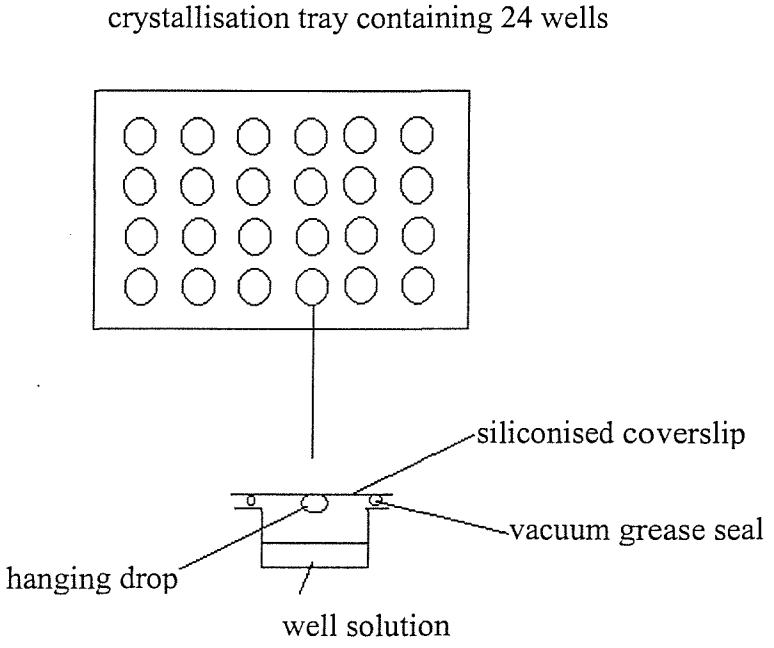
The sample was then crystallised using the hanging drop method using well solutions in grids of 200 mM Tris/HCl buffer pH 8-8.5, containing 500 µM $ZnCl_2$, 10 mM β-mercaptoethanol and saturated ammonium sulphate (AS) (2%-5%), with the drop containing 4 µl of well solution and 4µl of protein solution (table 2.1).

Table 2.1. A typical crystallisation grid for *E. coli* ALAD.

A	B	C	D
pH8.1 2%AS	pH 8.2 2%AS	pH 8.3 2%AS	pH 8.4 2%AS
pH8.1 3%AS	pH 8.2 3%AS	pH 8.3 3%AS	pH 8.4 3%AS
pH 8.1 4%AS	pH 8.2 4%AS	pH 8.3 4%AS	pH 8.4 4%AS
pH8.1 5%AS	pH 8.2 5%AS	pH 8.3 5%AS	pH 8.4 5%AS

Growth of ALAD crystals *via* the hanging drop method was achieved using the conditions illustrated in table 2.1 with crystals grown in the dark at 18°C. The crystals are typically 250-300 microns in diameter and take about two weeks to grow.

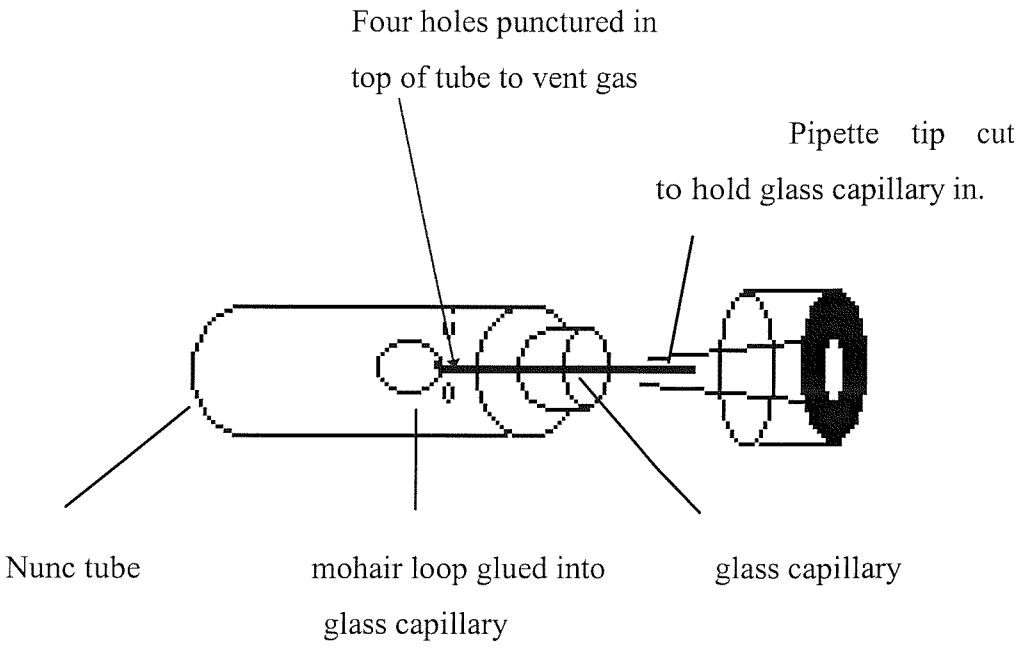
Fig 2.2 Hanging drop crystallisation method



2.5.2 Crystal freezing

In order to store crystals and prolong their useful life during data collection the technique of crystal freezing was employed (50). Crystals to be frozen were removed from their hanging drop using a mohair loop (51) and were transferred to cryoprotectant (50 mM Tris/HCl buffer pH 8.3 containing 40% glycerol, 30% ammonium sulphate, 500 μM $ZnCl_2$, β-mercaptoethanol) for 1 hour and then removed and held in a mohair loop attached to a glass capillary held in a Nunc tube lid (fig 2.3). The crystal in the loop was briefly immersed in liquid ethane for rapid cooling and then transferred to liquid nitrogen. Whilst immersed in liquid nitrogen the crystal and loop were placed into a Nunc tube (also immersed) and the lid fastened. The tubes could now be transferred to a Dewar storage container for transportation.

Fig. 2.3 Nunc tube used in crystal mounting and freezing



2.5.3 Data collection

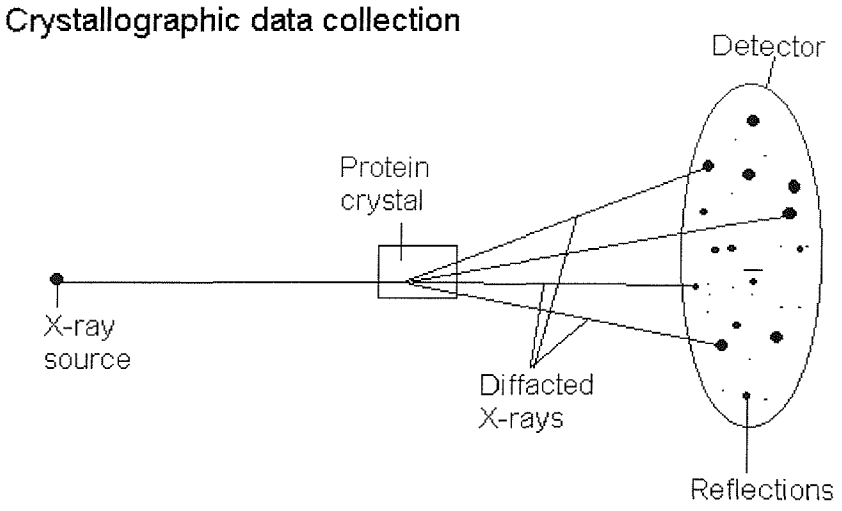
Once suitable crystals have been grown and frozen, data collection can be attempted. There are two main X-ray sources available for data collection, namely rotating anode generators and synchrotrons. X-Rays are generated at synchrotrons by the acceleration of electrons at relativistic velocities around a large ring. This results in the emission of white radiation from which a single wavelength can be chosen for crystal analysis.

Prior to data collection, the crystal is removed from storage and mounted onto a goniometer head, whilst at all times being kept frozen either by liquid nitrogen in the transport Dewar or a dedicated cryostream present at the X-ray source. The crystal is then carefully centred, making sure that it will be fully exposed to the X-ray beam throughout a 360° oscillation. Data can now be collected using the oscillation method (52). This involves exposing the crystal to the X-ray beam for a length of time whilst rotating it through small angles (typically 1.5°) and then stopping the exposure whilst the image plate is scanned (see X-ray detectors). The oscillation angle is chosen to maximise the number of reflections on each separate image without the reflections becoming overlapped.

2.5.4 X-Ray detectors

The only type of x-ray detector used in this project was the MAR research image plate. This is constructed from a 15 μm thick layer of BaFBr(Eu) mounted on a flexible plastic sheet (53). The Eu^{2+} is excited to Eu^{3+} by incident X-rays. These states can then be stimulated into emitting violet light by a He-Ne laser. The intensity of the emitted light is proportional to the number of absorbed X-rays. This allows the image to be read by scanning the plate with a fine laser beam and detecting the emitted light and erasing of the image with a bright light. MAR image plates can be of diameter 18 cm or 30 cm. A simplistic view of data collection can be seen in fig 2.4.

Fig 2.4



X-Rays are scattered by the electrons of the atoms in the protein, with the scattering being proportional to atomic number. The subsequent structure is that of the electron density of molecules in the asymmetric unit, the asymmetric unit being the individual unit that, when repeated an integral number of times, yields the entire unit cell.

The directions of these scattered X-rays, referred to as reflections, are determined by the crystal lattice, not the molecule. The three dimensional crystal lattice gives a three dimensional lattice of scattered R-rays in reciprocal space. The resulting reciprocal lattice has parameters that are the inverse of the crystal lattice, a result of Braggs law of diffraction:

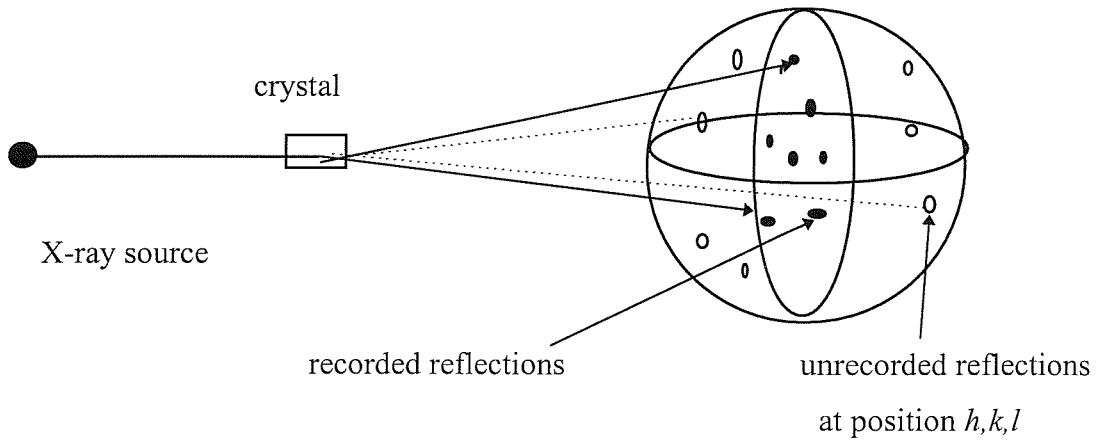
$$2d \sin \theta_n = n\lambda$$

Where d is the spacing between molecules in the crystal, θ_n is the angle of diffraction for the nth diffraction order, and λ is the wavelength of the radiation.

The individual reflections of the reciprocal lattice are generally designated by their Miller indices *h, k, l,* which correspond to their position in the reciprocal lattice. Intensity values are assigned to each reflection as a measure of the relative strength of the reflection. The *hkl* indices are counted from the origin (0,0,0) which lies in the direct path of the x-ray beam (fig. 2.5). Reflections with low indices lie near the origin, and those with high indices lie further away. Reflections that lie near the origin carry information about larger features of the molecules in the unit cell, with reflections far from the origin giving finer detail.

It is usually possible to determine the space group of the lattice and the dimensions of the unit cell from the diffraction pattern of a crystal. In this project this was done using the CCP4 (54) program IPDISP. Once the choice of cell dimensions and space group have been made, the complete dataset can be processed.

Fig. 2.5 Crystallographic data collection, showing reflections at one particular crystal orientation ● and those that could be measured at another orientation○



2.5.5 Data Processing

Data processing involves the transformation of the raw image data into a list of reflection indices (*h,k,l*) and assigning them corresponding amplitudes. The program

used to perform this task on this project was MOSFLM (55). The intensities from different images are then put onto the same scale to account for experimental conditions such as changes in X-ray beam intensity and radiation damage of the crystal during data collection. The program used for calculating scales in this study was SCALA (CCP4) and the scales were applied using Agrovata (CCP4) which adds together partially recorded reflections, monitors and rejects bad agreements between repeated measurements and symmetry equivalents, and averages them for output. The usual measure of the data precision is the merging R factor on intensity, R_{merge}, which is defined as,

$$R_{merge} = \Sigma_{hkl} \mid I\text{-}I_{AVE} \mid / \Sigma_{hkl} I_{AVE}$$

Where, I = measurement of intensity and I_{AVE} = mean intensity

The CCP4 program TRUNCATE was then used to calculate a final reflection file in MTZ format, consisting of a list of values for the structure factor (F) for each *hkl* indexed reflection.

2.5.6 Molecular replacement

Once a data set is collected and processed a technique known as molecular replacement (MR) can be used in order to calculate an electron density map (56). This technique uses the co-ordinates of a previously solved protein that is homologous to the unknown.

2.5.7 Isomorphous phasing model

When studying the proteins with a known structure, but with a small ligand attached, such as the *E. coli* reduced ALAD-ALA complex, the phases of the free protein can be used directly to compute $\rho(x,y,z)$ from native intensities of the new protein using the equation below. These are necessary as, although the reflections measured *via* data collection contain information about amplitude and frequency of the Fourier terms required to build an accurate electron density map, they will not reveal its phase, which is also vital.

$$\rho\,(x,y,z) = 1/V \sum_h \sum_k \sum_l |\, F_{hkl} \,|\, e^{-2\pi i\,(\,hx + ky + lz - \alpha\ \text{model hkl})}$$

Rebuilding the model based on the map should yield phases that differ from those of the original model ,revealing the desired structure.

2.5.8 Non isomorphous phasing models.

In cases where the phasing model is not isomorphous with the desired structure, such as the ALAD-PBG complex discussed later in this thesis, the problem of molecular replacement becomes more complex. In order to use a known protein as a phasing model it must be superimposed on the new structure in the correct position and orientation in the new unit cell. To simplify the search it is possible to split molecular replacement into two discrete processes: a rotation function to find the correct rotation angles and a translation function to provide the translation vector.

2.5.9 Rotation function

Rotation functions involve the comparison of Patterson maps (a map showing the vectors between atoms in a unit cell, instead of the absolute position of the atoms) from a known structure and the desired unknown structure. During this search, only *intra*-molecular vectors should be used and steps taken to exclude the use of *inter*-molecular vectors (vectors between molecules). To achieve this, a Patterson radius is chosen that will exclude the large majority of larger *intra*molecular vectors during the rotation function(57).

Once a radius has been chosen, structure factors for the known protein structure are calculated in an artificial P1 cell. These structure factors are then used in the rotation search using a Eullerian angle system (58) with a range of angles chosen to cover the volume of the space group of the unknown structure. At each angle of rotation the match between the probe and observed Patterson functions is calculated and recorded. After all angles have been recorded, the list is sorted and grouped into peaks which are reported in their size, relative to the root mean square of peak size or sigma. A good match will yield a large peak at several sigma values.

2.5.10 Translation function

Once a rotation function has been found, the translation function can be attempted to determine the translational parameters. This function involves varying the position of the correctly orientated molecules until the cross-Patterson vectors correlate with the Patterson of the unknown struture. This is achieved by using the T2 function (59) which calculates a three dimensional Patterson overlap function using all the inter-molecular vectors.

For a successful translation function to be performed it is essential for an satisfactory rotation function to have been carried out, as it is extremely sensitive to the accuracy of these rotation angles.

2.5.11 Refinement of co-ordinates

In this project, the program XPLOR (60) was used to perform all aspects of refinement. Once a co-ordinate file of the required molecule (including hydrogen atoms) and a structure file have been obtained, all aspects of refinement are possible.

2.5.12 Electron density maps

Different electron density maps may be drawn in order to yield different information about the structure under investigation. Below is a brief description of the two found to be most useful in this project:

2.5.13 Fo - Fc map

The difference Fourier, Fo (observed) - Fc (calculated) reveals peaks where density is not accounted for in the model used to calculate Fc and holes where there is too much density in the model. This map can be useful for finding corrections to the current model e.g. looking for missing waters, changes in mutants, ligand binding etc.

2.5.14 2Fo-Fc map

This is the sum of an Fo map plus an Fo-Fc map. This map shows density at the positions where there are correctly placed atoms in the model and also where atoms missing from the model should be. This map is less biased towards the model than a Fo - Fc map.

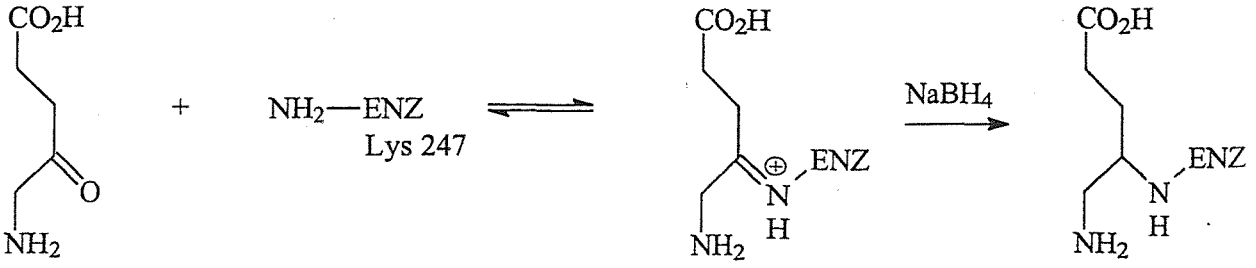
Chapter 3: Characterisation of *E. coli* ALAD A and P binding sites using X-ray crystallography.

The determination of the native structure of *E. coli* ALAD (26) allowed the opportunity to study protein-ligand interactions in order to characterise the active site further and to gain mechanistic information. However, as PBG production involves the dimerisation of two identical molecules of ALA, co-crystallisation of ALAD with additional ALA was not straightforward. To prevent enzymatic turnover of ALA present in the crystallisation solution it was necessary to immobilise an ALA molecule at the P-site of the enzyme *via* the reduction of the Schiff base initially formed between the ε-amino group of Lys 247 and the carbonyl of the ALA substrate molecule (61). Once fully reduced, it became possible to crystallise the reduced ALAD-ALA complex both in native crystallisation conditions and co-crystallised with ALA. This, in theory, would allow both the A- and P-site to be occupied by ALA without enzymatic turnover.

3.1 Formation of reduced *E. coli* ALAD-ALA complex

In order to immobilise ALA at the P-site of *E. coli* ALAD covalent binding of ALA to ALAD *via* reduction of Schiff base formed between ALA and invariant lysine 247 was necessary (scheme 3.1). To a volume of 10 ml of 1mg/ml of *E. coli* ALAD solution in 200mM potassium phosphate buffer, pH 7.0, containing 200μM $ZnCl_2$ and 10mM *β*-Mercaptoethanol 100 mM ALA solution was added to a final concentration of 7 mM. Immediately after the addition of ALA, the Schiff base complex was reduced with $NaBH_4$ (0.5mg/ml). The sample was left to react for 1 hour. The samples were then gel filtered into 10 mM potassium phosphate buffer, pH 7.0 containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol using a PD10 column (Pharmacia). Enzyme activity was then assayed. The activity, after labelling, was virtually zero even after extending the incubation time with substrate in the assay to 30 minutes.

Scheme 3.1 Covalent modification of the P-site *E. coli* ALAD with ALA and $NaBH_4$. ALA is immobilised at the P-site of *E. coli* ALAD *via* reduction of a Schiff base.



3.1.1 Crystallisation of reduced ALAD-ALA complex

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described (chapter 2). The hanging drops contained 5µl of reduced ALAD-ALA adduct solution (7mg/ml) and 5µl of well solution consisting of 200mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200µM $ZnSO_4$ and 6mM β-mercaptoethanol (fig 3.1).

3.1.2 Data collection and processing

The reduced ALAD-ALA complex crystallised in space group I422. Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the reduced ALAD-ALA complex were collected from a cryocooled crystal at SRS Daresbury Laboratories UK, at station 9.6 on an 18cm MAR image plate and were complete to 2.8 Å (fig 3.2). Data were processed with the program MOSFLM (55) and the CCP4 suite (53), for processing statistics see table 3.1. The P-site ALA molecule was located by calculation of difference Fourier map phased with the native *E.coli* ALAD structure (26). The structure was refined using X-PLOR (60) and were rebuilt using Quanta (62) running on Silicon-Graphics computers.

fig 3.1 Typical image from data collection on *E.coli* ALAD crystal (1° oscillation)

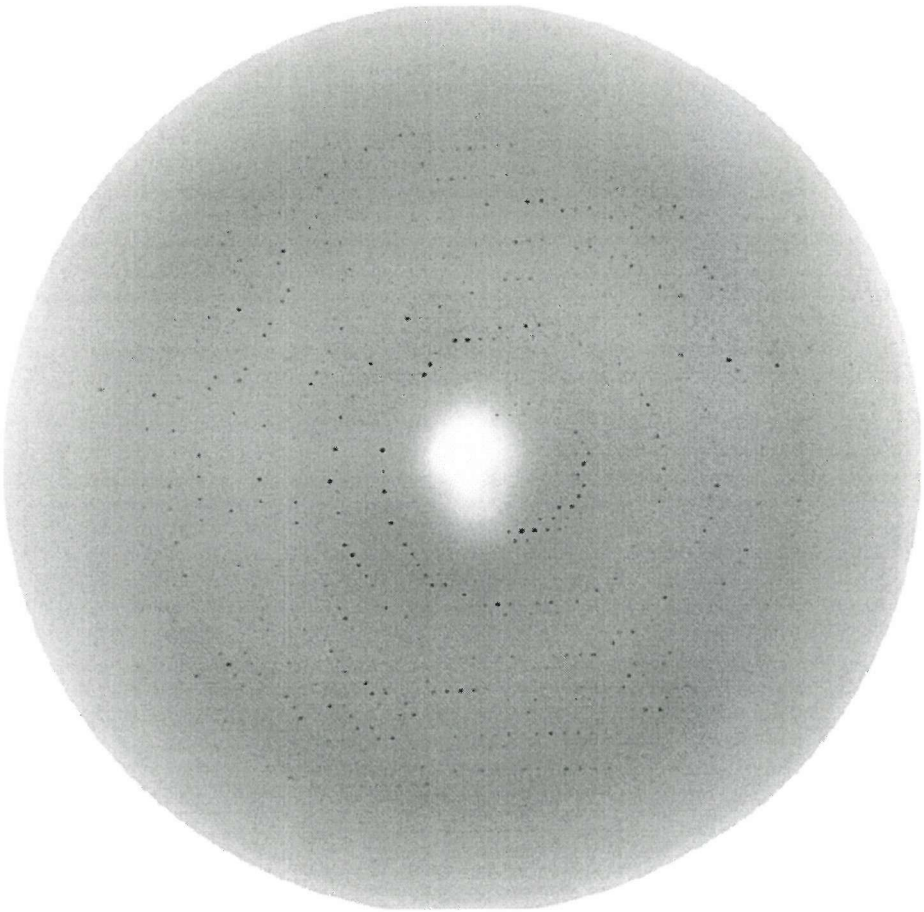
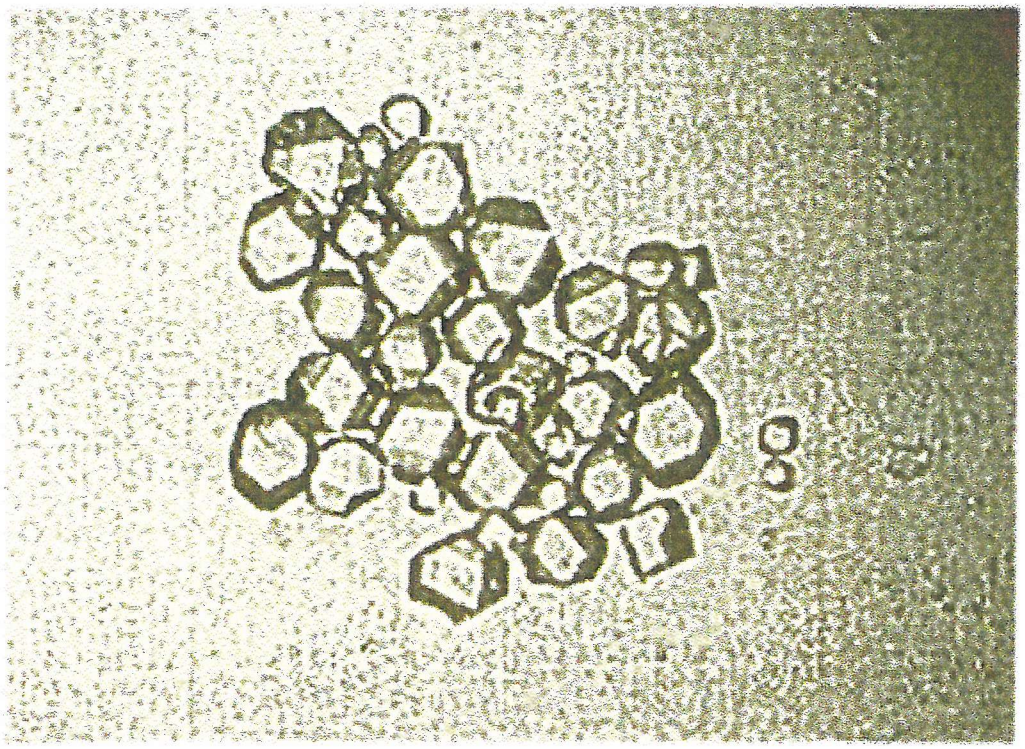


fig 3.2 Native *E. coli* ALAD crystals



300 microns

3.1.3 Quality of structure analysis

The structure of *E. coli* reduced ALAD-ALA complex was refined using data between 12.43 Å and 2.8 Å and had an R-factor of 24 % and a free R-factor of 30.1% (see table 3.1). The structure evaluation program PROCHECK was used to assess the quality of the final model. A Ramachandran plot produced by the program showed 89 % of residues within the allowed regions and 97 % of residues within the additional allowed boundary (fig 3.3).

3.1.4 Analysis of ALAD-ALA complex data

After processing of ALAD-ALA complex data and production of electron density maps it was possible to look at the active site and observe electron density differences between that of the native model and the complex. Inspection of this map at the active site revealed a large portion of difference density adjacent to Lys 247. An ALA molecule was then built into this density and the map underwent further refinement (fig 3.4).

3.1.5 Position and binding interactions of 5-ALA at the P-site

The 5-ALA residue is covalently bound to Lys 247 at the C-4 position *via* the reduction of the initially formed Schiff base by $NaBH_4$ and is positioned between the hydrophobic side chains of Phe 204 and Val 84. Tyrosines 312 and 270 are also involved in forming the binding site, with Phe 89 possibly making van der Waals contact with the C-5 carbon of the 5-ALA residue. Interestingly all of the residues mentioned above are all highly conserved within ALAD sequences, suggesting a high importance for hydrophobic aromatic interactions in P-site substrate binding.

Table 3.1 Data processing statistics for ALAD-ALA complex

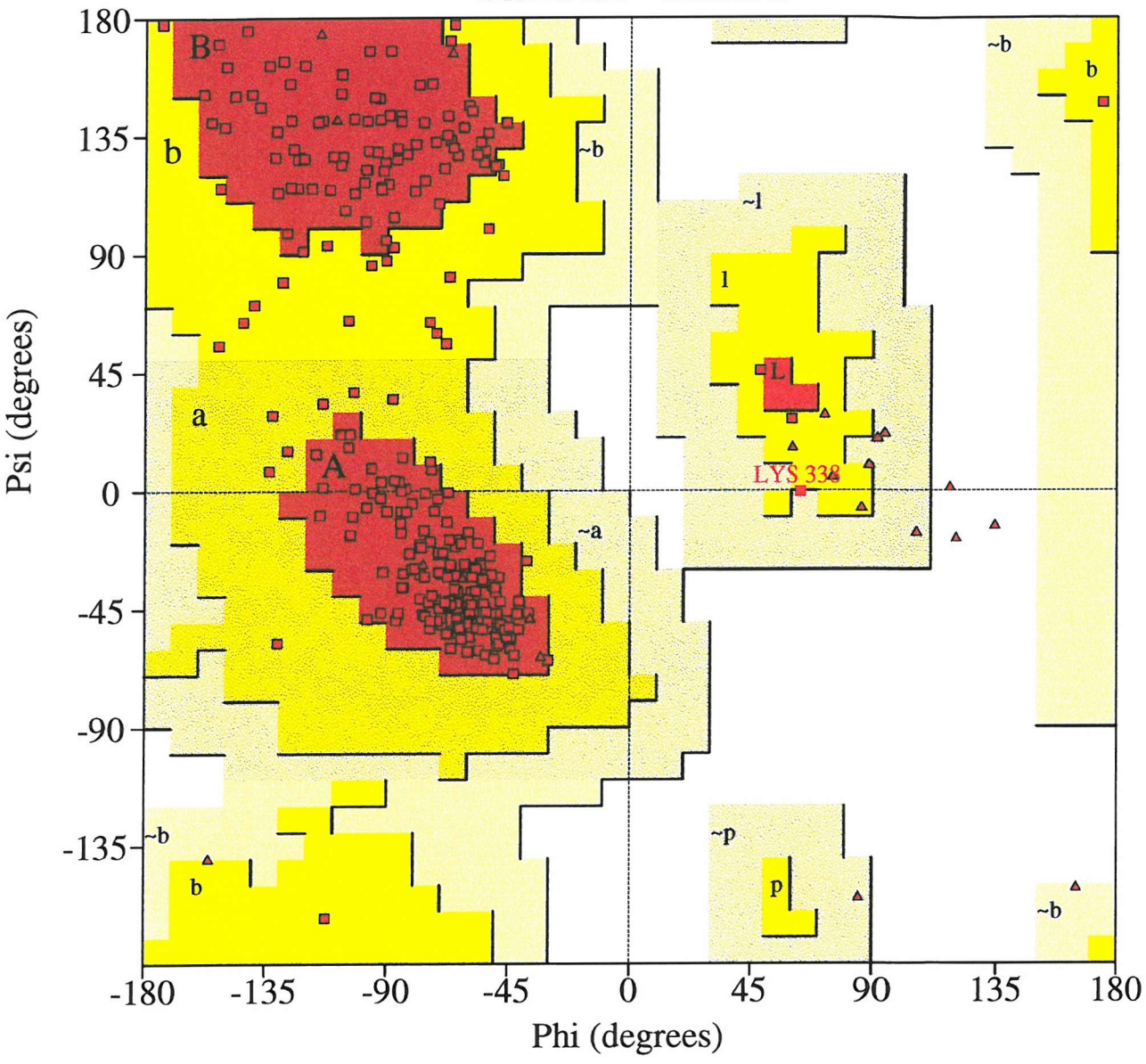
Space group	I422
a b (Å)	128.7
c (Å)	143.67
Entire data set	
Resolution range (Å)	12.43 -2.8
R-merge (%)	5.4
Completeness (%)	99
Overall (I/Iσ)	5.3
Fraction I> 3σ (I) %	84.0
Multiplicity	6.1
Outer Shell	
Resolution range (Å)	2.87 - 2.8
R-merge (%)	7.1
Completeness	96.3
I / σ I	1.9
Fraction I> 3σ (I) %	65.8
Multiplicity	6.1
Refinement	
R factor (%)	24.8
Free R factor (%)	31.2
Resolution range (Å)	15 - 2.8
Cut off	2 σ
Total number of reflections	15038
Number of atoms	3035

PROCHECK

Fig.3.3

Ramachandran Plot

ALAD-ALA



Plot statistics

Residues in most favoured regions [A,B,L]	252	88.1%
Residues in additional allowed regions [a,b,l,p]	33	11.5%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.3%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	286	100.0%
Number of end-residues (excl. Gly and Pro)	3	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	322	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig 3.4 Active site of *E. coli* ALAD with ALA attached to P-site Lys 247 *via* a reduced Schiff base.The proposed A-side ALA binding site is clearly visible, as is the catalytic zinc ion. The carboxyl group of ALA interacts with Ser 273 and Tyr 312 whilst Lys 195 is in close proximity to Lys 247.

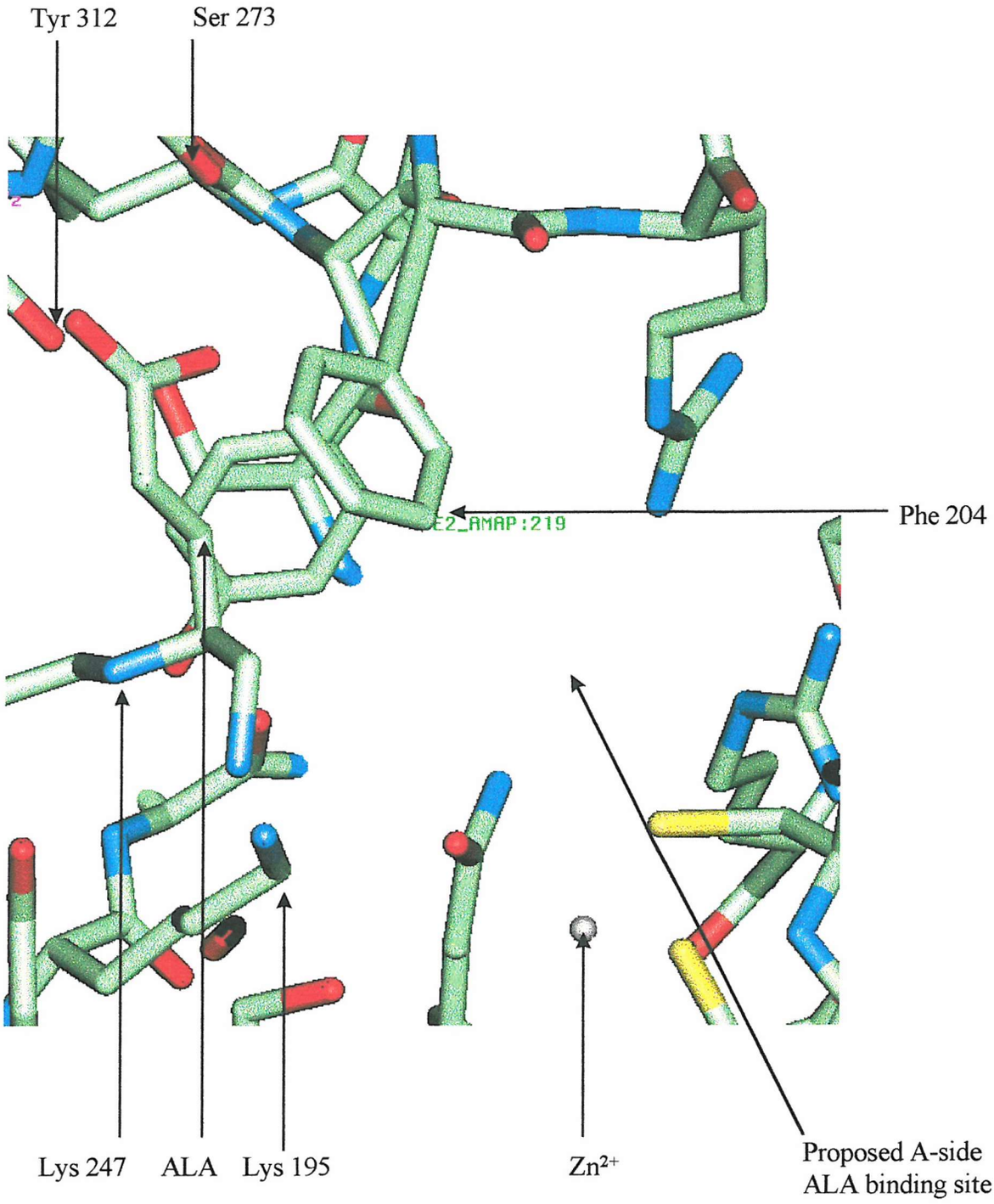
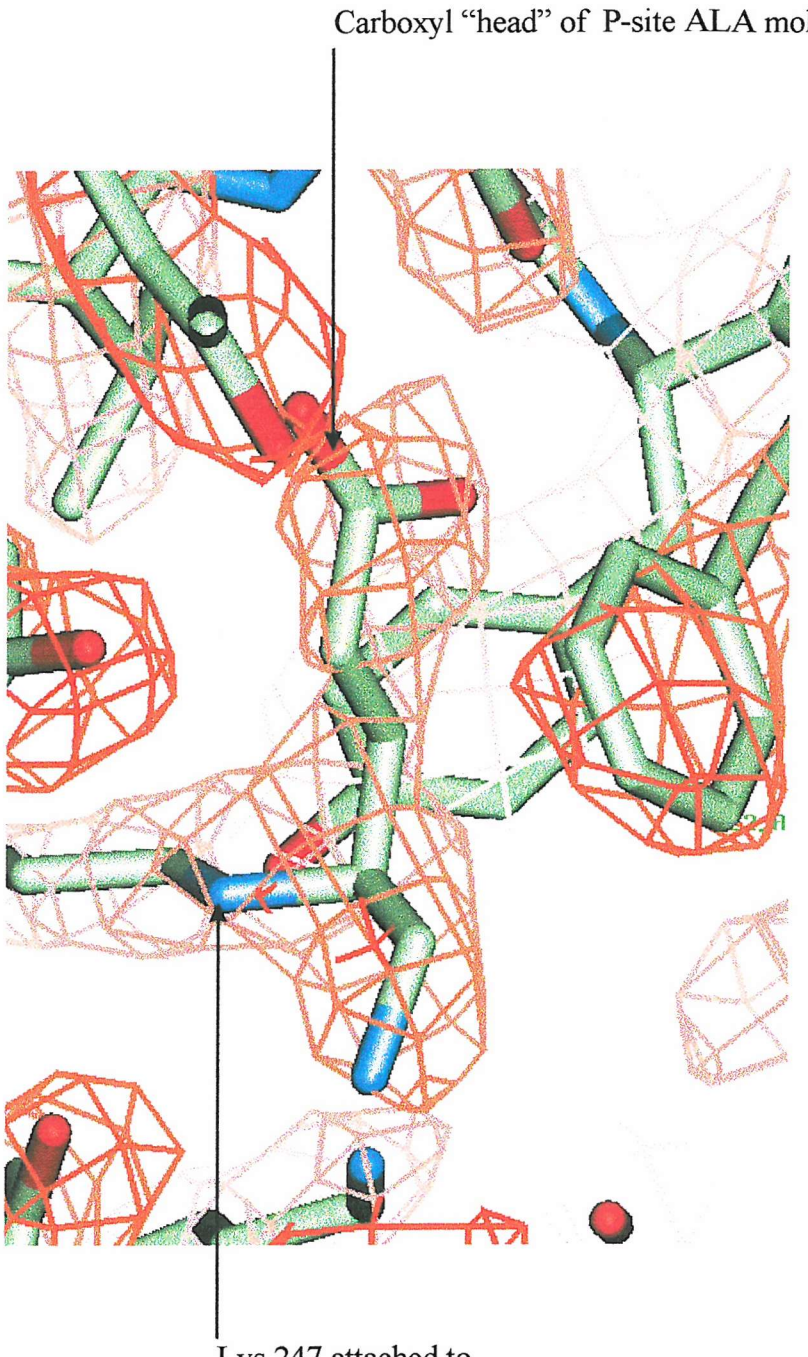


Fig 3.5 2Fo-Fc map of reduced ALAD-ALA complex at active site, with ALA modelled in as reduced Schiff base adduct. The position of the substrate molecule and interactions with P-site ALAD residues are clearly defined.



3.1.6 ALA carboxyl group binding at the P-site

The carboxyl group binding at the P-site of the enzyme is well defined by the structure. One carboxyl oxygen H-bonds to the hydroxyl of Ser 272 with the remaining carboxyl oxygen hydrogen bonding with the nitrogen on the main chain of the same residue. This second oxygen also forms a hydrogen bond with the phenolic oxygen of Tyr 312. The two residues mentioned above are also highly invariant and suggest a common binding structure among ALADs (fig 3.5).

3.1.7 ALA amino group binding at the P-site

The amino group of the bound 5-ALA molecule is positioned at the base of the active site At this position the amino group is within hydrogen bonding distance of Asp 118 a invariant residue which would be ionised at the optimum pH range of the enzyme and which itself H-bonds with invariant Ser 165 (fig 3.5). However there is some evidence to suggest that the position of the amino group of this substrate has some degree of flexibility (63) which would have important implications for catalysis. These are discussed later.

It is important that despite being linked to ALAD through a reduced Schiff base, the position of the ALA carbon skeleton is superimposable with the carbon skeleton of LA linked to ALAD through a non reduced Schiff base (26) suggesting that reduction of the Schiff base has little observable effect on the structure.

3.2 Crystallisation of reduced ALAD-ALA + ALA complex

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described in chapter 2. The hanging drops contained 5µl of reduced ALAD-ALA adduct solution (7mg/ml) and 5µl of well solution consisting of 200mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200µM $ZnSO_4$, 6mM β-mercaptoethanol and 5mM ALA.

3.2.1 Data collection and processing

The reduced ALAD-ALA complex crystallised in space group I422 (table3.2). Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the reduced ALAD-ALA+ALA complex were collected from cryocooled crystals on beamline 14 at the ESRF (Grenoble, France) on an 18cm MAR image plate and were complete to 3.2 Å. Data were processed with the program MOSFLM (55) and the CCP4 suite (1994), for processing statistics see table 3.2. The A-site ALA molecule was located by calculation of difference Fourier map phased with the native *E. coli* ALAD structure (26). The structure was refined using X-PLOR (60) and were rebuilt using Quanta (62) running on Silicon-Graphics computers.

3.2.2 Quality of structure analysis

The structure of *E. coli* reduced ALAD-ALA + ALA complex was refined using data between 12.63 Å and 3.2 Å and had an R-factor of 22.4 % and a free R-factor of 26.8% (see table 3.2). The structure evaluation program PROCHECK was used to assess the quality of the final model. A Ramachandran plot produced by the program showed 79.4 % of residues within the allowed regions and 99.3 % of residues within the additional allowed boundary (fig. 3.6).

3.2.3 Analysis of reduced ALAD-ALA + ALA complex data

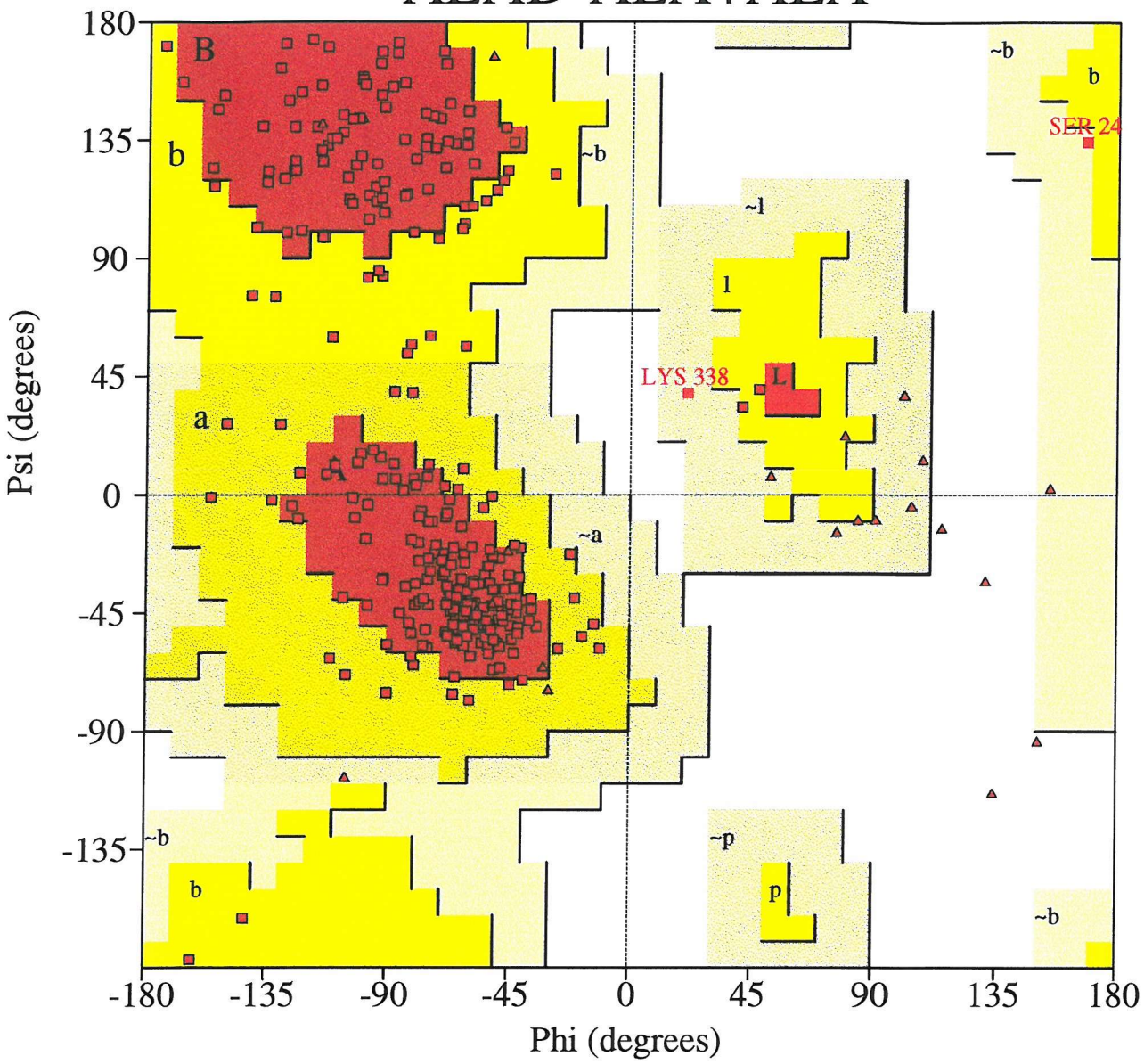
After processing of reduced ALAD-ALA + ALA complex data and production of electron density maps it was possible to look at the active site and observe electron density differences between that of the native model and the complex . Inspection of this map at the active site revealed a large portion of difference density adjacent to Lys 247 and further difference density at the putative A-side 5-ALA binding site. Two 5-ALA molecules were then built into this density and the map underwent further refinement (fig 3.7).

Table 3.2 Data processing statistics for ALAD-ALA + ALA complex

Space group	I422
a b (Å)	128.43
c (Å)	143.67
Entire data set	
Resolution range (Å)	12.63 - 3.2
R-merge (%)	11.3
Completeness (%)	94.2
Overall (I/Iσ)	4.6
Fraction I> 3σ (I) %	78,1
Multiplicity	4.1
Outer Shell	
Resolution range (Å)	3.27-3.2
R merge (%)	8.9
completeness	88.4
I / σ I	3.7
Fraction I> 3σ (I) (%)	48.3
Multiplicity	3.7
Refinement	
R factor (%)	22.4
Free R factor (%)	26.8
Resolution range (Å)	15 - 3.2
Cut off	2 σ
Total number of reflections	6734
Number of atoms	3035

PROCHECK
Fig.3.6

Ramachandran Plot
ALAD-ALA+ALA



Plot statistics

Residues in most favoured regions [A,B,L]	227	79.4%
Residues in additional allowed regions [a,b,l,p]	57	19.9%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	286	100.0%
Number of end-residues (excl. Gly and Pro)	3	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	322	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig. 3.7 2Fo-Fc map, contoured at 2σ, showing position of A- and P-site ALA binding to *E. coli* ALAD. The carboxyl group of A-site ALA forms a salt bridge with Arg 205 whilst the amino group of the substrate molecule interacts with Ser 165. This leaves the body of the molecule free to react with the active site zinc ion and Lys 195.

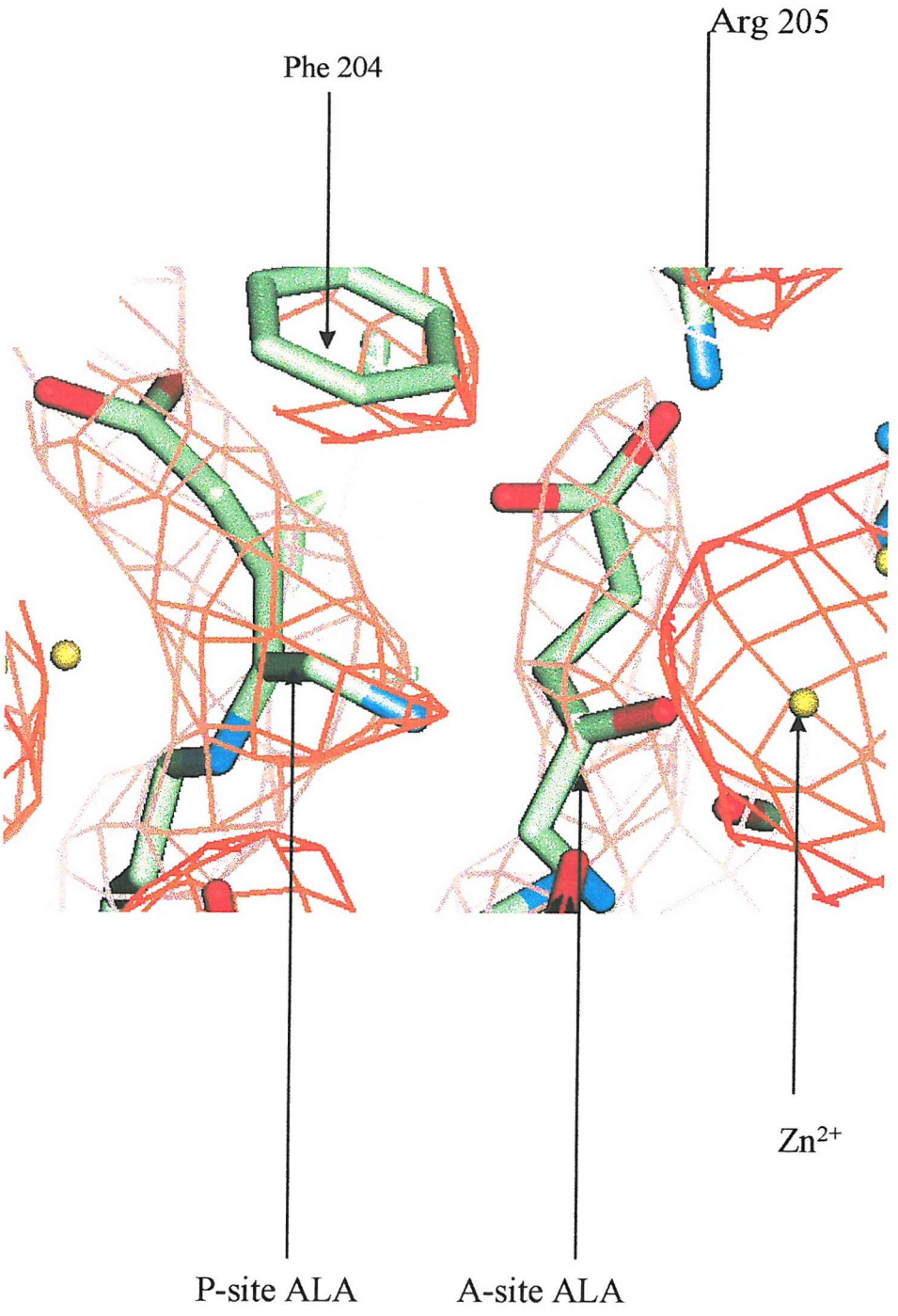
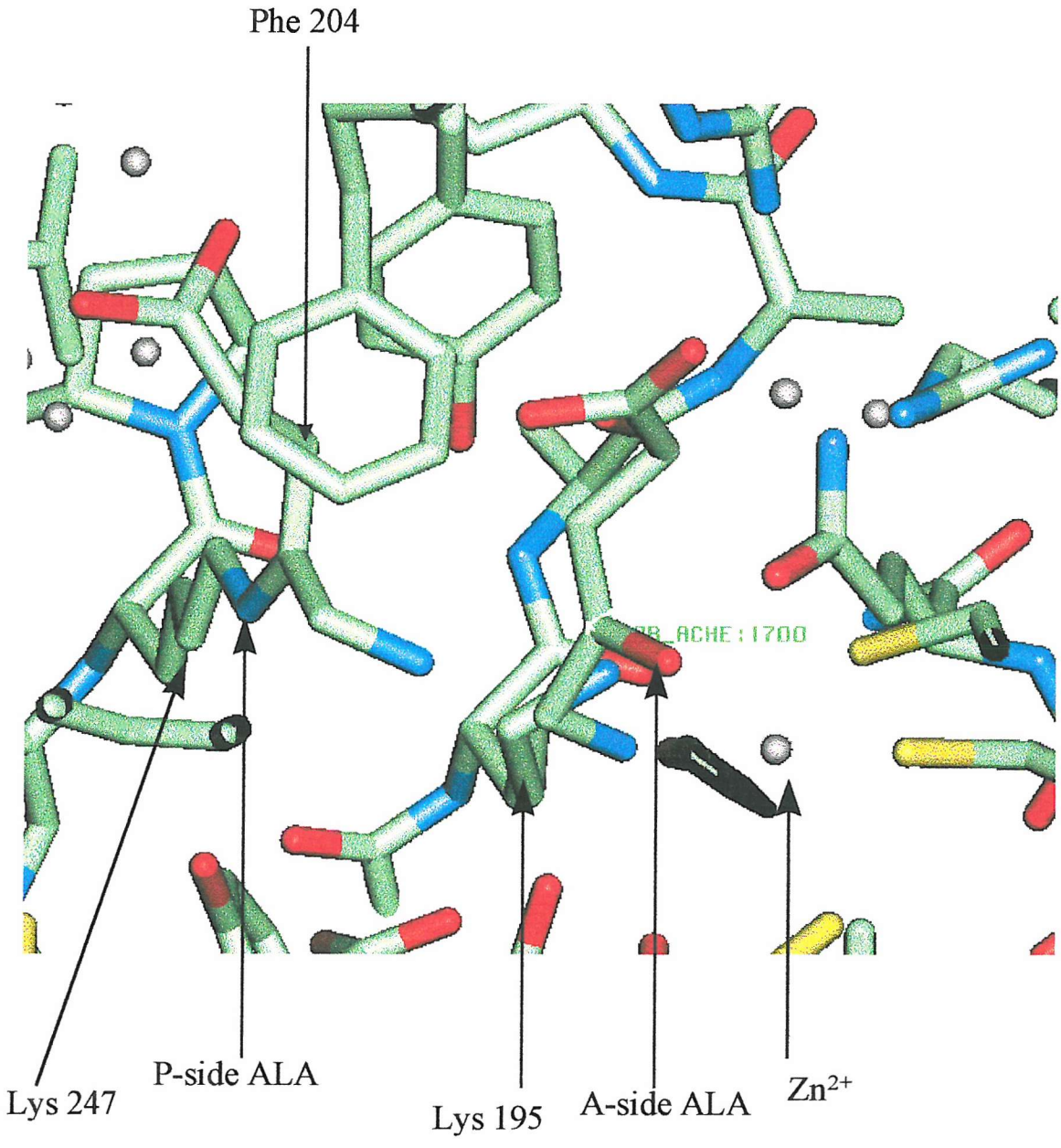


Fig. 3.8 Active site of *E. coli* ALAD with ALA at both A and P sites. ALA is bound to Lys 247 *via* a reduced Schiff base, allowing ALA to occupy the A site without PBG formation taking place. Phe 204 packs against the bound substrate molecules, shielding them from solvent



3.2.4 Carboxyl binding at A-site ALA

The A-site ALA binding site consists of a pocket partially made up by a number of conserved residues, with the zinc ion co-ordinated by cysteines 120, 122 and 130. The carboxyl group of the A-site ALA substrate molecule is positioned towards arginines 205 and 216 at the head of the A-binding site, forming a salt bridge with Arg 205. These arginine residues form part of a mobile section of the enzyme and are poorly resolved unless either substrate or inhibitor (LA) are bound (26). It is hugely significant that the binding of the P-site ALA to active site is essential for mobilising residues at the A-site to permit subsequent binding of A-site ALA. This positioning leaves the amino group of ALA close to the invariant Asp118 - Ser 165 pair, with the amino nitrogen H-bonding to the hydroxyl group of Ser 165 leaving the carbonyl oxygen in the vicinity of the zinc ion. This positioning allows the "head to head" condensation that is essential for PBG formation (fig 3.8).

3.3 Determination of ALAD-PBG complex structure in space group $P4_22_12$

To characterise further the active site of *E. coli* ALAD and to confirm the positions of both A- and P-site ALA determined in the two previously reported structures, the co-crystallisation of ALAD with PBG was attempted.

3.3.1 Crystallisation of the ALAD-PBG complex

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described (chapter 2.) The hanging drops contained 5µl of reduced ALAD solution (7 mg/ml) and 5 µl of well solution consisting of 200 mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200 µM $ZnSO_4$, 6 mM β-mercaptoethanol and 3 mM PBG.

3.3.2 Data collection and processing

The ALAD-PBG complex readily crystallised in the above conditions and yielded exceptionally large crystals (300 x 300 x 300 microns). Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v).

The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the ALAD-PBG complex were collected from a cryocooled crystal at SRS Daresbury Laboratories UK, at station 9.6 on an 18cm MAR image plate and were complete to 2.5 Å. Data were processed with the program MOSFLM (55) and the CCP4 suite (1994).

The dimensions of the unit cell of this new crystal were found to be almost identical to those determined in previous *E. coli* ALAD structures (a=130.46Å b=130.46Å c=144.09Å $\alpha=\beta=\gamma=90$). Examination of the processed data using CCP4 software revealed no systematic absences in *hkl* indicating that the ALAD-PBG complex had not crystallised in the I422 space group seen in previous *E. coli* ALAD structures. Data were initially processed in P422 and systematic absences further examined. The absence of every second reflection along 00l revealed a 4_2 screw axis. Due to an absence of data found when examining reflections along 0k0 and h00 the final space group was determined by translation function.

3.3.3 Estimation of initial phases by molecular replacement

3.3.4 Rotation function

Therefore the radius of integration for the rotation function was chosen to be 25 Å and the primitive cell enclosing the model was a=140 Å b=140 Å c=100 Å $\alpha=\beta=\gamma=$ 90. The resolution range of data used was 20-2.5 Å. The search model used for molecular replacement was the *E. coli* ALAD monomer with zinc ions removed.

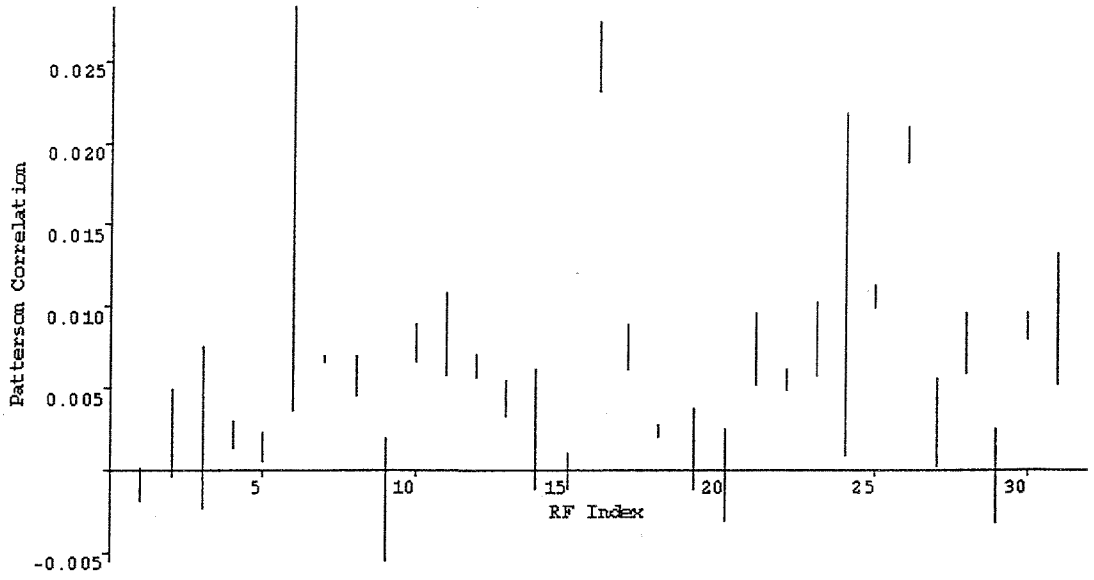
The rotation function was calculated using sharpened structure factors in X-PLOR, and PC refinement was carried out using the barrel and tail of the monomer as rigid bodies.

Each peak in fig. 3.9 represents one solution with the lower point of the line representing the Patterson correlation values before PC refinement and the upper point representing the value after PC refinement. The co-ordinates that these peaks represent are shown in table 3.3.

Table 3.3 Solutions of rotation function after PC refinement.

Peak number	**Solution after PC refinement** ($\theta_1,\theta_2,\theta_3$)
1	250.87, 29.112, 70.419
2	10.25, 34.283, 345.82
3	52.364, 165.63, 108.99

Fig 3.9. A plot of the results of a rotation function after PC refinement. Each vertical line represents one solution the lower end of which indicates the PC value before refinement and the upper end giving the value after refinement.



3.3.5 Translation function

The solution from peak 1 was chosen (representing a rotation solution of 250.87 29.112 70.419) to calculate a translation function and produced a peak 31 sigma

above the mean. The packing of this solution was viewed using MOLPACK (CCP4) and showed sensible contacts. The translation function was used to finally determine the space group in which the complex had crystallised, indicating space group $P4_22_12$.

3.3.6 Refinement and model building

The initial R-factor of the model after molecular replacement was 43.6% (R_{free}= 49.2%). After rigid body refinement within X-PLOR using data from 15-2.5 Å this value dropped to 31.4% (R_{free} = 35.7%)

2Fo-Fc and Fo-Fc electron density maps were generated using phases calculated from the model co-ordinates to determine the position of zinc ions and PBG molecules present. The maps revealed zinc ions at expected sites but difference density was also present for PBG at the active site, into which PBG was modelled. At this point refinement was halted as further data on this complex had been collected to 2.0 Å which was processed in preference.

3.4 Crystallisation of the ALAD-PBG complex in space group I422

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described (chapter 2). The hanging drops contained 5µl of ALAD-PBG adduct solution (7 mg/ml) and 5µl of well solution consisting of 200 mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200 µM $ZnSO_4$ and 6 mM β-mercaptoethanol.

3.4.1 Data collection and processing

The ALAD-PBG complex crystallised in space group I422 Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the ALAD-PBG complex were collected from a cryocooled crystal at SRS Daresbury Laboratories UK, at station 9.6 on an 18cm MAR image plate and were complete to 2.0 Å. Data were processed with the program MOSFLM (55) and the CCP4 suite (53), for processing statistics see table

3.4.. The PBG molecule was located by calculation of difference Fourier map phased with the native *E .coli* ALAD structure (26). The structure was refined using X-PLOR (60) and were rebuilt using Quanta (62) running on Silicon-Graphics computers.

3.4.2 Quality of structure analysis

The structure of E.*coli* ALAD-PBG complex was refined using data between 10.23 Å and 2.0 Å and had an R-factor of 23.7 % and a free R-factor of 28.9% (see table 3.4). The structure evaluation program PROCHECK was used to assess the quality of the final model. A Ramachandran plot produced by the program showed 90.3 % of residues within the allowed regions and 99.3 % of residues within the additional allowed boundary (fig 3.10).

3.4.3 Crystallisation of ALAD in I422 and $P4_22_12$

Although crystallised simultaneously, data collected on separate crystals of the ALAD-PBG complex revealed crystallisation in two differing space groups, namely I422 and $P4_22_12$. Analysis of the two monomers in the asymmetric unit using the CCP4 program LSQKAB (ref) revealed that they are related by a translation of 0.4386, 0.561, 0.415 which indicates a loss in crystallographic symmetry. This explains the presence of two monomers in the asymmetric unit and the transition from space group I422 to $P4_22_12$. It is probable that the shift in space group occurred during crystal freezing.

3.4.4 Analysis of ALAD-PBG complex data

After processing of ALAD-PBG complex data and production of electron density maps it was possible to look at the active site and observe electron density differences between that of the native model and the complex . Inspection of this map revealed a large feature of difference density that occupied the centre of the active site. A PBG molecule was then built into this density and the map underwent further refinement. (fig 3.11).

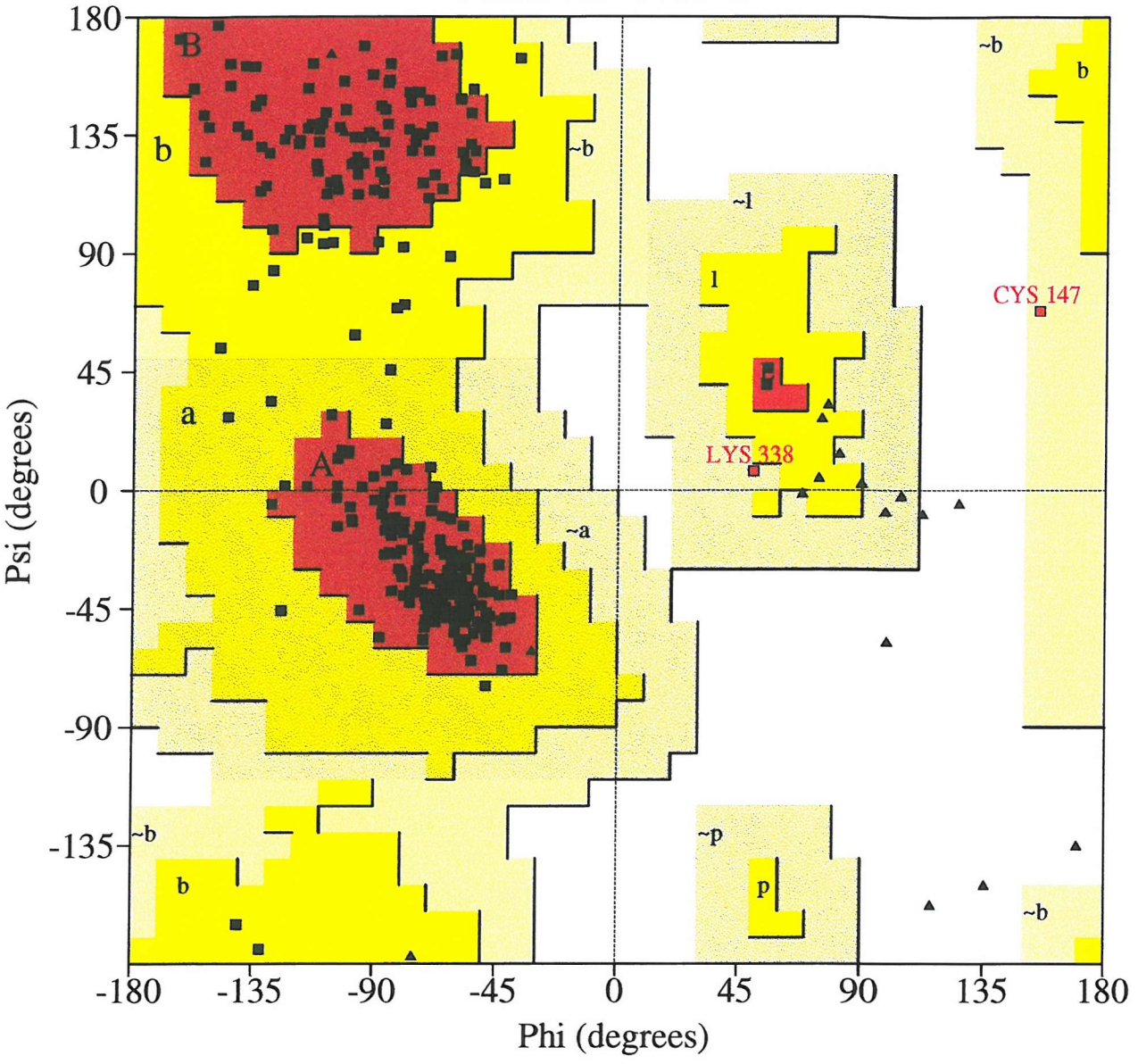
Table 3.4 Data processing statistics for ALAD-PBG complex

Space group	I422
a b (Å)	130.46
c (Å)	144.09
Entire data set	
Resolution range Å	36.47 - 2.0
R-merge (%)	12.9
Completeness (%)	89.5
Overall (I/Iσ)	3.4
Fraction I> 3σ (I) (%)	62.4
Multiplicity	3.1
Outer Shell	
Resolution range (Å)	2.36 -2.3
R merge (%)	9.8
completeness	81.8
I / σ I	0.7
Fraction I> 3σ (I) (%)	18.3
Multiplicity	2.2
Refinement	
R factor (%)	23.7
Free R factor (%)	28.9
Resolution range	15 - 2.5
Cut off	2 σ
Total number of reflections	49042
Number of atoms	6106

PROCHECK

Fig 3.10

Ramachandran Plot
ALAD-PBG



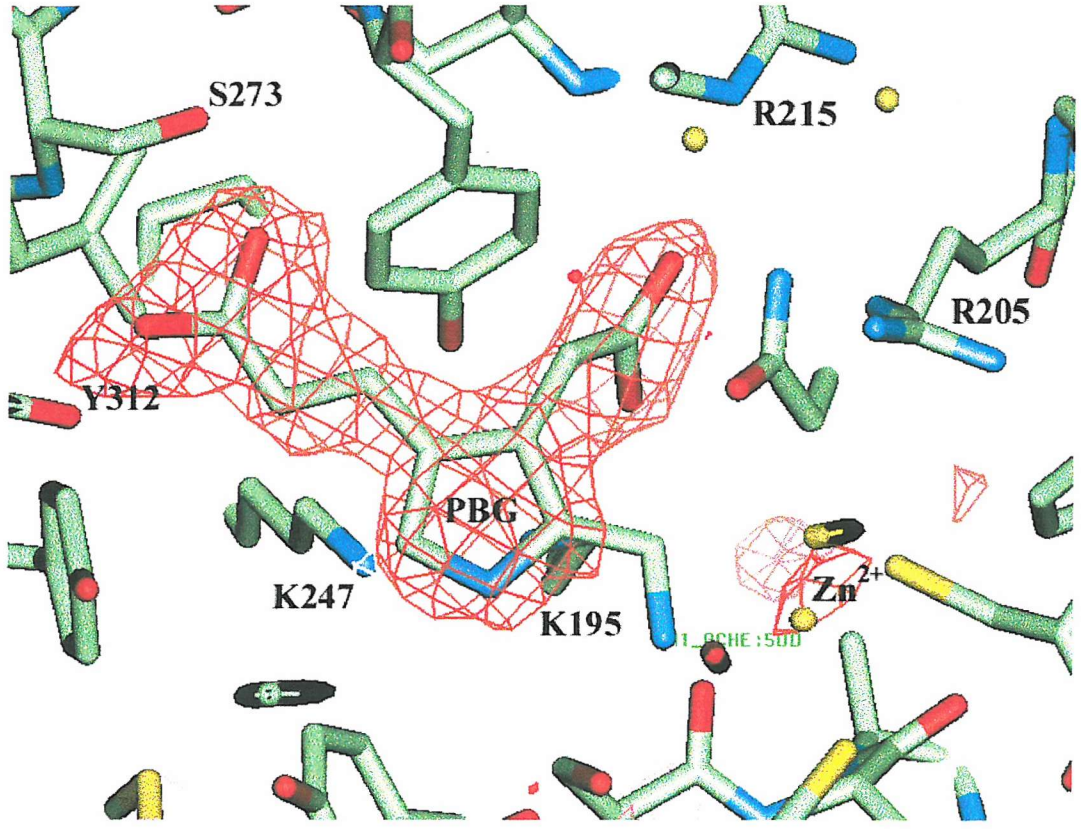
Plot statistics

Residues in most favoured regions [A,B,L]	260	90.3%
Residues in additional allowed regions [a,b,l,p]	26	9.0%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	288	100.0%
Number of end-residues (excl. Gly and Pro)	6	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	327	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

ALAD-PBG_01.ps

fig 3.11 Fo-Fc map, contoured at 2σ, showing position of PBG at the active site of *E. coli* ALAD. PBG has been built into difference density. The position of the A and P side chains confirms earlier identification of the A- and P- ALA binding sites within the active site cavity.



3.4.5 Position and binding interactions of PBG

The PBG molecule is clearly defined at the active site of *E. coli* ALAD with the pyrrole body of the molecule occupying the centre of the active site and the propionate and acetate side chains occupying the P- and A-sites respectively.

3.4.6 Propionate side chain binding at the P-site

The carboxyl group of the PBG propionate side chain makes identical interactions with the P-site as the carboxyl groups of ALA and LA previously described. One carboxyl oxygen H-bonds to the hydroxyl of Ser 272 with the remaining carboxyl oxygen hydrogen bonding with the nitrogen on the main chain of the same residue. This second oxygen also forms a hydrogen bond with the phenolic oxygen of Tyr 312.

3.4.7 Acetate side chain binding at the A-site

The A-site binding of the carboxyl group of the PBG acetate side chain is identical to that seen with the previously described structure of reduced ALAD-ALA + ALA. The carboxyl group of the acetate side chain is positioned towards arginines 205 and 216 at the head of the A-binding site, forming a salt bridge with Arg 205.

3.4.8 Amino side chain binding at the A-site

Once the PBG molecule had been modelled into the difference density and undergone refinement it became apparent that the position of the A-site ALA amino side chain of the molecule was poorly defined. This is in keeping with the reduced ALAD-ALA + ALA structure where the amino group was also poorly resolved and suggests a degree of mobility for this group.

3.4.9 Interaction of pyrrole group of PBG with active site.

The pyrrole ring system of the PBG molecule is positioned at the center of the active site cavity and is well defined by difference electron density. The pyrrole nitrogen is positioned in close proximity to residues Ser 165 and Asp 118. The highly conserved active site loop residue, Phe 204, is well defined with its side chain packed against the product molecule creating a hydrophobic environment. The key

lysine residues 195 and 247 are aligned at the base of the PBG molecule with Lys 247 on the P side and Lys195 on the A side (fig 3.11).

The position of the PBG molecule at the active site is consistent with the positions of the ALA molecule at both the A and P sites observed in the reduced ALAD-ALA and reduced ALAD-ALA + ALA structures, previously outlined in this chapter, and allows confidence in further characterisation of the active site.

3.5 Characterisation of A and P binding sites and implications for PBG formation

The structural solutions of reduced ALAD-ALA, reduced ALAD-ALA+ALA and ALAD-PBG allow the comprehensive characterisation of both the P- and A-ALA binding sites within the active site cavity of *E. coli* ALAD, increasing understanding of both substrate binding and PBG formation.

The initial reaction between ALAD and ALA is the binding of substrate ALA at the P-site of the active site and the formation of a Schiff base between the C-4 carbon of ALA and the ε-nitrogen of Lys 247. The incoming ALA is correctly positioned at the P-site by hydrogen bond formation between the carboxyl group of ALA and conserved residues Tyr 312 and Ser 273, with the amino group of ALA interacting with Asp 118.

The Schiff base formation is made possible by the positioning of Lys 195 adjacent to Lys 247. It is probable that the proximity of these two amino groups result in the suppression of the pK_a of Lys 247 by Lys 195, leaving Lys 247 unprotonated and able to act as a nucleophile. Significantly, NMR studies have demonstrated the amino group of P-site ALA is deprotonated (64,65) which is of great importance when considering the subsequent Schiff base formed between this amino group and the C-4 carbon of A-site ALA during PBG synthesis.

Once ALA has bound to the P-site of ALAD Phe 204, which is positioned in the flexible active site loop, packs tightly against the methylene groups of the substrate and shields it from solvent. This hydrophobic interaction between ALA and Phe

204 appears to be essential to PBG formation as the active site loop is now more stable allowing remaining loop residues to function in the formation of the A-side ALA binding site.

Once the active site loop is secure the A-site ALA molecule is able to bind. The carboxyl group of the second substrate forms a salt bridge with Arg 205, a residue found in the flexible active site loop, with the amino group of the molecule being close to the conserved Asp 118 Ser 165 interacting pair at the base of the active site. This leaves the body of the A-site substrate in a position to interact with the active site zinc ion and Lys 195 which may play a further catalytic role. The two substrate molecules are now ideally placed for PBG formation.

Chapter 4. Mutation of active site residues of *E. coli* ALAD

Following characterisation of the active site of *E. coli* ALAD by X-ray crystallography (chapter 3), a structural understanding of the A- and P-site ALA binding sites became more clear, and mutation of residues directly implicated in substrate binding and catalysis was possible.

This chapter describes the site directed mutagenesis of key A- and P-site residues, in *E. coli* ALAD, their purification and characterisation.

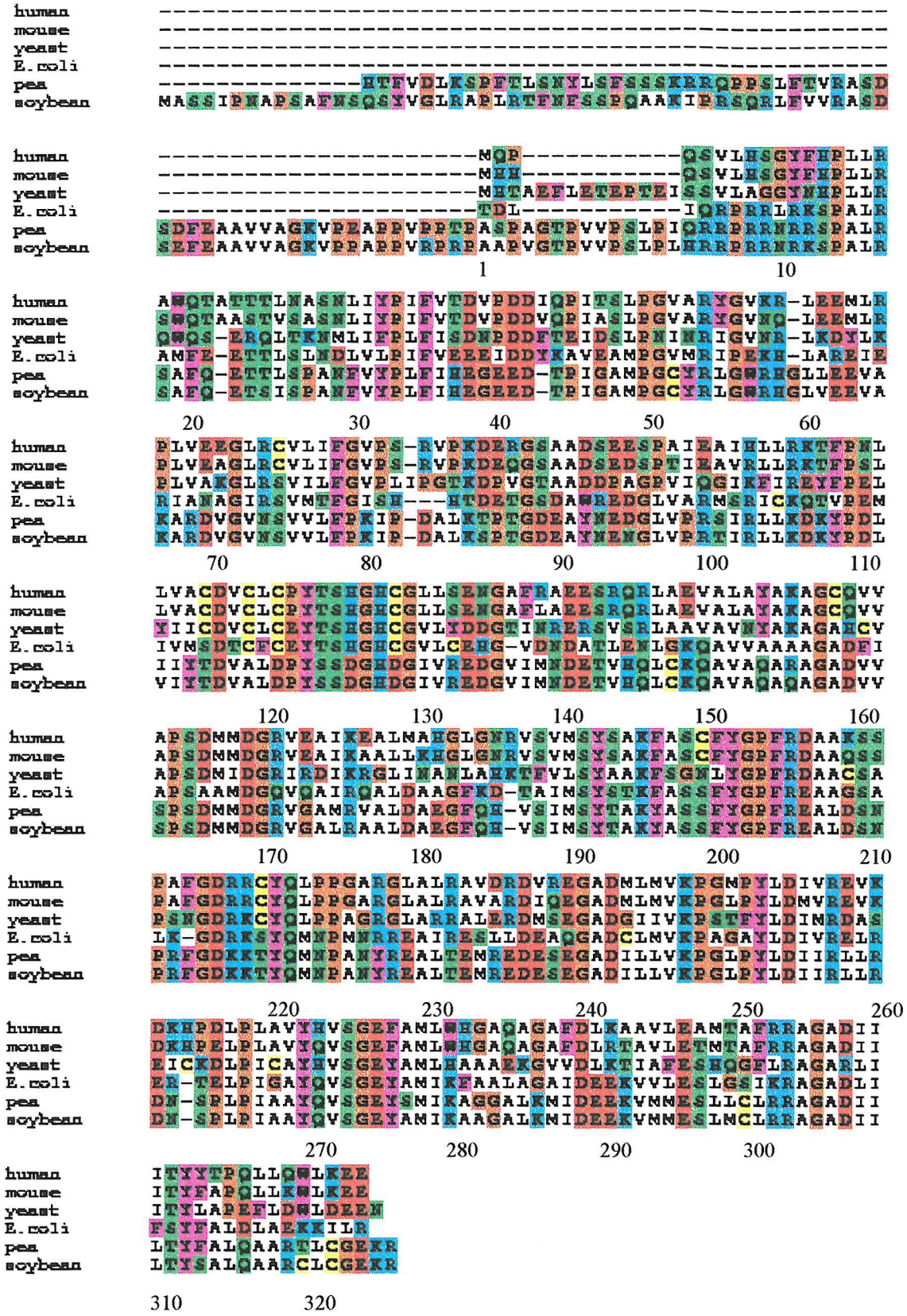
4.1 Rationale in choosing *E. coli* ALAD residues for targeting by site directed-mutagenesis

The residues chosen for mutation were selected on the basis of their likely contribution to substrate binding and catalysis. Sequence alignment revealed an invariance of targeted residues between species (Fig 4.1) All proposed mutations are listed in table 4.1 and their proposed functions described in sections 4.2 and 4.3. A diagram of the active site of *E. coli* ALAD showing the position of residues selected for site directed mutagenesis and H-bonding between substrate ALA and active site residues can be seen in fig 4.2.

Table 4.1 *E. coli* ALAD mutants prepared for analysis

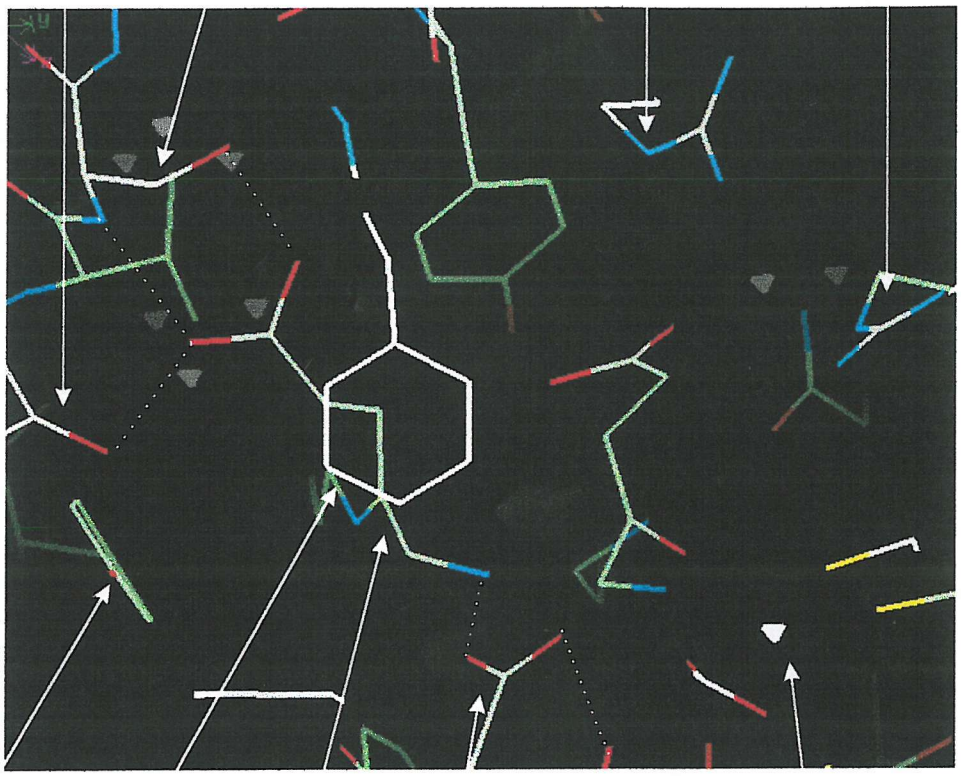
wild type residue		mutant residue
Tyrosine 312	⇒	Phenylalanine
Serine 273	⇒	Threonine
Tyrosine 192	⇒	Phenylalanine
Aspartate 118	⇒	Glutamate, Asparagine and Alanine
Serine 165	⇒	Threonine
Arginine 215	⇒	Lysine
Phenylalanine 204	⇒	Isoleucine

Fig 4.2 Alignment of ALAD amino acid sequences from human, mouse, yeast, *E. coli*, pea and soybean sources.



Key to colour scheme: Blue = basic residues, Red = acidic residues, Green = small polar residues Pink = aromatic residues, Yellow = cysteine residues, Orange = glycine and proline residues.

Fig. 4.2 Active site of *E. coli* ALAD with ALA at both the P- and A-sites. All mutated residues are labelled and H-bonds between ALA molecules and active site residues are shown.

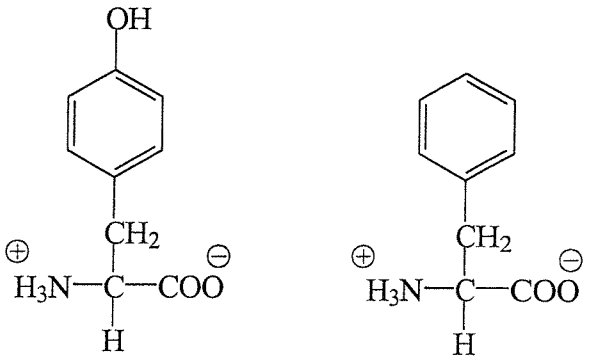


4.2 Description and rationale of P-site mutations

4.2.1 Tyr 312 ⇒ Phe

This residue is located at the top of the P-site ALA binding site and is thought to be involved in binding P-site ALA by forming a hydrogen bond with the carboxyl group of the substrate molecule. The mutation of Tyr 312 to Phe (scheme 4.1) will remove the hydroxyl group involved in the hydrogen bond but will cause minimal disruption to active site structure, as both are large aromatic residues.

Scheme 4.1 Chemical structure of tyrosine and phenylalanine.

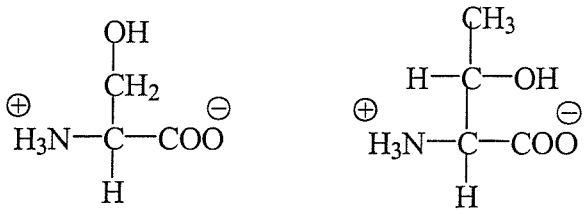


Tyr Phe

4.2.2 Ser 273 ⇒ Thr

As with Tyr 312, Ser 273 also appears to be involved in binding P-site ALA by hydrogen binding to the carboxyl group of the substrate. This residue was mutated to threonine (scheme 4.2) as the -OH group is retained but the additional methyl group may cause changes in interaction with -COOH group of ALA and in the binding of ALA with Tyr 312.

Scheme 4.2 Chemical structure of serine and threonine



Ser Thr

4.2.3 Tyr 192 ⇒ Phe

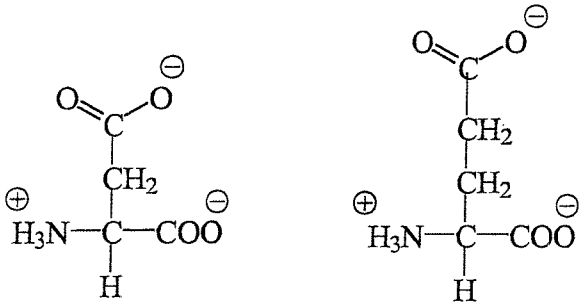
Tyr 192 is positioned near the P-site ALA binding site, with its hydroxyl group possibly forming a hydrogen bond with the amino group of bound P-site ALA. The mutation to phenylalanine was chosen to cause minimal structural disruption whilst removing the hydrogen bonding ability of the residue.

4.2.4 Asp118 ⇒ Glu

Asp118 is situated at the base of the active site, closely associated with Ser 165. The residue is well placed to interact with the amino group of P-site ALA and could possibly have a catalytic role in PBG formation. Glutamate was chosen as a suitable mutant residue (scheme 4.3) as it can also provide a carboxyl group at the base of the active site, but its increased side chain length would disrupt interactions between substrate ALA and the adjacent residue Ser 165.

A similar mutation was carried out in the study of the invariant *E. coli* PBG deaminase active site cleft residue Asp 84 (66). Substitution of this residue by glutamate resulted in a reduction of enzyme activity to 1 % of wild type activity, with the mutant being able to form stable enzyme intermediate complexes. However, mutation of asp 84 to alanine or asparagine resulted in complete loss of enzyme activity.

Scheme 4.3 Chemical structure of aspartate and glutamate

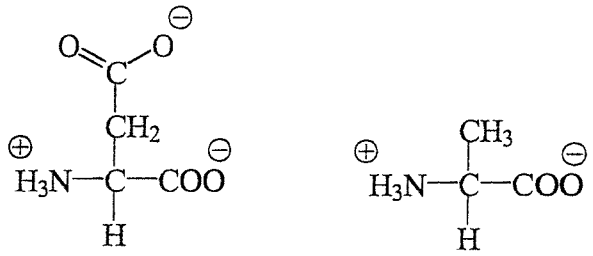


Asp Glu

4.2.5 Asp 118 ⇒ Ala

The rationale behind this mutation is similar to that for Asp118 ⇒ Glu. In this case however the carboxyl group at the base of the active site is removed (scheme 4.4) allowing further characterisation of the contribution of Asp 118 to ALA binding and catalysis.

Scheme 4.4 Chemical structure of aspartate and alanine

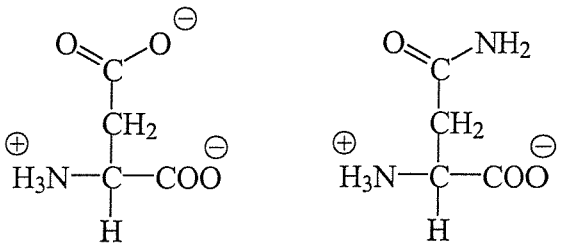


Asp Ala

4.2.6 Asp 118 ⇒ Asn

To characterise further the role of Asp 118 the residue was mutated to asparagine (scheme 4.5). Although unable to act as an acid or base, the ability of asparagine to donate and receive hydrogen bonds might allow the mutant residue to interact with P-site ALA and thus establish whether the role of Asp 118 is in substrate binding only.

Scheme 4.5 Chemical structure of aspartate and asparigine



Asp Asn

4.3 Description and rationale of A-site mutations.

4.3.1 Ser 165 ⇒ Thr

As previously mentioned Ser 165 is located at the base of the active site and is adjacent to Asp 118. The reduced ALAD-ALA + ALA structure shows the position of the amino group of A-site ALA in close proximity to Ser 165. The mutation of this residue to threonine could, through the introduction of an additional methyl group lead to disruption in the vicinity of Asp 118 and provide information about the role of the serine and aspartate interaction in binding substrate and catalysis.

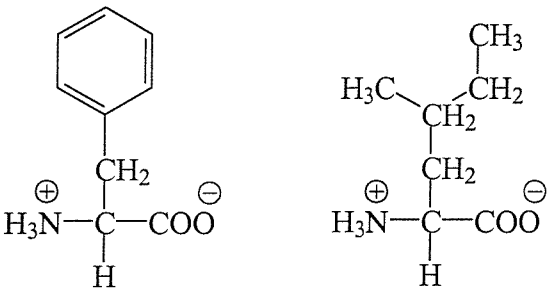
4.3.2 Ser 165 ⇒ Ala

The mutation of Ser 165 to alanine will prevent interaction between this residue and A-site ALA and also Asp 118. It is hoped that the characterisation of the Ser 165 and Asp 118 mutants will clarify the role of these closely associated and invariant residues.

4.3.3 Phe 204 ⇒ Ile

Phe 204 is located in the flexible loop appears to dock into active site on substrate binding. Whilst invisible in the native ALAD structure, the loop is clearly defined in ALAD-LA and reduced ALAD-ALA structures. On binding substrate or inhibitor at the P-site, Phe 204 packs against the methylene groups of the bound substrate molecule. This shields the substrate from solvent and brings in other residues from the flexible loop, forming an important contribution to the A-site ALA binding site. Mutation of Phe 204 to isoleucine (scheme 4.6) will partially retain the hydrophobic nature at this position and its characterisation give insight into its importance in anchoring the active site loop in place.

Scheme 4.6 Chemical structure of phenylalanine and isoleucine

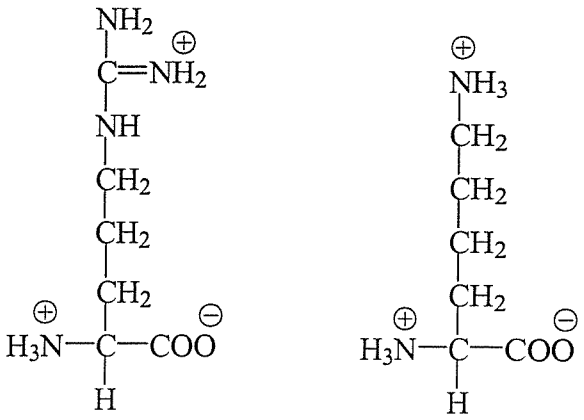


Phe Ile

4.3.4 Arg 205 ⇒ Lys

Arg 205 and Arg 215 form the binding site for the A-site ALA carboxyl group and are both located on the flexible loop covering the active site. It is thought that Arg 205 forms a salt bridge with the carboxyl group of the A-site substrate molecule. This interaction is of great importance as it ensures the binding of both molecules of ALA in a "head to head" formation which is critical for PBG formation. The mutation of this residue to lysine (scheme 4.7) will retain the positive charge of Arg 205 to the A-site ALA carboxyl group binding region.

Scheme 4.7 chemical structure of arginine and isoleucine



Arg Lys

All residues chosen for site mutagenesis are invariant throughout known ALAD sequences. An alignment of ALAD sequences can be seen in fig 4.1. All mutants were prepared by Dr. M. Sarwar and the fidelity of all DNA manipulations checked by DNA sequencing. All prepared mutants are shown in table 4.1

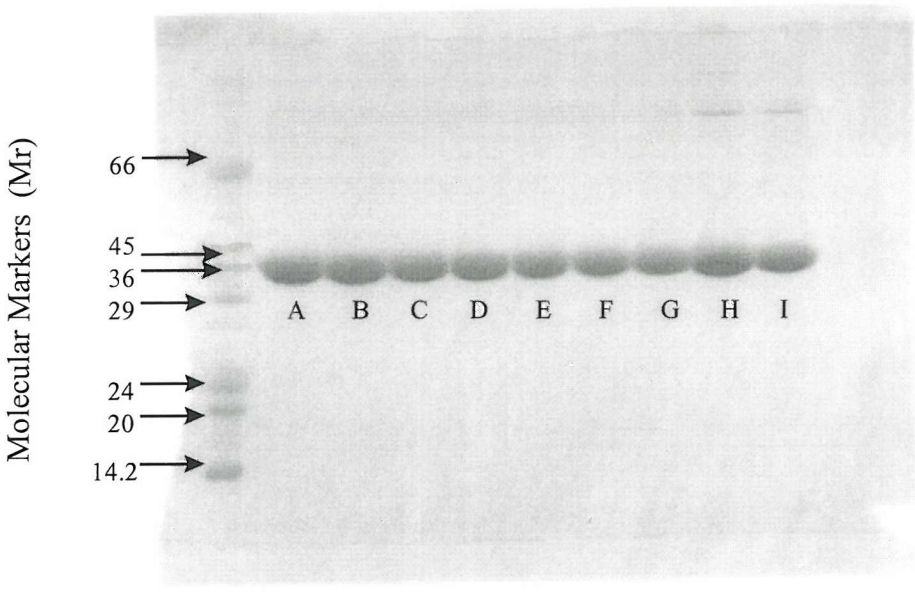
4.4 Purification and characterisation of *E. coli* ALAD mutants

Strains of *E. coli* expressing mutant ALAD enzymes were grown as previously described in materials and methods (chapter 2). The mutant enzymes were purified and their kinetic parameters assayed using the same method as for wild type (chapter 2). All enzymes were analysed by SDS-PAGE and appeared identical to wild type in terms of their mobility.(fig. 4.3)

4.5 Mass spectrometric analysis of *E. coli* ALAD mutants.

The purified ALAD mutants were buffer exchanged into 10 mM Tris/HCl pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol and the concentration of the enzyme was adjusted to 1 mg/ml. An equal volume of 50% acetonitrile, 48% water and 2% formic acid was added to the protein solutions. Typically a volume of 500 μl of enzyme was prepared and 20 μl was injected into an triple quadrupole electrospray mass spectrometer in positive ion mode. Data were acquired at a cone voltage of 40 V, capillary voltage of 3.5 kV and HV lens voltage of 0.30 kV between

Fig. 4.3 SDS-PAGE analysis of purified *E. coli* ALAD mutants run with Dalton VI markers. Gel shows all enzymes purified to > 95% homogeneity. The position of enzyme bands on the gel is consistent with *E. coli* ALAD monomer Mr 35,500.



Key to gel notation.

A. Wild type *E. coli* ALAD
B. *E. coli* ALAD Y312F mutant
C. *E. coli* ALAD S273T mutant
D. *E. coli* ALAD Y192F mutant
E. *E. coli* ALAD D118E mutant
F. *E. coli* ALAD S165T mutant
G. *E. coli* ALAD R215K mutant
F. *E. coli* ALAD F204I mutant

Table 4.2 Results of kinetic analysis of *E. coli* ALAD mutants

Mutant	K_m (mM)	V_{max} (μM PBG/mg/hr)	k_{cat} (S^{-1})	% Wild type k_{cat}
Wild type	1.9	36.8	0.356	100
P-site mutations				
Y312F	1.7	6.3	0.062	17
S273T	2.2	4.3	0.041	12
Y192F	2.3	21.2	0.205	61
D118A	2.2	0.03	2.7×10^{-4}	< 1
D118E	1.8	7.1	0.068	11
D118N	2.5	0.06	5.4×10^{-4}	<1
A-site mutations				
S165T	2.4	3.7	0.0358	10
S165A	1.8	1.1	0.01	4
R215K	7.9	25.7	0.249	70
F204I	17.6	12.9	0.125	35

4.4 Implications of mutant analysis.

4.4.2 Analysis of P-site mutations

Tyr 312 ⇒ Phe

The proposed role of this residue is to form part of the ALA P-site of the active site cavity. The mutation of tyrosine to phenylalanine causes the loss of the hydroxyl group involved in a hydrogen bond with the carboxyl group of P-site ALA. The position of ALA at the P-site is vital, since the carbonyl group of the substrate molecule must be adjacent to the ε-nitrogen of Lys 247 to facilitate formation of the initial Schiff base. The large reduction in enzyme activity, a decrease to 17% of wild type activity, caused by this mutation indicates the importance of Tyr 312 in P-site ALA binding. Significantly, although a large reduction of enzyme activity was observed, there was no alteration in the K_m value in comparison to wild-type indicating that the P-site has a higher affinity for ALA than the A-site. This in agreement with observations made by Spencer *et al.*, (18) who investigated the ALA binding properties of the A-site ALA binding site of *E. coli* ALAD using rate of dialysis analysis with a reduced Schiff base adduct. Details of mass spectrometric and kinetic analysis are shown in figures 4.6 and 4.7 and table 4.2.

Ser 273 ⇒ Thr

As with Tyr 312, the role of Ser 273 is to form the binding site for the carboxyl group of P-site ALA. The mutation of this residue to threonine causes significant reduction in enzyme activity to 12% of wild type activity, indicating poorer substrate binding. It is possible that the addition of a larger residue at this position may sterically hinder the binding of ALA at the P-site as well as alter interactions between the -OH group of the residue and the carboxyl group of P-site ALA. Details of mass spectrometric and kinetic analysis are shown in figure 4.8 and table 4.2.

Fig 4.4 Mass spectrometric analysis of wild type *E. coli* ALAD. Data shows Mr of 35495 ± 2.64 which is in agreement with the expected Mr of 35497.

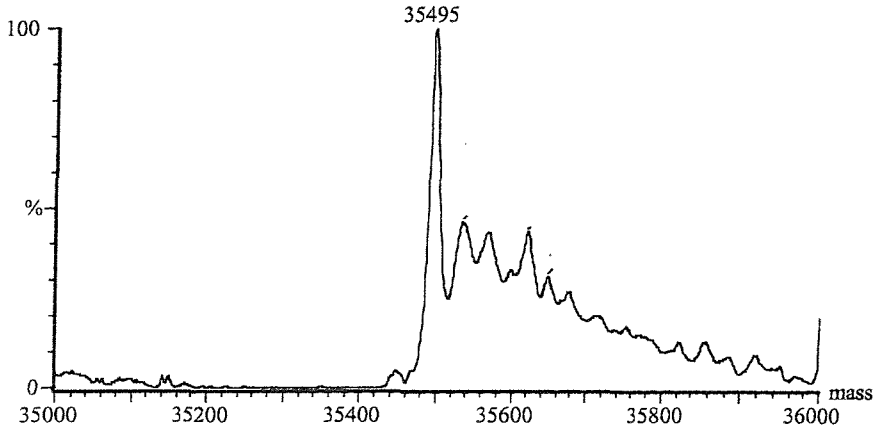


Fig 4.5 Graph showing kinetic analysis of *E. coli* ALAD.
Km = 1.9 ± 0.22 mM; Vmax = 36.8 ± 1.5 μmoles PBG/mg/hr.

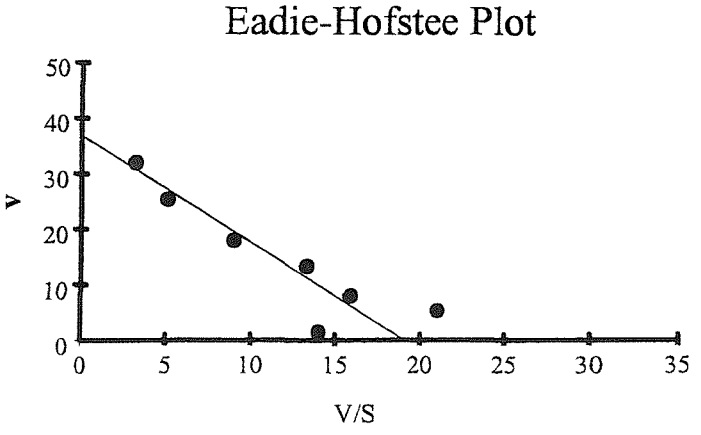


Fig.4.6 Mass spectrometric analysis of *E. coli* ALAD mutant Y312F. Data shows Mr of 35481 ± 3.19.which is in agreemant with the expected Mr of 35482.

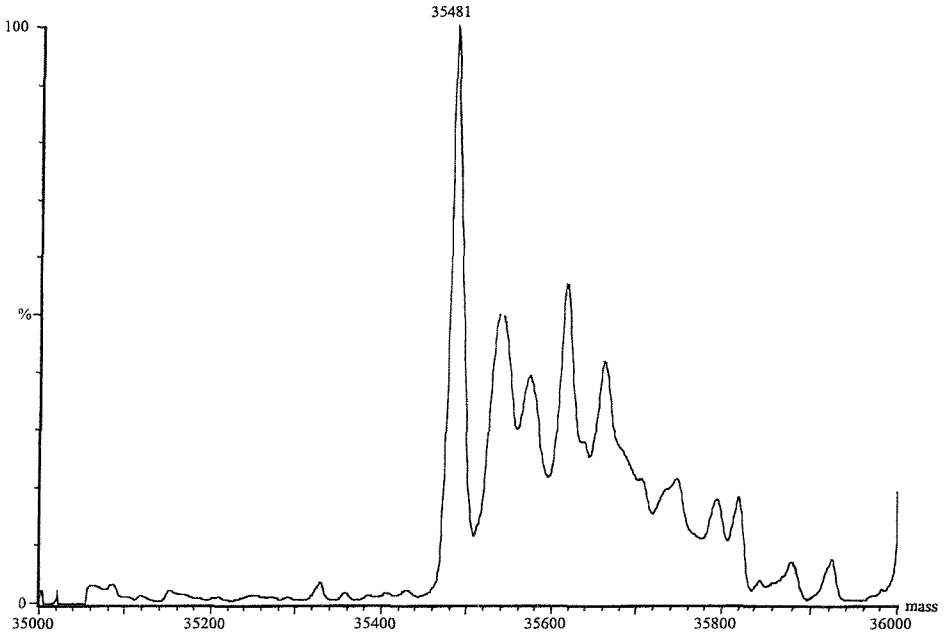


Fig. 4.7 Graph showing kinetic anaylsis of *E. coli* ALAD mutant Y312F. K_m = 1.7 ± 0.4 mM; V_{max} = 6.3 ± 0.5 μmoles PBG/mg/hr.

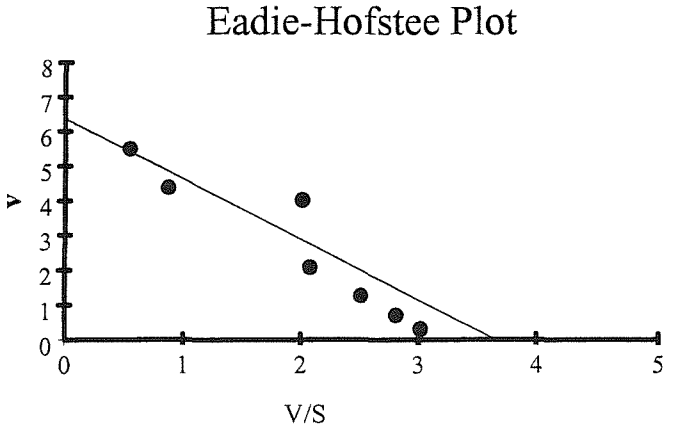
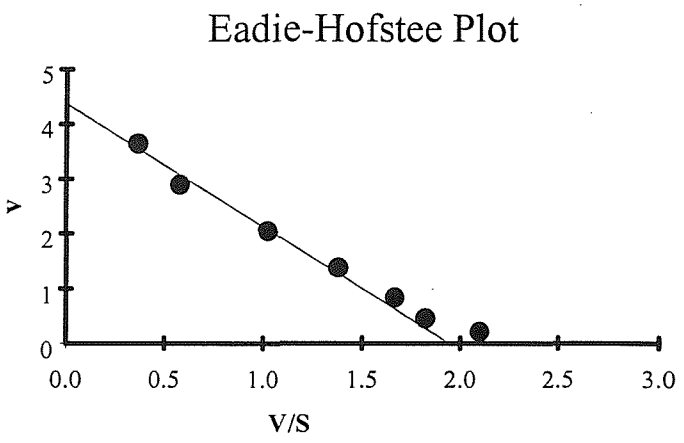


Fig. 4.8. Graph showing kinetic analysis of *E. coli* ALAD mutant S273T. $K_m = 2.2 \pm 0.18$ mM; $V_{max} = 4.3 \pm 0.109$ μmoles PBG/mg/hr



Tyr 192 ⇒ Phe

Although mutation of this residue causes significant loss in activity, decreased to 61% of wild type activity, it is not thought that the residue directly binds the amino group of the P-site ALA substrate molecule. The importance of this invariant residue is thought to be its contribution to the highly polar active site cavity. Details of mass spectrometric and kinetic analysis are shown in figures 4.9 and 4.10 and table 4.2.

Asp 118 ⇒ Ala

Mutation of Asp 118 to alanine had a dramatic effect on enzyme activity, reducing activity to less than 1% of wild type. The loss of the carboxyl group would prevent H-bond formation with the amino group of P-site. It is thought that the interaction between Asp 118 and the amino group of P-site ALA is necessary for Schiff base formation between the two substrate molecules. The role of this residue in the catalysis of PBG formation is further discussed in chapter 6. Details of mass spectrometric and kinetic analysis are shown in figures 4.11 and 4.12 and table 4.2.

Asp 118 ⇒ Glu

Mutation of aspartate 118 to glutamate reduced activity of the enzyme to 11% of wild type activity. The relatively higher levels in activity in comparison to the Asp 118 ⇒ Ala mutant suggest that the mutant glutamate residue, despite its extra size is able to function partly in the same role as aspartate. Details of mass spectrometric and kinetic analysis are shown in figures 4.13 and 4.14 and table 4.2.

Asp 118 ⇒ Asn

Mutation of Asp 118 to asparagine reduced enzyme activity to less than 2% of wild type activity. The results of this mutant analysis emphasis the importance of the carboxyl group of Asp 118. Details of mass spectrometric and kinetic analysis are shown in figures 4.15 and 4.16 and table 4.2.

Fig.4.9 Mass spectrometric analysis of *E. coli* ALAD mutant Y192F. Data shows Mr of 35482 ± 2.97 which is in agreement with the expected Mr of 35482.

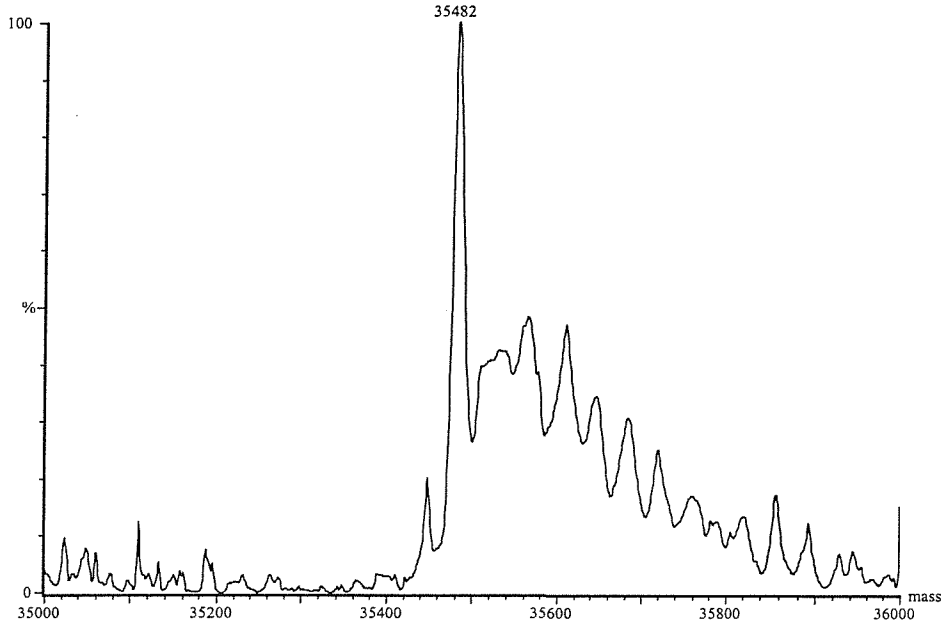


Fig. 4.10 Graph showing kinetic analysis of *E. coli* ALAD mutant Y192F. K_m = 2.3 ± 0.29 mM V_{max} = 21.2 ± 0.9 μmoles PBG/mg/hr.

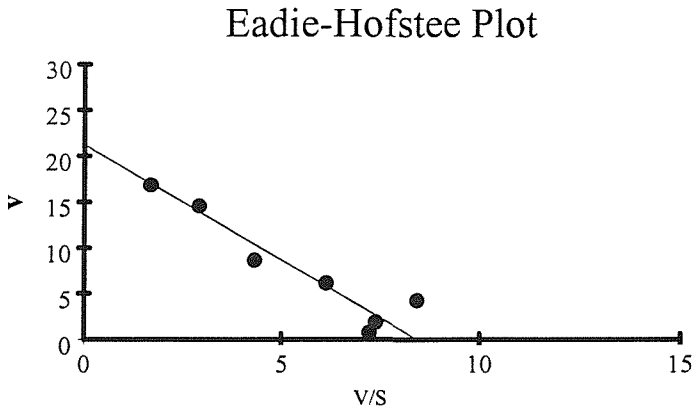


Fig.4.11 Mass spectrometric analysis of *E. coli* ALAD mutant D118E. Data shows Mr of 35508 ± 2.37 which is agreement with the expected Mr of 35510.

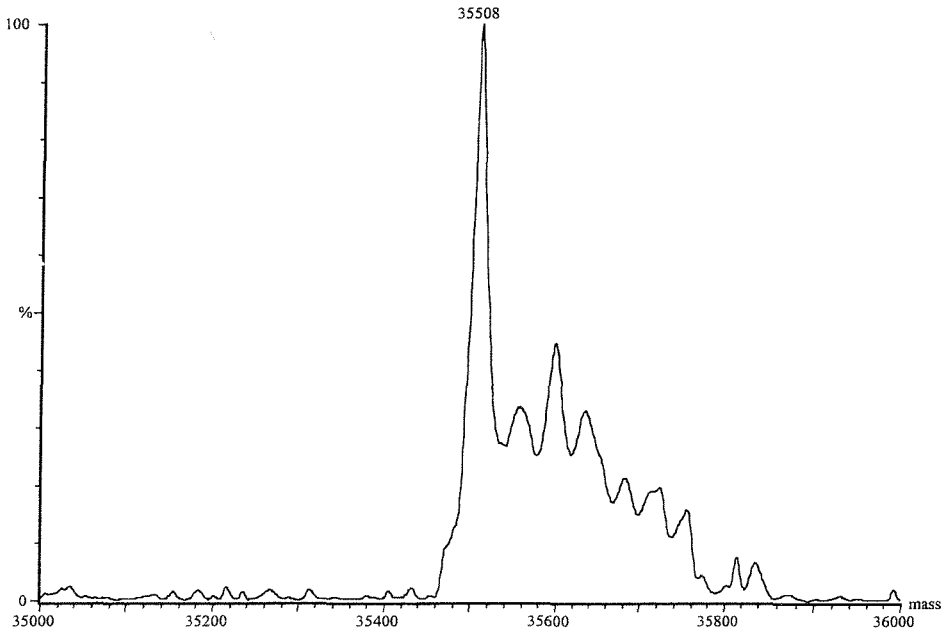


Fig. 4.12 Graph showing kinetic analysis of *E. coli* ALAD mutant D118E. K_m = 1.8 ± 0.2 mM; V_{max} = 7.1 ± 0.28 μmoles PBG/mg/hr.

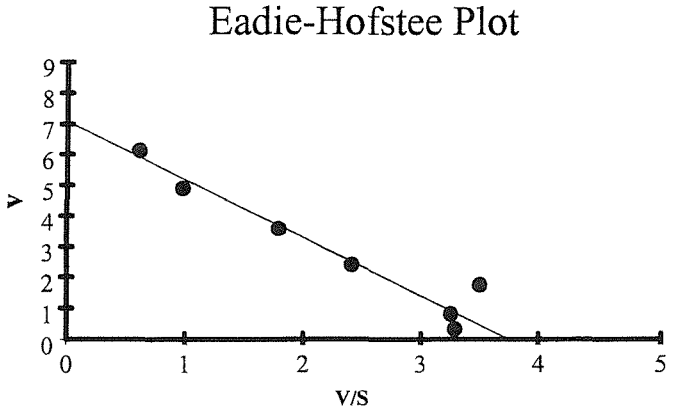


Fig.4.13 Mass spectrometric analysis of *E.coli* ALAD mutant D118A. Data shows Mr of 35454 ± 1.12 which is in agreement with the expected mass of 35453

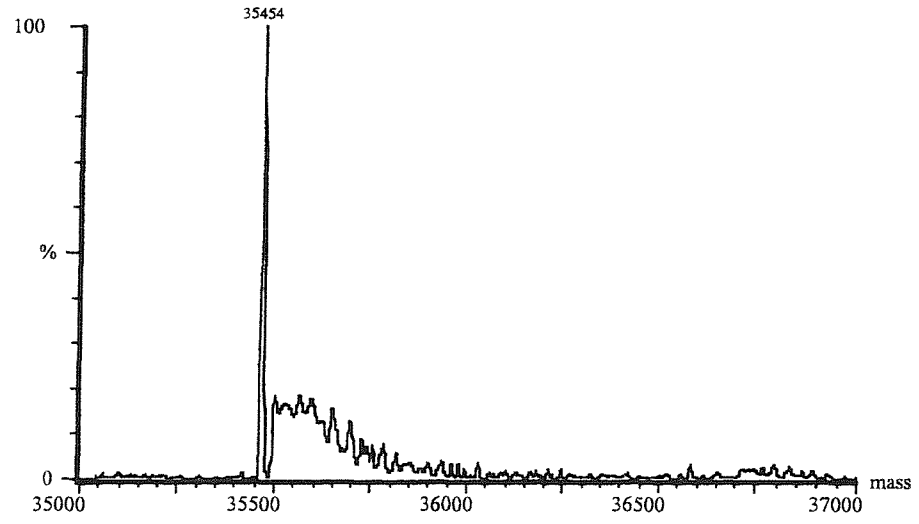


Fig. 4.14 Graph showing kinetic analysis of *E. coli* ALAD mutant D118A. K_m = 2.2 ± 0.3 mM ; V_{max} = 0.03 ± 0.01 μmoles PBG/mg/hr.

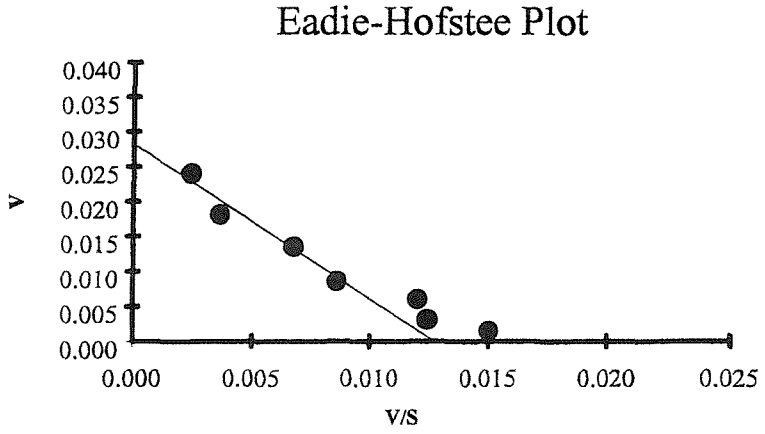


Fig.4.15 Mass spectrometric analysis of *E. coli* ALAD mutant D118N. Data shows Mr of 35496 ± 0.57 which is in agreement with the expected Mr of 35497.

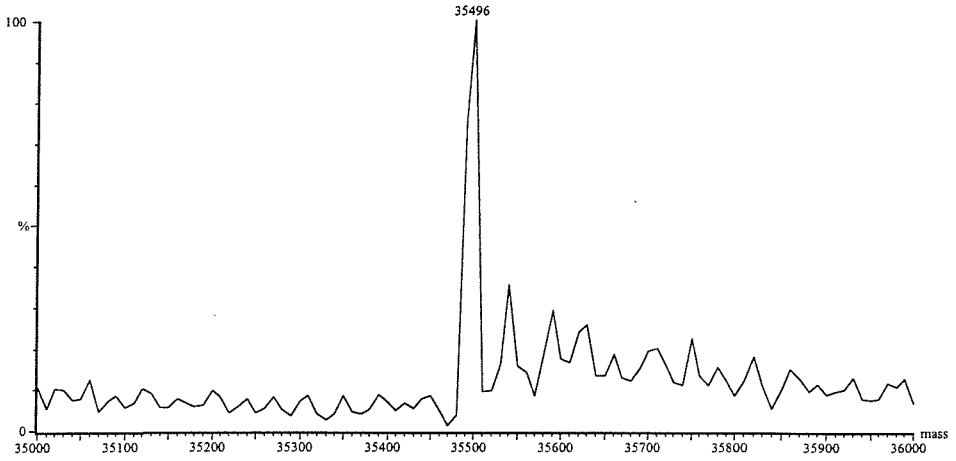
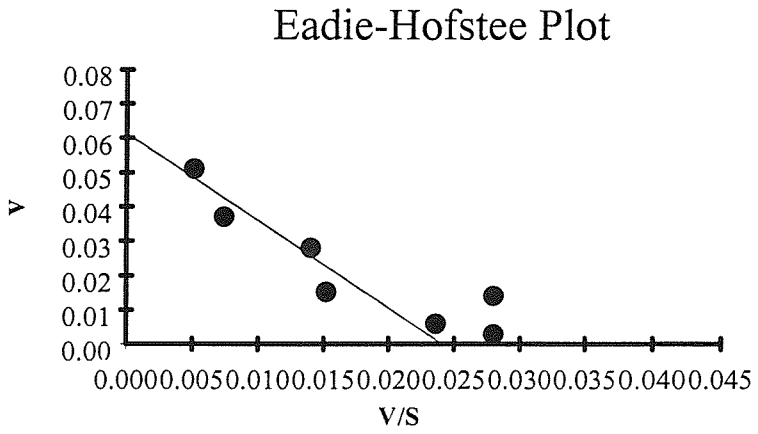


Fig. 4.16. Graph showing kinetic anaylsis of *E. coli* ALAD mutant D118N. $K_m = 2.5 \pm 0.2$ mM ; $V_{max} = 0.06 \pm 0.02$ μmoles PBG/mg/hr



4.4.2 A-site mutations

Ser 165 ⇒ Ala and Ser 165 ⇒ Thr

The mutation of Ser 165 caused a decrease in enzyme activity to 4% (Ser 165 ⇒ Ala) and 10 % (Ser 165 ⇒ Thr). In comparison to wild type no increase in K_m is observed in either case. It is probable that this residue does not form any significant bond with the amino group of the A-site substrate molecule. The loss in enzyme caused by the mutation of this residue can be attributed to disruption of interactions with Asp 118 which is involved co-ordinating the amino group of P-site ALA. Details of mass spectrometric and kinetic analysis are shown in figures 4.17 and 4.18 (Ser 165 ⇒ Ala) and 4.19 and 4.20 (Ser 165 ⇒ Thr), and table 4.2.

Arg 205 ⇒ Lys

The loss in activity to 70% of wild type, and increase in K_m from 1.93 mM in wild type to 7.8 mM is consistent with the proposal that this residue forms a salt bridge with that carboxyl group of the A-site substrate molecule. The loss in activity is fairly small however, indicating that the positively charge lysine residue may be a reasonable substitute for arginine. The increase in K_m value demonstrates the lowered affinity for ALA of the mutated A-site ALA binding site. Details of mass spectrometric and kinetic analysis are shown in figures 4.21 and 4.22 and table 4.2.

Phe 204 ⇒ Ile

Once substrate ALA is bound at the P-site, Phe 204 packs against the methylene groups of the bound molecule. The result of this interaction is to shield the substrate from solvent water and to stabilise the flexible loop covering the active site which contributes to the A-site ALA binding. The mutation of this residue causes a loss in enzyme activity to 35% of wild type and an increase in K_m to 17.6 mM. It is probable therefore that the decreased hydrophobic interaction between Ile 204 and the bound P-site substrate molecule causes an increase in active site loop mobility which destabilises the A-site ALA binding site. The loop region contains residues Arg 204 and Arg 215 which are both involved in forming the A-site ALA carboxyl group binding region, causing a decrease in affinity for ALA and an increase in K_m

value. Details of mass spectrometric and kinetic analysis are shown in figures 4.23 and 4.24 and table 4.2.

Fig.4.17 Mass spectrometric analysis of *E.coli* ALAD mutant S165A. Data shows Mr of 35480 ± 3.12 Da. which is in agreement with expected Mr of 35482.

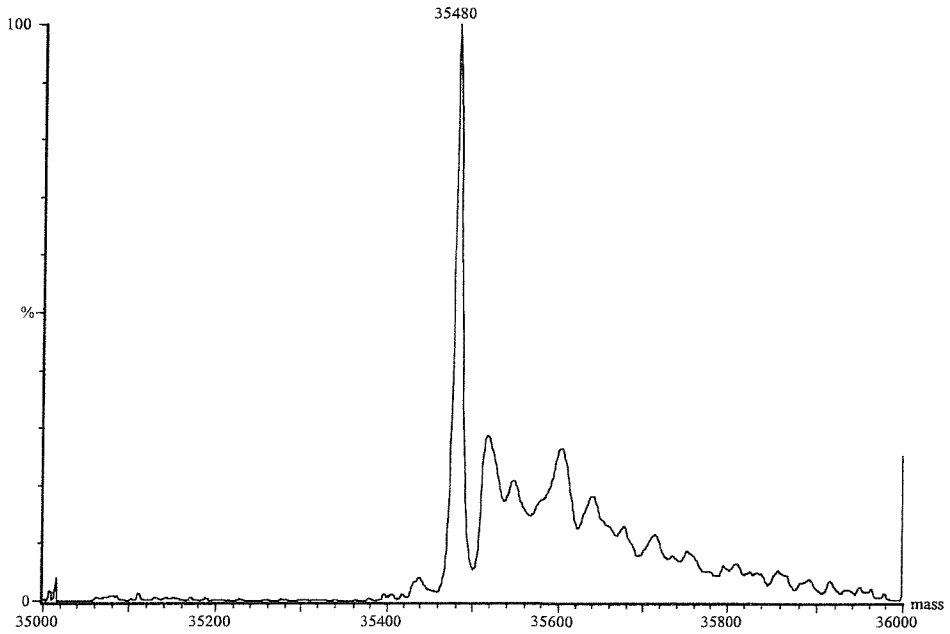


Fig. 4.18. Graph showing kinetic analysis of *E. coli* ALAD mutant S165A. K_m = 1.8 ± 0.3 mM; V_{max} = 1.1. ± 0.07 μmoles PBG/mg/hr

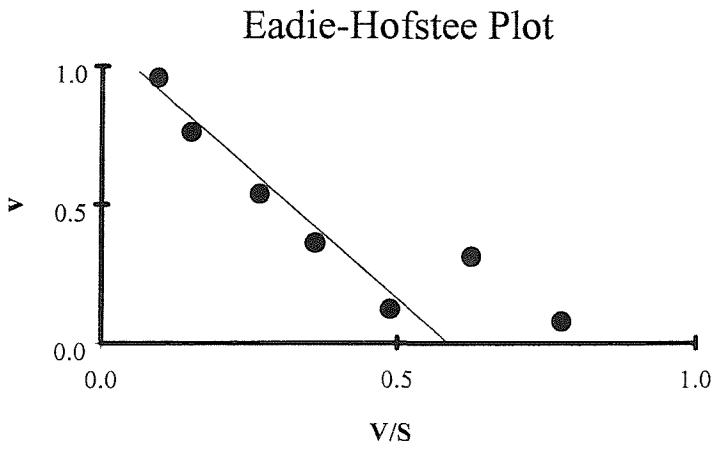


Fig.4.19 Mass spectrometric analysis of *E.coli* ALAD mutant S165T. Data shows Mr of 35510 ± 2.54.which is in agreement with expected Mr of 35512.

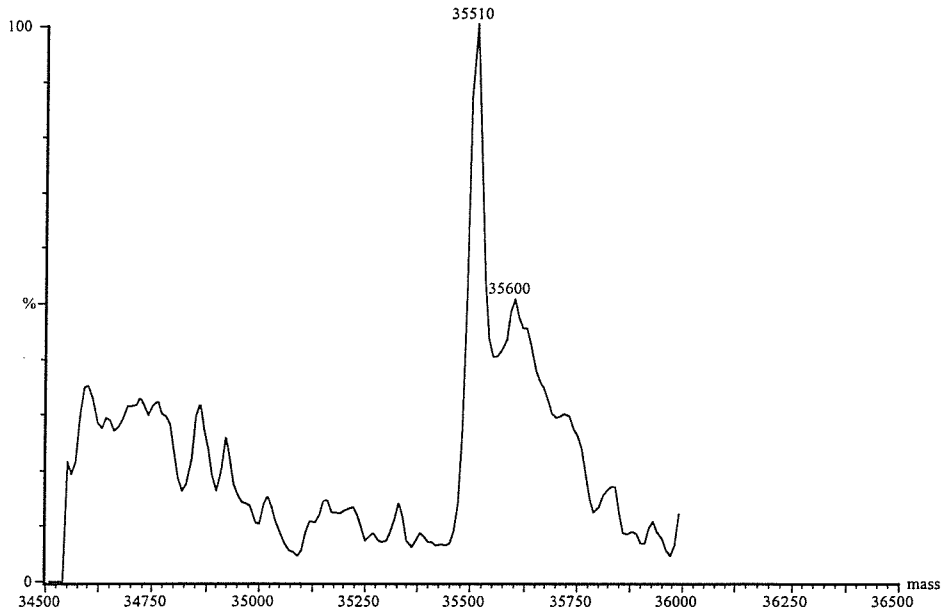


Fig. 4.20. Graph showing kinetic analysis of *E. coli* ALAD mutant S165T. K_m = 2.2 ± 0.2 mM ; V_{max} = 4.3 ± 0.12 µmoles PBG/mg/hr.

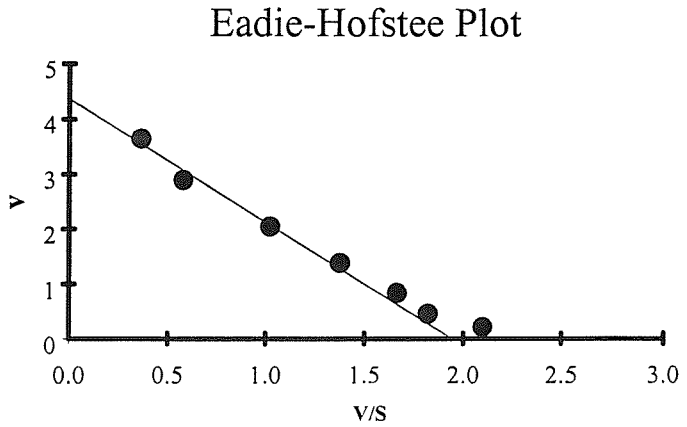


Fig.4.21 Mass spectrometric analysis of *E.coli* ALAD mutant R204K. Data shows Mr of 35466 ± 3.42 which is in agreement with expected Mr of 35469.

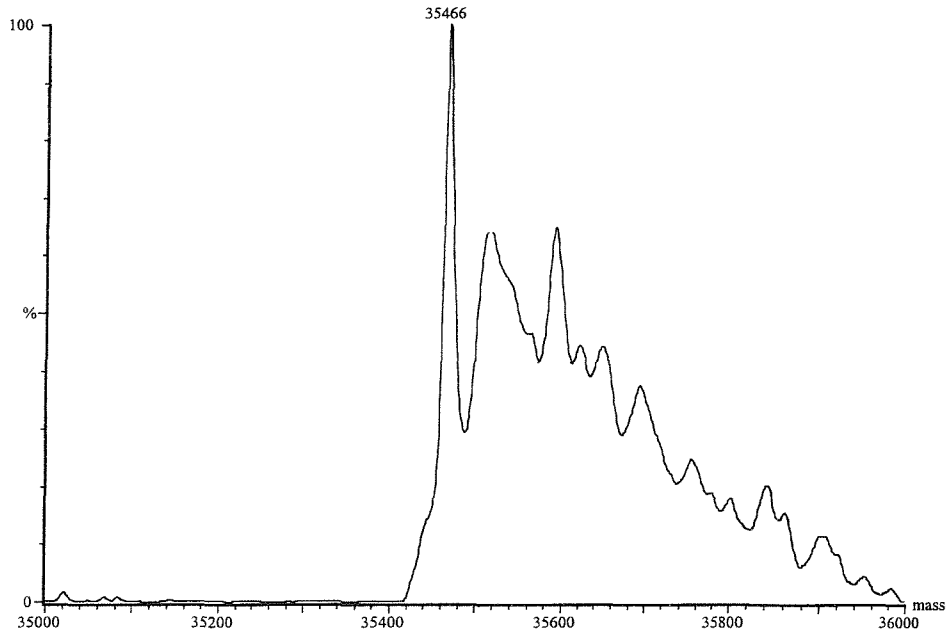


Fig. 4.22 Graph showing kinetic analysis of *E. coli* ALAD mutan R205K. K_m = 7.9± 0.39 mM; V_{max} = 25.7 ± 0.72 μmoles PBG/mg/hr

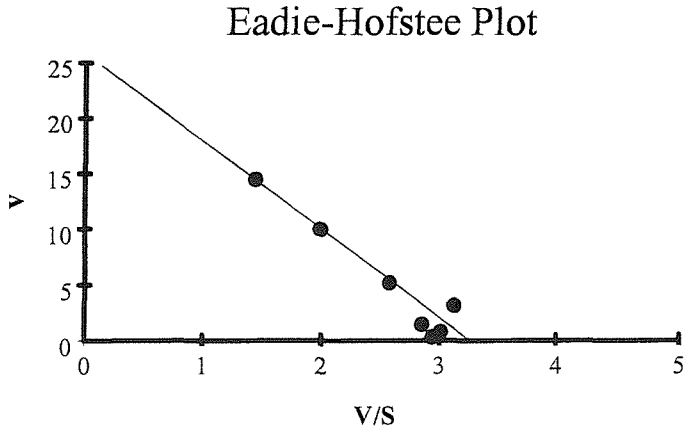


Fig.4.23 Mass spectrometer analysis of *E.coli* ALAD mutant F204I. Data shows Mr of 35465 ± 1.07 which is in agreement with expected Mr of 35464.

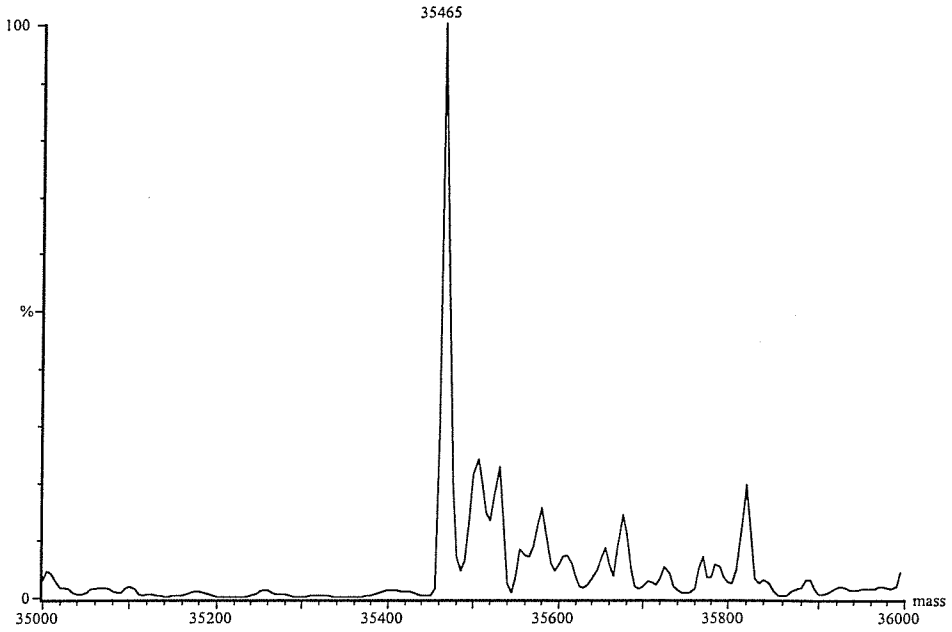
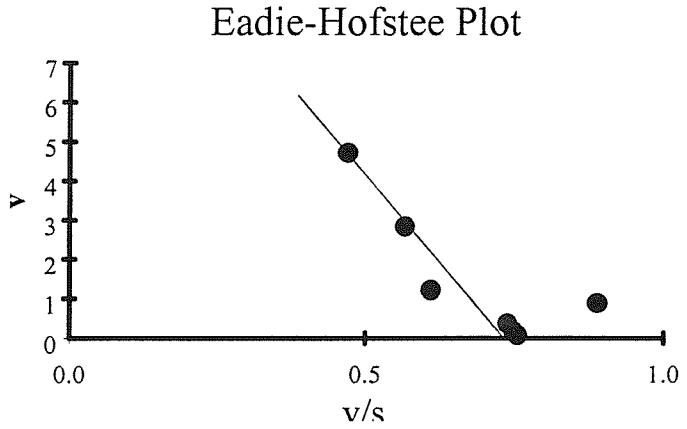


Fig. 4.24. Graph showing kinetic analysis of *E. coli* ALAD mutant F204I. K_m = 17.6 ± 3.5 mM; V_{max} = 12.9 ± 1.8 μmoles PBG/mg/hr.



Chapter 5. Investigation into the contribution of lysine 247 and lysine 195 in the mechanism of PBG formation catalysed by *E. coli* ALAD.

Following X-ray analysis of *E. coli* ALAD, detailed consideration of the residues in the active site suggested the apparent mechanistic importance of two adjacent lysine residues, Lys 247 and Lys 195. These two residues thus became the subject of a comprehensive investigation. Although the function of Lys 247 in forming a Schiff base with the P-side ALA has been established (61), the absolute importance of this residue in PBG formation and the influence of the adjacent Lys 195 was unclear. The roles of both these residues have been investigated using a combination of site directed mutagenesis, chemical modification and X-ray crystallography. The first step of this investigation involved the generation of Lys 247 and Lys 195 mutants in order to assess the importance of these residues to enzyme activity. The mutants studied in this chapter were kindly generated and characterised by sequencing by Dr. Muhammad Sarwar.

Three mutations of Lys 247 were made K247A, K247C and K247N. The rationale for generating the three mutants was as follows: K247A represents a mutation in which a small, relatively uncontroversial residue is substituted. Alanine has been used in a large number of mutations in a variety of studies. The mutation to cysteine was designed for future modification with bromoethylamine to create a "pseudolysine" side chain. The mutation to asparagine was designed to introduce an amide $-NH_2$ potentially to hydrogen bond to the P-site substrate carbonyl oxygen in a spatially related way to the lysine Schiff base.

5.1 Purification and characterisation of *E. coli* Mutants K247A, K247C and K247N

Mutant *E. coli* ALAD genes were expressed in a hem B^- TB1 strain of *E. coli* as previously described in materials and methods. All mutants were grown in the presence of hemin to ensure full growth. The mutant enzymes were purified and their kinetic parameters were assayed using the same method as for wild type (see

methods). All enzymes were analysed by SDS-PAGE and appeared identical to wild type.

5.1.1 Mass spectrometric analysis of *E. coli* ALAD mutants

The purified ALAD mutants were buffer exchanged into 10 mM Tris/HCl buffer pH 7.5 containing 50 μM and 10 mM β-mercaptoethanol. The concentration of the enzymes was adjusted to 1 mg/ml and an equal volume of 50% acetonitrile, 48% water and 2% formic acid solution was added to the protein solutions. Typically 500 μl of enzyme was prepared. A volume of 20 μl was injected into a triple quadrupole electrospray mass spectrometer in positive ion mode. Data were acquired at a cone voltage of 40 V, capillary voltage of 3.5 kV and HV lens voltage of 0.30 kV between masses of 500 m/z and 2000 m/z. The raw data obtained from the above values were then subjected to maximum entropy analysis according to the Micromass schedule. Analysis of mutants K247A and K247C gave Mr.'s of 35445 ± 3.16 (fig 5.1) and 35477 ± 2.57 (fig 5.2) respectively. These masses are accurate to the predicted mass of the mutants within the tolerances of the equipment used. The mass spectrometer analysis of mutant K247N will be discussed later in this chapter.

Fig. 5.1 Mass spectrometric analysis of *E.coli* ALAD mutant K247A. Data shows a Mr. of 35445 ± 3.16 Expected Mr = 35441

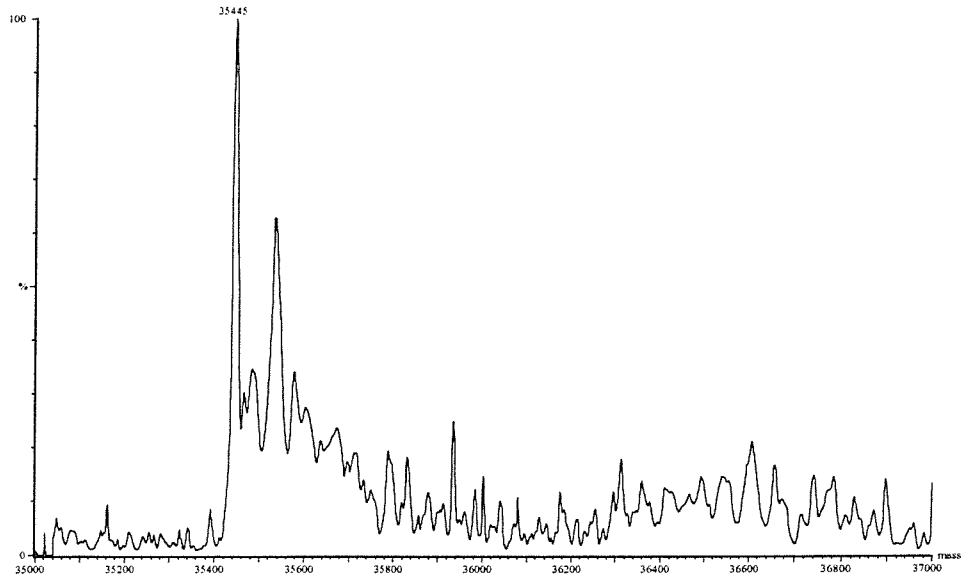
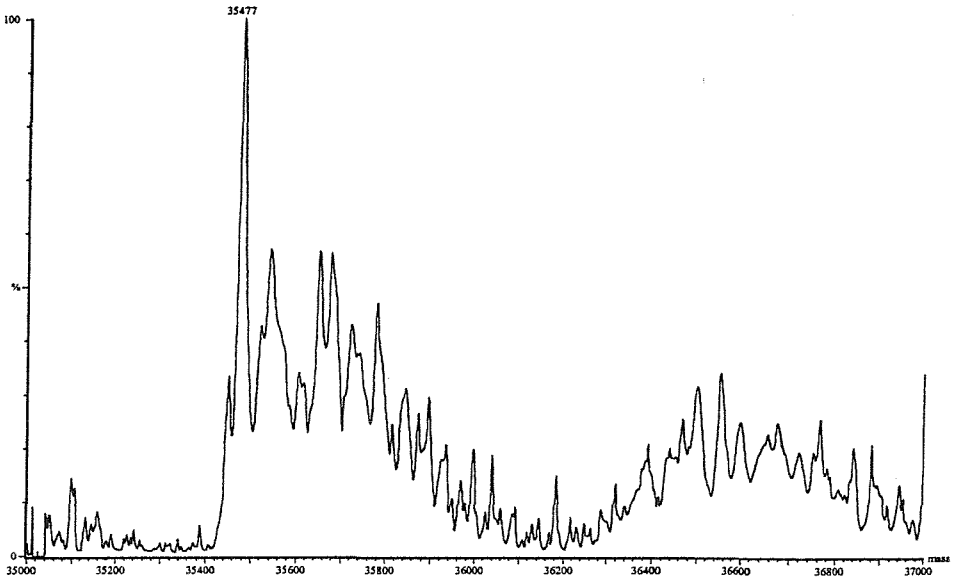


Fig.5.2 Mass spectrometer analysis of *E.coli* ALAD mutant K247C. Data shows a mass of 35477 ± 2.57 Da. Expected mass = 35473 Da



5.1.2 Kinetic analysis of *E. coli* ALAD mutants K247A, K247C and K247N.

Mutants K247A and K247C proved to be completely inactive and no kinetic parameters could be determined. The inability of these mutants to form a Schiff base with ALA demonstrates how crucial this first step is in the ALAD catalysed formation of PBG.

Although unable to form a Schiff base with ALA, the K247N mutant was not completely inactive with 0.01% of activity, with respect to the wild type, remaining. This activity was also evident during the growth of the mutant, since haemin was not vital for successful growth whereas mutants K247A and K247C were unable to grow unless haemin was present in the media. A comparison of mutant growth in the absence of haemin versus the activity of the mutant is shown in table 5.1. The mutant K247N has been co-crystallised with laevulinic acid and

Table 5.1 Growth of ALAD mutants, relative to wild type, in presence and absence of haemin, where + + + denotes normal growth and - - - denotes complete absence of bacterial growth.

ALAD mutant	**Bacterial growth + Haemin**	**Bacterial growth -Haemin**	**relative activity**
wild type	+ + +	+ + +	100 %
K247A	+ + +	- - -	0 %
K247C	+ + +	- - -	0 %
K247N	+ + +	+ - -	0.01 %
K195A	+ + +	+ + -	0.1 %
K195C	+ + +	+ + -	0.1 %
D118A	+ + +	+ + +	1 %

the X-ray structure solved. The structure provides evidence to suggest that the amide -NH_2 group group of the mutant Asn 247 residue is capable of interacting with the carbonyl oxygen of P-site ALA *via* a hydrogen bond and facilitating PBG formation. The structure of this K247N mutant and its characterisation are discussed later in this chapter.

5.1.3 Purification and characterisation of *E. coli* ALAD mutants K195A and K195C

The rationale for the formation of the Lys 195 mutants K195A and K195C is similar to that of the K247A and K247C mutants. K195A replaces Lys 195 with the small and unreactive alanine residue whilst the mutation to cysteine was designed to allow future modification of the side chain with bromoethylamine.

The purification, mass spectrometer analysis and kinetic analysis of *E. coli* ALAD mutants K195A and K195C were carried out as described for the K247A, K247C and K247N mutants. Both K195A and K195C required growth media supplemented with hemin to achieve full growth. (table 5.1).

5.1.4 Results and discussion of *E. coli* ALAD K195 mutant analysis

Both K195A and K195C showed very low levels of activity with values of approximately 0.1% of that of wild-type. It is probable that the absence of the positively charged amino group of Lys 195 allows Lys 247 to exist in a protonated state at physiological pH. In this protonated state, Lys 247 would be insufficiently nucleophilic to form a Schiff base with the carbonyl carbon of the ALA substrate molecule at the P-site. The effect of Lys 195 on the pK_a of Lys 247 has been investigated further and is discussed later in this chapter.

Fig.5.3 Mass spectrometric analysis of *E.coli* ALAD mutant K195A. Data shows a Mr of 35446 ± 3.08 Da. Expected Mr = 35441 Da.

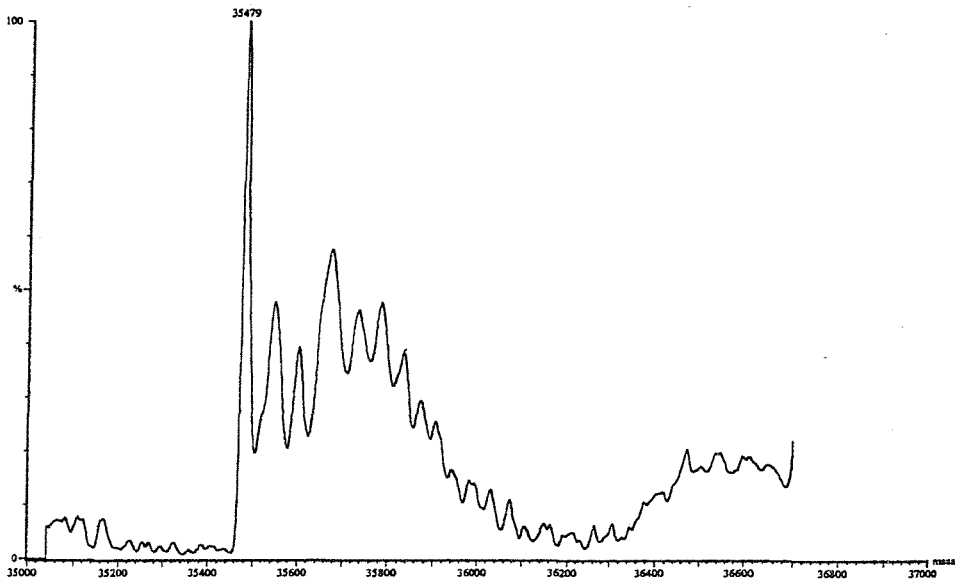
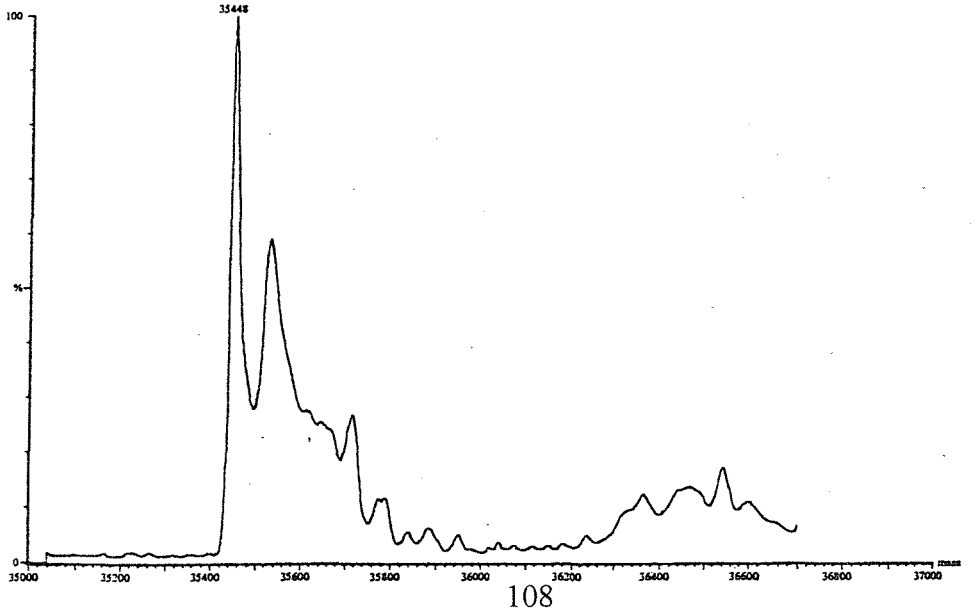


Fig. 5.4 Mass spectrometric analysis of *E.coli* ALAD mutant K195C. Data shows a Mr of 35476 ± 2.67. Expected Mr = 35473



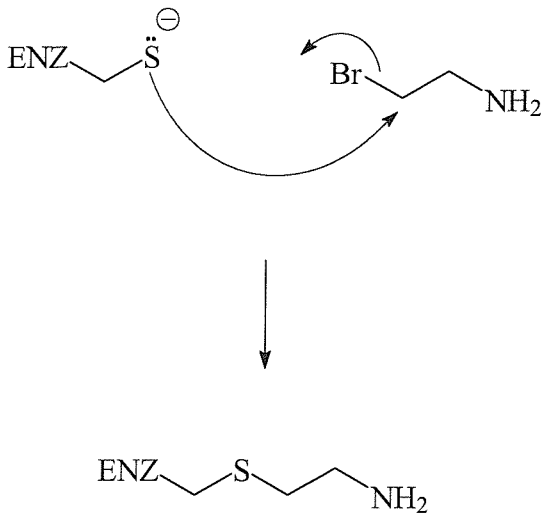
5.2 Alkylation of wild type *E. coli* ALAD, and *E. coli* ALAD mutant K247C and *E. coli* ALAD mutant K195C with 2-bromoethylammonium bromide.

The *E. coli* ALAD mutants K247C and K195C were purified and characterised (section 5.1) revealing a dramatic drop in activity for both enzymes, and emphasising the vital role played by both Lys 247 and Lys 195 in PBG formation.

As mentioned above, cysteine mutants had been generated with a view to designing "rescue " experiments to regenerate a lysine equivalent by alkylation with bromoethyl ammonium bromide. In similar experiments Highbarger *et. al.* (67) regenerated acetoacetate decarboxylase mutants, K116C and K115C, with the alkylating agent 2-bromoethyl-ammonium bromide (2-BEAB). As with ALAD, both of these lysine residues are essential and invariant and are thought likely to be adjacent to each other in the active site. Lys 115 of AAD is known to form a Schiff base with the acetoacetate substrate molecule in the initial enzyme-substrate complex. All known mutations of this residue result in complete inactivation of the enzyme. Lys 116 of AAD is thought to suppress the pK_a of Lys 115, allowing it to exist in its unprotonated $-NH_2$ form which is vital for Schiff base formation. Although not completely crucial for enzyme activity, mutation of Lys 116 to cysteine reduces enzyme activity to 2% of wild type levels. Alkylation of AAD mutants K115C and K116C with 2-BEAB restored their activity to 8% and 55% of wild type activity respectively.

As the proposed functions of Lys 115 and Lys 116 in AAD are so analogous to those predicted for Lys 247 and Lys 195 in *E. coli* ALAD, restoration of mutants K247C and K195C with 2-BEAB was attempted (scheme 5.1).

Scheme 5.1 Alkylation of cysteine by 2-bromoethylammonium bromide (2-BEAB). The side chain of the cysteine attacks the 2-BEAB molecule, displacing the Br as Br^- leading to the attachment of an ethylamine group to the sulphur of the cysteine residue. This forms a thialysine group enabling the modified cysteine to act as a "pseudo lysine" residue with the potential of restoring enzyme activity.



5.2.1 Modification of wild type *E. coli* ALAD and mutants K247C and K195C with 2-BEAB.

Unlike with AAD, the treatment of *E. coli* ALAD with 2-BEAB posed the threat of alkylating the cysteine residues at the A-side of the active site that bind the essential zinc ion. Alkylation of these residues is known to result in loss of Zn^{2+} from the active site and prevent any activity gained from modification of the mutant cysteine residue being detected due to enzyme deactivation. After a series of experiments with a range of concentrations of 2-BEAB being added to mutant enzyme solution it was found that maximum rescue of activity occurred with 2-

BEAB being added to enzyme solution in an 1:1 molar ratio i.e. one molecule of 2-BEAB per monomer of ALAD. This level of addition was shown to have negligible effect on the activity of wild type ALAD whereas increased levels of addition caused progressively lower levels of restoration. Development of these optimum modification conditions were carried out by Lucy Turner, a 3rd year project student under the authors supervision.

Wild type *E. coli* ALAD and mutants K247C and K195C were concentrated to 10 mg/ml in 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$, using an Amicon 10 ml concentrating cell with an Amicon PM-10 membrane. stoichiometric amounts of 2-BEAB were mixed with the enzyme solution by the addition of freshly prepared 20 mM 2-BEAB in 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$. After 12 hours the enzyme solution underwent buffer exchange with a Sephadex G-25M column in order to remove any unreacted 2-BEAB into 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ 10mM β-mercaptoethanol. The aminoethylated ALADs designated WT-EA, K247C-EA and K195C-EA were subjected to kinetic analysis, as previously described in this thesis. Results of these analysis can be seen in table 5.2 and figs. 5.5 and 5.6

Table 5.2 kinetic analysis of WT-EA, K247C-EA, and K195C-EA modified ALAD

Enzyme	V_{max} (μMPBG/hr./mg)	K_m (mM)	K_{cat} (S^{-1})
WT	36.8	1.9	0.356
WT-EA	34.62	2.1	0.335
K247C	--	--	--
K247C-EA	1.8	2.9	0.02
K195C	0.04	9.6	5×10^{-4}
K195C-EA	3.4	2.3	0.033

Fig. 5.5. Graph showing kinetic anaylsis of modified *E. coli* ALAD mutant K247C-EA. K_m = 2.9 ± 0.22 mM; V_{max} = 1.8 ± 0.15 μmoles PBG/mg/hr.

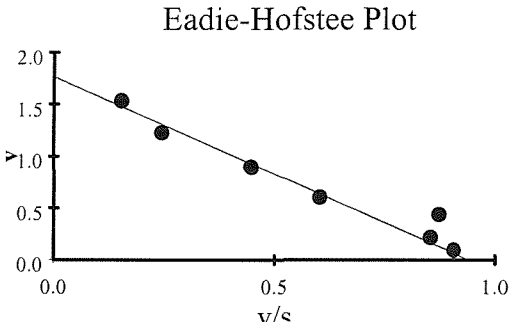
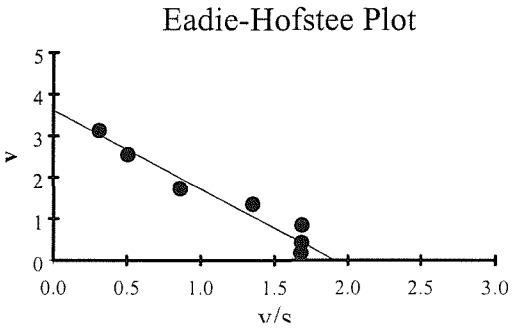


Fig. 5.6. Graph showing kinetic anaylsis of modified *E. coli* ALAD mutant K195C-EA. K_m = 2.3± 0.2 mM; V_{max} = 3.37 ± 0.068 μmolesPBG/mg/hr.



5.2.2 Results and discussion of rescue experiment

5.2.3 Restoration of K247C mutant.

As previously discussed in section 5.1 the *E. coli* K247C mutant possessed no detectable activity revealing the vital role of Lys 247 in ALAD-catalysed PBG formation. However, after modification with 2-BEAB the enzyme activity was restored to 5% that of wild type ALAD. Under the same modification conditions the wild type enzyme showed no significant loss in activity.

The restored activity demonstrates that modification of the mutant cysteine residue to a thialysine (Scheme 5.1) has been partially successful and that the enzyme is able to function catalytically. The low level of restored activity could be caused by one of two reasons:

1) The modification of mutant residue K247C is not complete due to the relatively small amounts of 2-BEAB added to the protein solution in order to prevent further non-specific alkylation of catalytically vital cysteines at the active site. The location of the C247 residue in a hydrophobic pocket in the active site would also hinder modification.

2) The difference of side chain lengths between lysine and thialysine (7.61 Å and 8.11 Å respectively) will cause a change in position of the amino group of residue 247 in the active site. A small change in position could have a dramatic effect on the ability of the enzyme to react with the substrate ALA molecule.

5.2.4 Restoration of K195C mutant

After reaction with 2-BEAB, the K195C mutant recovered 10% of wild-type enzyme activity, demonstrating the importance of an amino group in close proximity to Lys 247. The increased restoration of activity in comparison to the of the K247-EA enzyme may be due to the fact that C195 is nearer to the surface of the active site and may be more accessible to the reagent than C247. The exact length of the thialysine side chain may also be less vital than when considering the modification of K247C since the role of K195 is to suppress the pka of K247 rather than form a Schiff base with the substrate ALA.

Chemical rescue of the activity of mutant enzymes by exogenous organic compounds that mimic the side chain of an active site residue was demonstrated originally for the inactive K258A mutant of *E. coli* aspartate aminotransferase by use of primary amines by Toney and Korch, 1989 (68). Other reports of rescue experiments include the restoration of activity of the D237A mutant in the lactose permease of *E. coli* with an alkylsulfonate (69) and the chemical rescue of the R325A mutant of Asparagine synthase from *E. coli* with guanidinium hydrochloride (70).

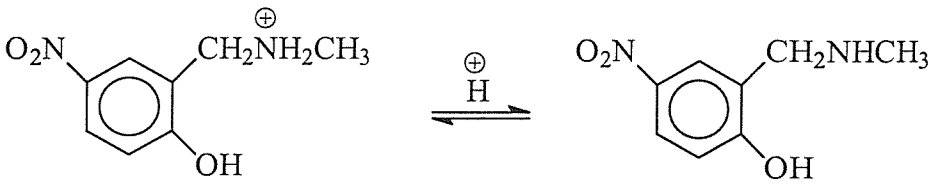
5.3 Use of 5-nitrosalicylaldehyde as an intramolecular reporter by covalent modification of *E. coli* ALAD and X-ray crystallography.

For lysine 247 to form a Schiff base with the C-4 carbon atom of the ALA substrate molecule, the ε-amino nitrogen must be in its unprotonated $-NH_2$ state. However, the pK_a of lysine in free solution at 25°C is 10.5. This would leave lysine 247 in its protonated, $-NH_3+$, state at the pH optimum of 8 for the *E. coli* ALAD. It has been proposed (26) that the adjacent lysine 195, which is in a more polar environment than lysine 247, is protonated and depresses the pKa of lysine 247 thus maintaining it in an unprotonated state and allowing it to act as a nucleophile.

In order to measure the pKa of the amino group of lysine 247, a reporter molecule, 5-nitrosalicylaldehyde (5-NSA), was employed. The use of 5-NSA to investigate the pKa of a Schiff base forming lysine in acetoacetate decarboxylase (AAD) was previously employed by Westheimer *et. al.*, (71) where it was demonstrated that the enzyme forms a Schiff base with active site lysine 116. This Schiff base was reduced with sodium borohydride, covalently binding the 5-NSA molecule to the active site lysine. Once the Schiff base had been reduced, it was possible to measure spectrophotometrically the pK_a of the resulting secondary amine group formed between Lys116 and the reporter molecule at 410nm (scheme 5.2). It was demonstrated that the pK_a of the secondary amine group had been decreased by

4.5 units (to 6.0) in comparison to the standard molecule N-methyl-2-hydroxy-5-nitrobenzylamine that models the AAD-5-NSA adduct (scheme 5.2). This behaviour was the basis of the suggestion that perturbation in the pK_a of Lys 115 was caused by a local positive charge in the active site. Kokesh and Westiemer (72) suggested that the sequence proximal lysine in the active site peptide was providing the positive charge responsible for dereasing the pk_a of its sequence neighbour allowing it to form a Shciff base at physiological pH. The measurment of the decreased pk_a of Lys 115 was in agreement with a previous estimation of the pK_a of Lys 115 *via* a titration experiment with an inhibitor molecule (73).

Scheme 5.2. Protonation of the N-methyl-2-hydroxy-5-nitrobenzylamine model molecule. The model mimics the secondary amine formed between Lys 247 and 5-NSA on reduction with $NaBH_4$. The pK_a of this group in free solution is ~10.5.



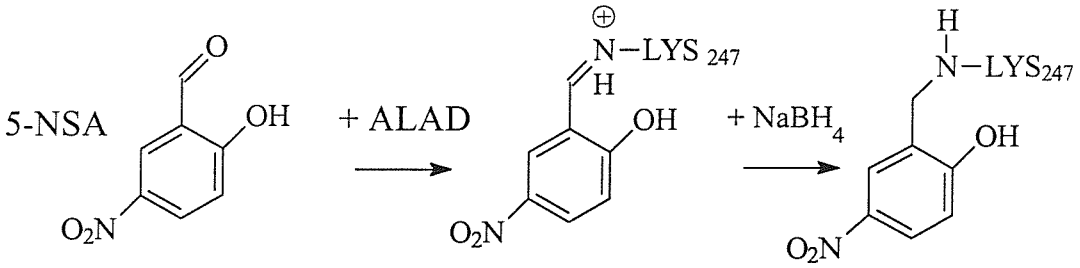
As with *E. coli* ALAD, it is proposed that AAD has two adjacent lysine residues at its active site (Lys 115 and 116) and it had been suggested (71) that it is the proximity of the amino group of lysine 115 that causes the pK_a depression of lysine 116.

5.3.1 Formation of *E. coli* ALAD-5-NSA complex

E. coli ALAD was cooled to 4°C and concentrated to 20 mg/ml in 100 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol. An aliquot of protein solution (1ml) was mixed with 0.1 ml of precooled 100 mM 5-NSA and left to react for 5 minutes. The solution was then treated with 0.01 ml of 0.1 M sodium borohydride solution. After 10 minutes, the sample was passed through a 10 x 1 cm Sephadex G25 column that had been previously equilibrated with 50 mM potassium phosphate, pH 8.0, buffer containing 50 μM $ZnCl_2$ and 10mM β-mercaptoethanol. The protein was eluted

from the column with the same buffer. As the labelled protein is yellow, its progress down the column could be followed visually.

Scheme 5.3. Formation of the ALAD-5-NSA adduct. The 5-NSA molecule binds to ALAD *via* a Schiff base formed with Lys 247. This is then reduced by $NaBH_4$ forming a secondary amine, the pK_a of which can be measured spectrophotometrically.

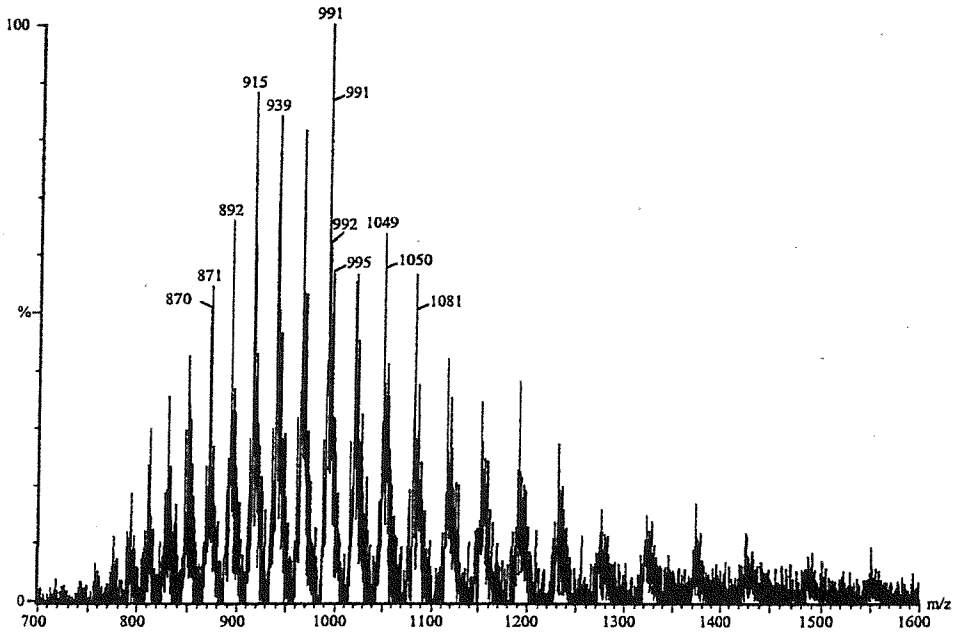


5.3.2 Mass spectrometric analysis of *E. coli* ALAD-5-NSA complex

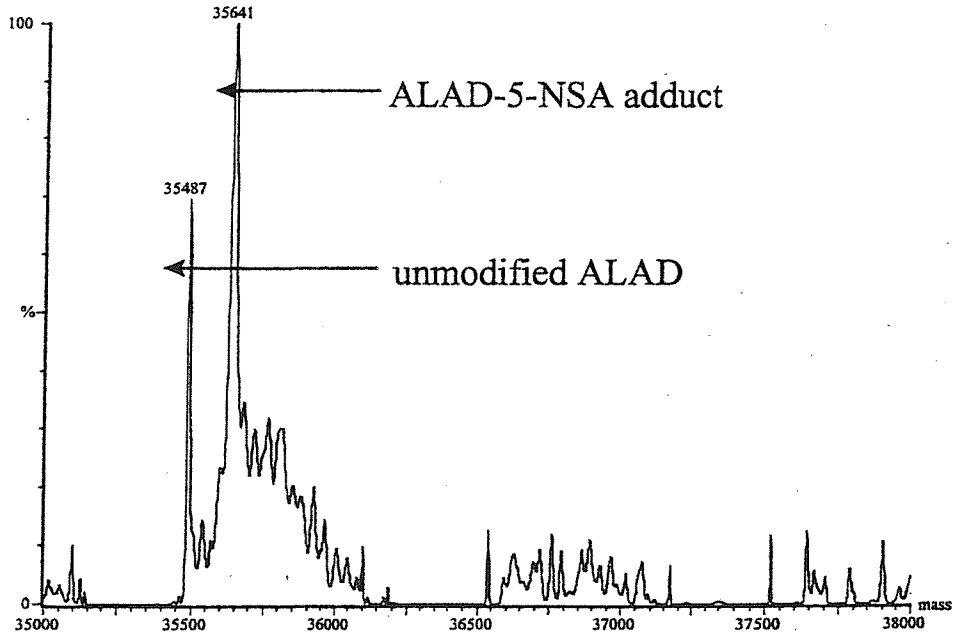
The reduced ALAD-5-NSA complex was buffer exchanged into 10 mM Tris/HCl, pH 8.0, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol and the concentration of the enzyme was adjusted to 1 mg/ml. An equal volume of 50% acetonitrile, 48% water and 2% formic acid was added to the protein solution. A volume of 20 μl was injected into a triple quadrupole electrospray mass spectrometer (micromass quatro II) in positive ion mode. Data were acquired at a cone voltage of 40 V, capillary voltage of 3.5 kV and HV lens voltage of 0.30 kV between masses of 500 m/z and 2000 m/z. The raw data obtained from the above values were then subjected to maximum entropy analysis according to the Micromass schedule. The results of the mass spectometric analysis showed modification of ALAD, presumably at a single site (fig 5.7). In order to ascertain that the reporter molecule was binding to the correct residue, Lys 247, crystallisation of the complex was attempted.

Fig 5.7. Mass spectrometer data from the *E. coli* ALAD-5-NSA adduct. **A** shows initial spectrum, where two distinct species are clearly present. **B** shows spectrum after maximum entropy analysis, showing mass of unmodified ALAD and mass of ALAD-5-NSA adduct (+154 Da).

A



B



5.3.3 Crystallisation of reduced *E. coli* ALAD-5-NSA complex

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described (chapter 2). The hanging drops contained 5 µl of reduced ALAD-5-NSA adduct solution (7mg/ml) and 5 µl of well solution consisting of 200mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200 µM $ZnSO_4$ and 6 mM β-mercaptoethanol. The crystals grew to 300 microns in diameter and were yellow in colour, indicating the incorporation of the 5-NSA reporter molecule (fig. 5.8). Although mass spectrometric analysis revealed modification of ALAD a significant portion of the sample remained unmodified (fig. 5.7)

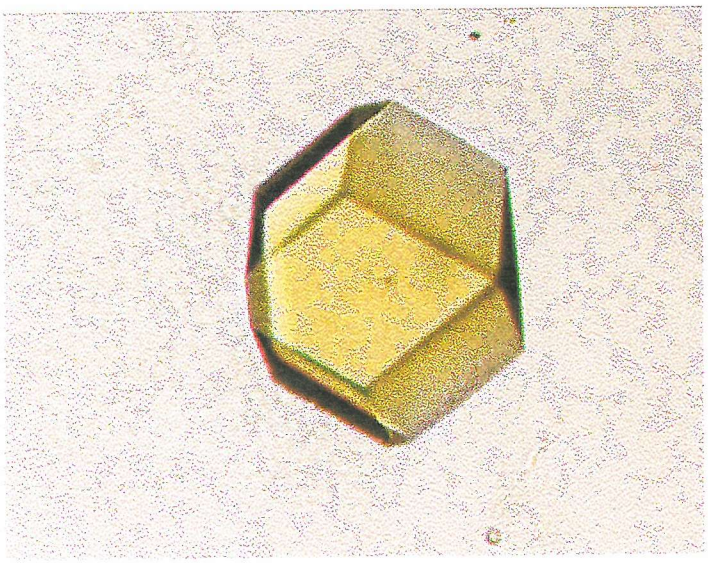
5.3.4 Data collection and processing of the reduced *E. coli* ALAD-5NSA complex

The reduced ALAD-5-NSA complex crystallised in space group I422 (table 5.3). Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the reduced ALAD-5-NSA complex were collected from a cryocooled crystal at Southampton University, on an 18 cm MAR image plate and were complete to 3.3 Å. Data were processed with the program MOSFLM (55) and the CCP4 suite (1994). For processing statistics see table 5.3. The 5-NSA molecule was located by calculation of difference Fourier map phased with the native *E. coli* ALAD structure (26). The structure was refined using X-PLOR (60) and were rebuilt using Quanta (62) running on Silicon-Graphics computers.

5.3.5 Quality of structure analysis of the reduced *E. coli* ALAD-5NSA complex

The structure of the *E. coli* reduced ALAD-5-NSA complex was refined using data between 15 Å and 3.3 Å and had an R-factor of 25.7 % and a free R-factor of 31.6% (see table 5.3). The structure evaluation program PROCHECK was used to assess the quality of the final model. A Ramachandran plot produced by the

Fig. 5.8 *E. coli* ALAD-5-NSA crystals. The yellow colour of the crystals indicates the incorporation of the 5-NSA reporter molecule.



300 microns

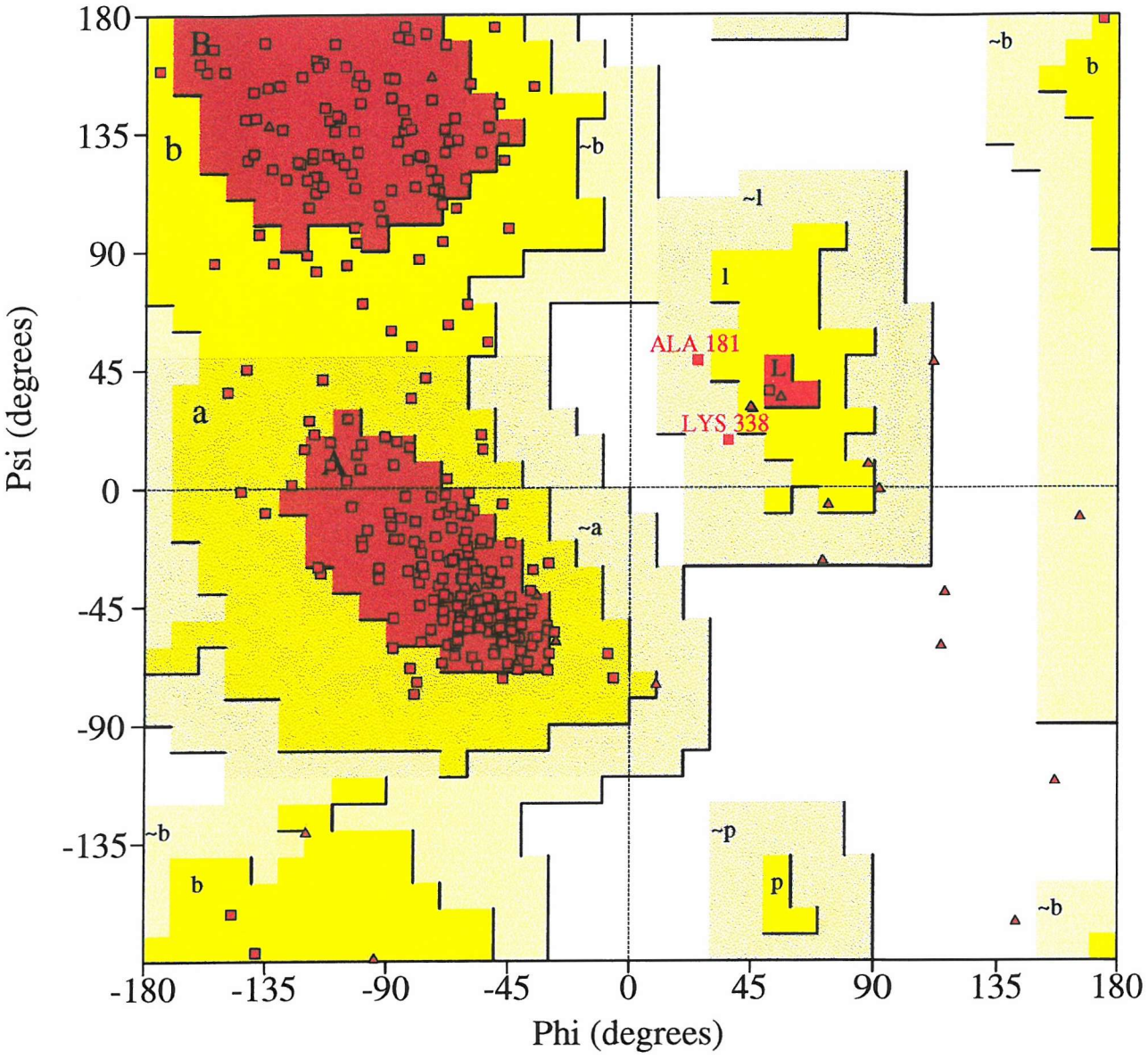
Table 5.3. Processing statistics for *E. coli* ALAD in complex with 5-NSA.

Space group	I422
a b Å	129.76
c Å	142.90
Entire data set	
Resolution range Å	15 - 3.3
R-merge (%)	7.6
Completeness (%)	97.0
Overall (I/Iσ)	6.7
Fraction I> 3σ (I)	83.4
Multiplicity	5.3
Outer Shell	
Resolution range	3.42 - 3.3
R merge (%)	10.7
I / σ I	6.1
Fraction I> 3σ (I)	65.1
Multiplicity	5.3
Refinement	
R factor (%)	25.7
Free R factor (%)	29.41
Resolution range	15-3.3
Cut off	2σ
Total number of reflections	9126
Number of atoms	3035

PROCHECK

Fig 5.9

Ramachandran Plot
ALAD-5NSA



Plot statistics

Residues in most favoured regions [A,B,L]	230	79.9%
Residues in additional allowed regions [a,b,l,p]	56	19.4%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	288	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	323	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

LAD-5NSA_01.ps

program showed 80 % of residues within the allowed regions and 99 % of residues within the additional allowed boundary (fig 5.9).

5.3.6 Analysis of *E. coli* ALAD-5-NSA complex data

After processing of ALAD-5-NSA complex data and production of electron density maps, it was possible to look at the active site and observe electron density differences between that of the native model and the complex. Inspection of this map at the active site revealed a large feature of difference density adjacent to Lys 247. A 5-NSA molecule was then built into this density and the molecule underwent further refinement. (5.10, 5.11)

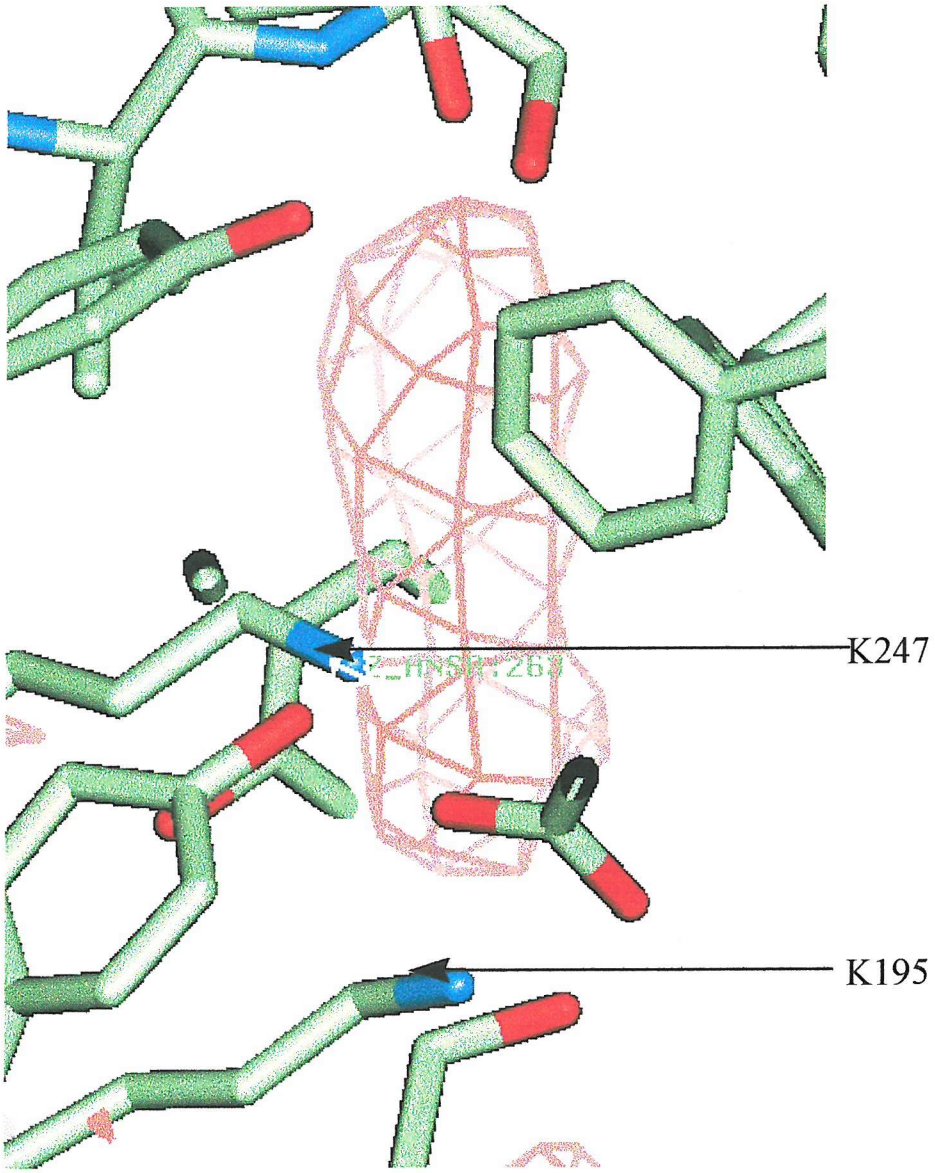
5.3.7 Position and binding interactions of 5-NSA at the P-site of *E. coli* ALAD.

The 5-NSA residue is covalently bound to Lys 247 *via* the reduction of the initially formed Schiff base by $NaBH_4$ and is positioned between the hydrophobic side chains of Phe 204 and Val 272. The nitro group of the 5-NSA molecule is in a similar position to the carboxyl group of the ALA substrate molecule (chapter 3). The nitro group forms hydrogen bonds with the side chain of Tyr 312 and the hydroxyl group of Ser 273, the same residues that form hydrogen bonds with the carboxyl group of either ALA or LA molecules at the P-site.

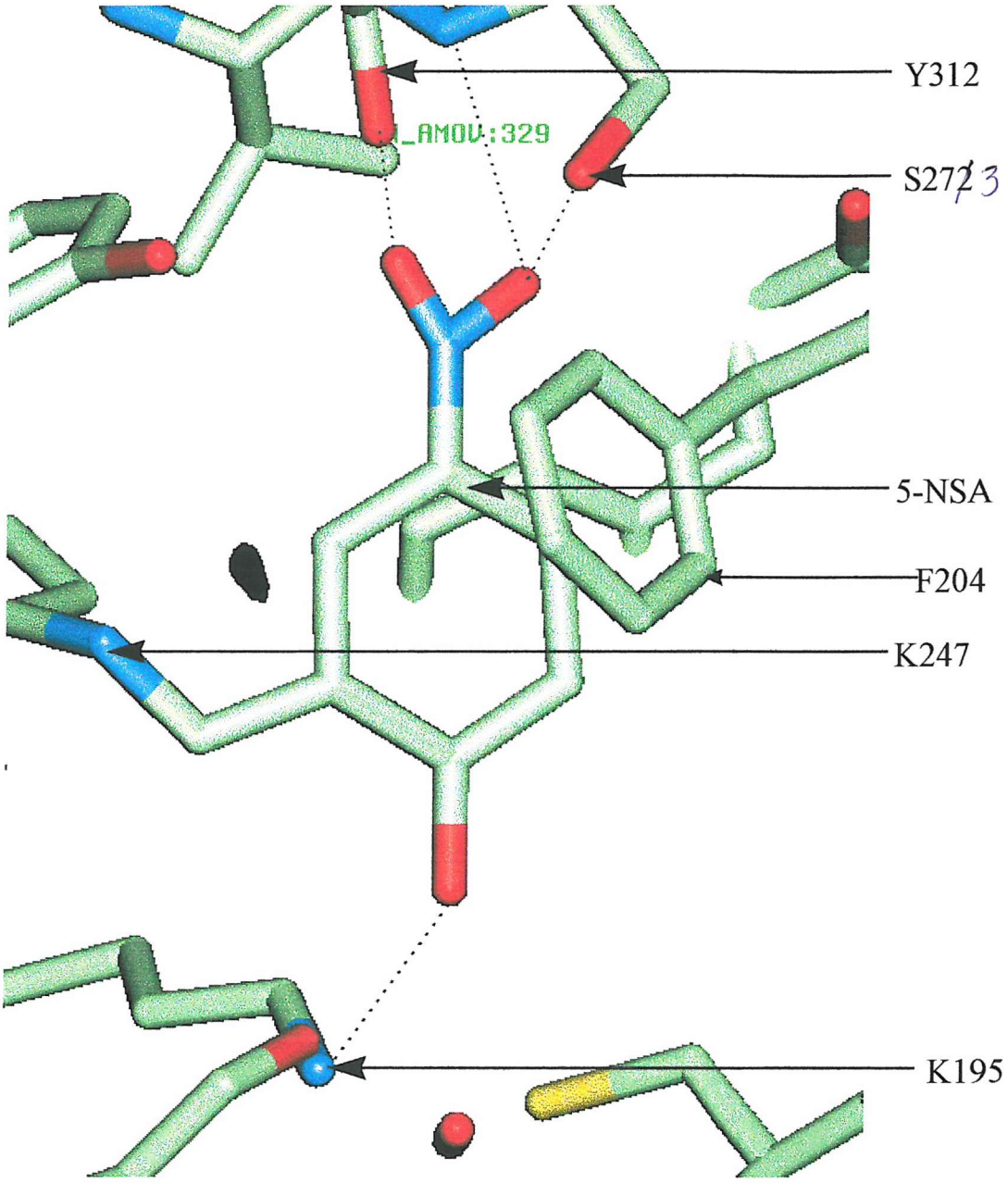
The phenol OH group of the reporter molecule forms hydrogen bonds with the amino group of Lys 195. As with previous *E. coli* ALAD-substrate and ALAD-inhibitor complex structures, invariant residue Phe 204 has packed against the inhibitor molecule forming hydrophobic interactions. Residue Phe 204 is part of the flexible active site loop which becomes rigid when Phe 204 stacks against the 5-NSA molecule at the P-site of the active site.

The solution of the ALAD-5-NSA structure coupled with mass spectrometric data showing one 5-NSA molecule bound per monomer confirms that the reporter molecule is bound to Lys 247. This ensures that pK_a values obtained can be attributed to the environment of this specific residue.

Fig. 5.10 Fo-Fc map of active site of *E. coli* ALAD-5-NSA adduct contoured at 2 sigma. The portion of difference density adjacent to Lys 247 clearly shows the position of the 5-NSA reporter molecule.



5.11. 5-NSA reporter molecule bound *via* a reduced Schiff base at the active site of *E. coli* ALAD. The interactions formed between the nitro group of 5-NSA and Ser 272 and Y312 are similar to those formed with the carboxyl group of ALA and LA



5.3.8 Determination of the pK_a of the secondary amine group of the *E. coli* ALAD-5-NSA adduct.

Spectroscopic measurements were made with a Hitachi 2000 spectrophotometer. To obtain a spectrum, 0.8 ml of buffer was added to each of two 1ml quartz cells. One of the cells was then used to provide a baseline blank for the wavelength scan. The cell was removed and 10 µl of 13 mg/ml ALAD-5-NSA in 50 mM potassium phosphate buffer, pH8.0, containing 50 µM $ZnCl_2$ and 10 mM β-mercaptoethanol was added. A 10 µl sample of buffer was added to the second cell as a control. The contents of each cell were mixed and the absorption spectrum between 300 nm and 600 nm was determined for each (fig 5.13). Spectra were determined for pH values in the range 6.0-10.5 (table 5.4) using potassium phosphate buffers at pH intervals of 0.5 and the pK_a value for protonation of the secondary amine group of the ALAD-5-NSA adduct determined (scheme 5.2). The experiment was repeated with unmodified *E. coli* ALAD as a control.

Fig. 5.12 Absorption spectra for ALAD-5-NSA complex at pH 10.0(A) and pH 5.9(B)

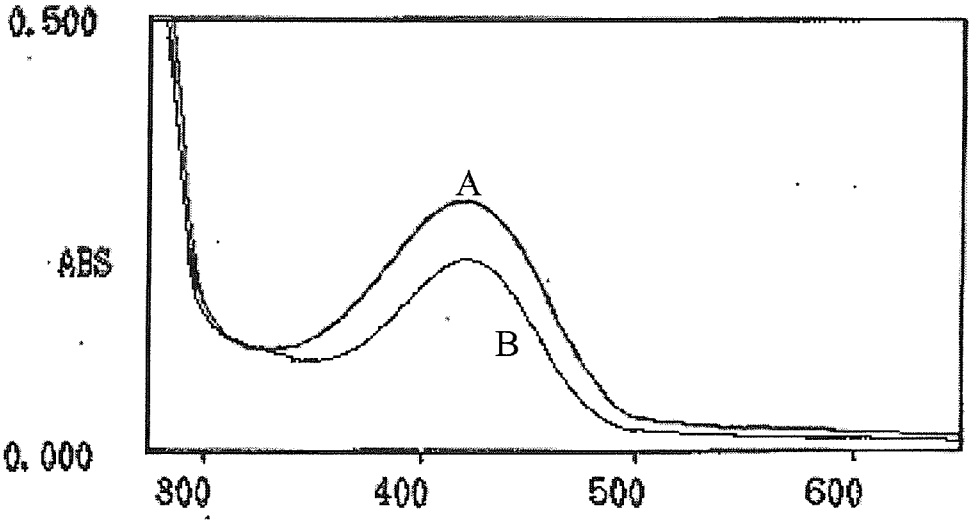
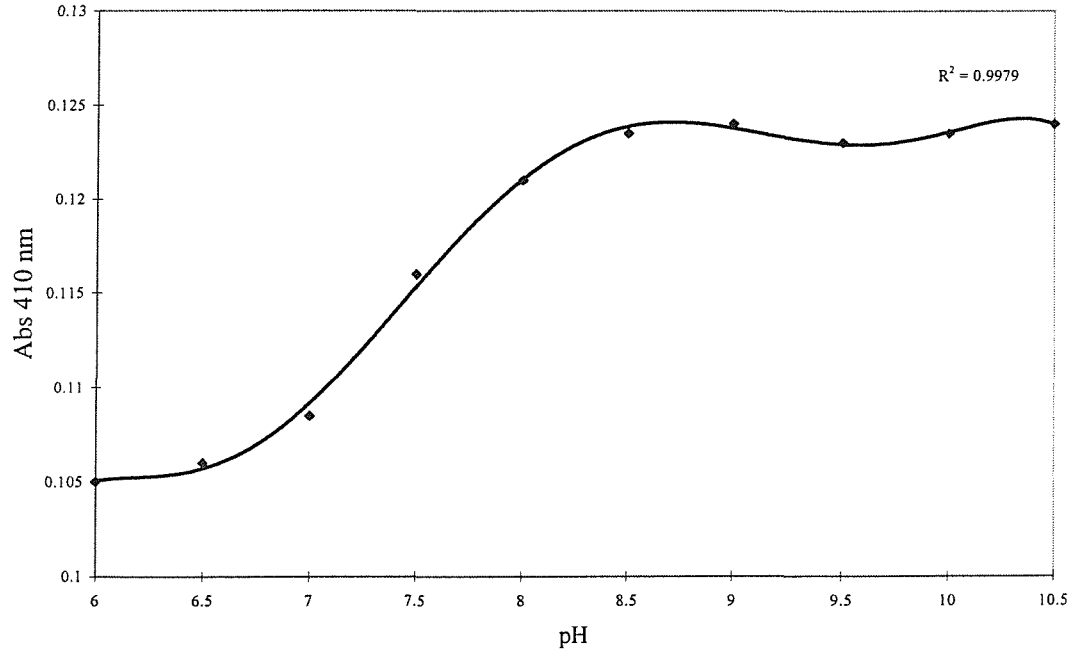


Table 5.4 Absorption at 410 nm of ALAD-5-NSA adduct over range of pH values. Table shows average of three assays.

pH	Abs. 410nm
6	0.105
6.5	0.106
7	0.1085
7.5	0.116
8	0.121
8.5	0.1235
9	0.124
9.5	0.123
10	0.1235
10.5	0.124

Scheme 5.4 Graph showing titration of ALAD-5-NSA adduct at 410nm. The pK_a of the secondary amino group in the 5-NSA modified enzyme is estimated to be 7.5



5.3.9 Results of E. coli ALAD-5-NSA adduct analysis

The data from mass spectrometric and crystallographic analysis of the reduced ALAD-5-NSA adduct show a single addition of 5-NSA *via* a reduced Schiff base to invariant residue lysine 247.

The ionisation constant (pK_a) of the secondary amine group formed between K247 and the reporter 5-NSA molecule followed at 410 nm was determined to be 7.5, which is 3.0 units lower than that of the non enzymic model.

The change in pK_a can be ascribed to the electrostatic effect of the nearby positive charge of K195 which, as previously mentioned, is in close proximity to K247. These findings are in close parallel with AAD which also has a lysine residue that forms a Schiff base with the substrate molecule, and is also thought to have an adjacent lysine residue.

The sharing of a single charge between Lys 247 and Lys 195 can be compared with the active site of aspartic proteinases, which have two aspartic side chains in close proximity at the catalytic centre. Aspartic proteases are a group of homologous enzymes which include pepsin, chymosin, cathepsin D and related microbial enzymes such as endothiapepsin. All aspartic proteases have acidic pH optima and have two aspartates involved in the catalytic mechanism. The elucidation of the X-ray structure of the active site of endothiapepsin at pH 4.5 (Pearl *et. al.* 1984) revealed a symmetrical hydrogen bond network involving the side chains of the two catalytic (74) aspartate residues (Asp 32 and Asp 215). The two carboxyl groups of the aspartates are 2.9 Å apart. As with the key lysines in ALAD it is thought that the two proximal aspartate residues share different pK_a values with Asp 32 having an abnormally high value (~5) and the pKa of Asp 215 being abnormally low (~2), causing Asp 215 to exist in its deprotonated state in the native enzyme. The negative charge carried by the deprotonated Asp 215 causes polarisation of a solvent molecule tightly bound to both aspartate carboxyls allowing it to attack the scissile bond of the substrate. (75)

The lowered pK_a of *E. coli* ALAD K247 is mechanistically important. If a Schiff base is to be formed between K247 and the carbonyl group of ALA, the amino group of the lysine must be in its neutral, unprotonated form so it can act as a nucleophile. If K247 had a similar pK_a to that of lysine in free solution (10.5) then only one residue in 1000 would be in the unprotonated form at the ALAD pH optimum of 8.0 which would have effects on enzyme activity. Interestingly, when the lysine adjacent to K247, K195, is mutated to an alanine residue the V_{max} and the K_{cat} of the enzyme drops to 0.1% of that measured for wild type ALAD. This adds further evidence that it is the positive charge on K195 that is depressing the pK_a of K247.

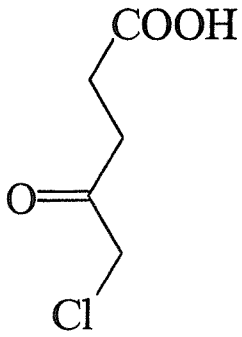
5.4 Inhibition of *E. coli* ALAD by 5-chlorolaevulinic acid.

5.4.1 Studies of chlorolaevulinic acid inhibition of ALAD

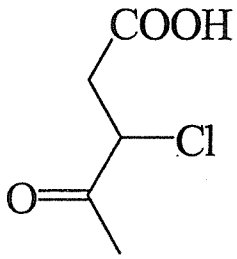
The use of α-haloketones to alkylate active site nucleophiles has been well described. Studies on the inactivation of bovine ALAD by 5- and 3-chlorolaevulinic (Scheme 5.5) acid were initially carried out by Jordan and Seehra (76). 3-Chlorolaevuinic acid (3-CLA) was shown to be an inferior inactivating agent to 5-chlorolaevuinic acid 5-CLA which was a highly effective competitive inhibitor with a K_i = 11μM. When 5-CLA levels were raised to 5mM the enzyme was irreversibly inhibited. Inhibition by 3-CLA can be prevented by the addition of LA, but only ALA is able to protect the enzyme against inactivation by 5-CLA.

Scheme 5.5 showing the chemical structures of 5-Chlorolaevulinic acid and 3-chlorolaevulinic acid. All 5-chlorolaevulinic acid used in this study was synthesised by Jack Cheung.

5-Chlorolaevulinic acid



3-Chlorolaevulinic acid



The dual inhibition characteristics of 5-CLA, competitive at low concentrations and irreversible inactivator at high, is of great interest. Work carried by Jaffe (77) further characterised the nature of the interaction between 5-CLA and ALAD. Jaffe showed that after inactivation by high concentrations of 5-CLA, ALAD was still able to bind four molecules of ALA per octamer, with a K_d of 60 μM. This K_d value was between those found for the A (242 μM) and P (3.8 μM) sites of native enzyme, and was thought to be indicative that 5-CLA binds irreversibly to the A site leading to a reduced level of binding at the P-site. Treatment of the 5-CLA modified enzyme with ALA and $NaBH_4$ showed the covalent linkage of 4 substrate molecules per octamer, again hinting at the A site binding of 5-CLA.

The Zn^{2+} binding properties of the 5-CLA-modified enzyme were also investigated. It was discovered that 5-CLA blocked the complexion of the octameric protein to the catalytic Zn^{2+} binding site. It was proposed that the 5-CLA binds to a cysteine in the active site which is involved in Zn^{2+} binding. The conclusions from this research suggest that at low concentrations, 5-CLA acts as a competitive inhibitor at the P-site, whilst at higher concentrations the A site is permanently modified *via* covalent interactions with a cysteine residue.

5-CLA inhibition of ALAD from *Bacillus subtilis* was investigated by Appleton *et. al.,* (19) by kinetic and mass spectrometer analysis. In this study ALAD was incubated with increasing concentrations of 5-CLA (2-100 mM) and the resulting masses of the modified enzymes were obtained. At low concentrations of 5-CLA the enzyme was mono-alkylated whereas at the highest concentrations of 5-CLA, tri- and tetra-alkylations could be seen. The kinetic analysis of ALAD inhibition with 5-CLA revealed no saturation kinetics indicating that the alkylation that caused the inactivation did not occur at the active site. The conclusions drawn from this study were that 5-CLA is not an active-site specific inhibitor but is a non specific alkylating agent acting generally on cysteine residues in the enzyme.

In light of the recent solution of the X-ray structures of ALAD from different species (25,26,27), the opportunity to characterise the interaction between ALAD and 5-CLA using crystallographic techniques arose.

5.4.2 Crystallisation of ALAD-5-CLA complex

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described in chapter 2. The hanging drops contained 5µl of ALAD solution (7mg/ml) and 5µl of well solution consisting of 200 mM Tris/HCl buffer, pH 8.0-8.4, containing 2% ammonium sulphate, 200 µM $ZnSO_4$, 6mM β-mercaptoethanol and 5 mM 5-CLA (synthesised by Dr. Kwai-Ming Cheung).

5.4.3 Data collection and processing

The ALAD-5-CLA complex crystallised in space group I422 (table 5.5). Cryoprotection was employed for data collection at 100 K by addition of crystals to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystals were then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the ALAD-5-CLA complex were collected from cryocooled crystals on beamline 14 at the ESRF (Grenoble, France) using an 18 cm MAR image plate and were complete to 3.6 Å. Data were processed with the programme MOSFLM (55) and the CCP4 suite (1994), for processing statistics see table 5.5. The CLA molecule was located by

calculation of difference Fourier map phased with the native *E. coli* ALAD structure (26). The structure was refined using X-PLOR (60) and rebuilt using Quanta (62) running on Silicon-Graphics computers.

5.4.4 Quality of structure analysis

The structure of *E. coli* ALAD-5-CLA complex was refined using data between 15 Å and 3.6 Å and had an R-factor of 24 % and a free R-factor of 30.1% (see table 5.5). The structure evaluation program PROCHECK was used to asses the quality of the final model. A Ramachandran plot produced by the program showed 75 % of residues within the allowed regions and 99 % of residues within the additional allowed boundary (fig 5.13)

5.4.5 Analysis of ALAD-5-CLA complex data

After processing of ALAD-5-CLA complex data and production of electron density maps it was possible to look at the active site and observe electron density differences between that of the native model and the complex. Inspection of this map at the active site revealed a large feature of difference density adjacent to Lys 247. A 5-CLA molecule was then built into this density as a Schiff base between Lys 247 and the carbon atom at position 4 of the 5-CLA molecule and the map underwent further refinement. (fig. 5.14)

Initial positioning of 5-CLA at the active site was unsuccessful. The carboxyl "head" and 5-carbon "body" of the molecule were well defined. However the inferred position of the Cl atom at the base of the molecule contained little difference density and was already occupied by the amino group of K195.

On further examination of the Fo-Fc map, it became apparent that the Lys 195 molecule had undergone alkylation, and that the ε-NH_2 group now covalently bound to the C-5 carbon of the inhibitor molecule. The new position of the inhibitor molecule was modelled in and refined further, confirming that the Cl atom of 5-CLA had been eliminated during reaction with the ALAD. (fig. 5.15)

Table5.5. Processing statistics for *E. coli* ALAD in complex with 5-CLA.

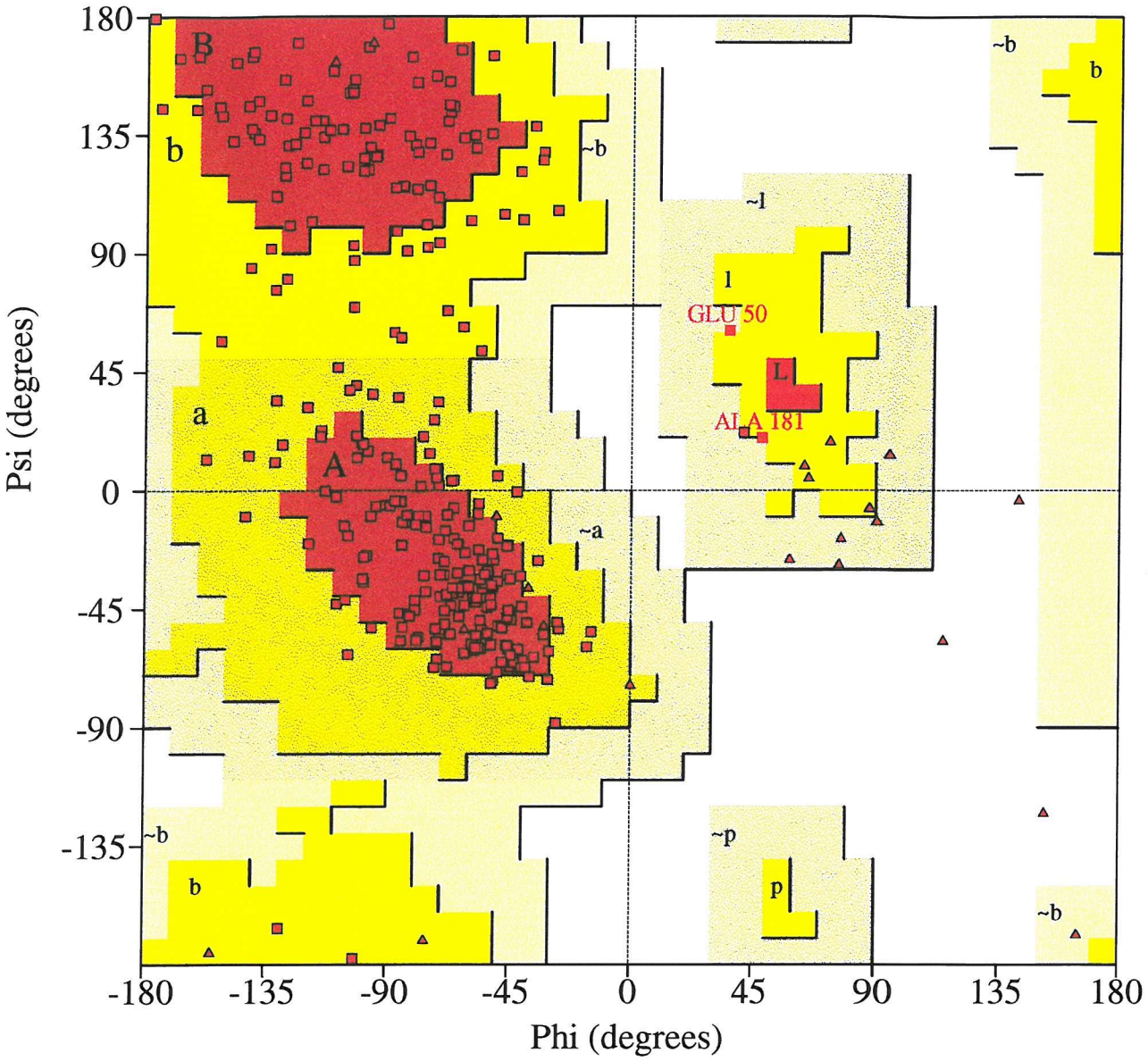
Space group	I422
a b (Å)	132.13
c (Å)	146.6
Entire data set	
Resolution range Å	48.24 - 3.6
R-merge (%)	6.2
Completeness (%)	97.2
Overall (I/Iσ)	6.4
Fraction I> 3σ (I)	79.1
Multiplicity	4.9
Outer Shell	
Resolution range	3.67 - 3.6
R merge (%)	9.4
I / σ I	2.4
Fraction I> 3σ (I)	67.9
Multiplicity	6.8
Refinement	
R factor (%)	23.6
Free R factor (%)	30.1
Resolution range	15-3.6
Cut off	2σ
Total number of reflections	12098
Number of atoms	3035

PROCHECK

Fig 5.13

Ramachandran Plot

ALAD-CLA



Plot statistics

Residues in most favoured regions [A,B,L]	214	74.8%
Residues in additional allowed regions [a,b,l,p]	70	24.5%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	286	100.0%
Number of end-residues (excl. Gly and Pro)	3	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	322	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

LAD-CLA_01.ps

5.4.6 Mass spectrometric analysis of *E. coli* ALAD-5-CLA complex

In order to verify the proposed mode of action of 5-CLA, the ALAD-5-CLA complex was subjected to mass spectrometer analysis. Native enzyme (5mg/ml) was incubated with 10 mM 5-CLA in 50 mM potassium phosphate buffer, pH 8.0, containing 50 μM $ZnCl_2$ and 1 mM β-mercaptoethanol, for 3 hours at 4°C. The resulting ALAD-5-CLA complex was used for analysis. The results of this analysis showed multiple alkylation of the enzyme at up to four separate sites (Fig 5.16). If alkylation of the enzyme by 5-CLA was occurring at cysteine residues, a Mr. increase of 116. can be expected. However, the mass of the mono-alkyated complex was just 103 Da. greater than that of the wild type. This smaller mass is consistent with the forming of a Schiff base with the carbonyl carbon of 5-CLA and Lys 247 which would eliminated water and account for the lower Mr. of the complex by 18. After this Schiff base formation alkylation of Lys 195 can occur eliminating the Cl atom and covalently linking the inhibitor molecule to the enzyme, irreversibly. It is important to note that mass spectrometric analysis of ALAD + ALA or ALAD + LA without reduction does not reveal the ES or EI complex.

The further mass peaks revealed by the analysis correspond to alkylation of cysteine residues by 5-CLA. No evidence alkylation of was found at cysteine residues in the active site in the X-ray structure, however, this may have been due to the lower concentration of 5-CLA (5 mM)used in these studies compared tot he higher concentration (10 mM) used for the mass spectrometric analysis.

5.4.7 5-CLA carboxyl group binding at P-site

The carboxyl group binding at the P-site of the enzyme is well defined by the structure. One carboxyl oxygen H-bonds to the hydroxyl of Ser 272 with the remaining carboxyl oxygen hydrogen bonding with the nitrogen on the main chain of the same residue. This second oxygen also forms a hydrogen bond with the phenolic oxygen of Tyr 312. This binding is identical to that of the carboxyl group of LA and ALA.

Fig 5.16 Mass spectroscopic analysis of *E. coli* ALAD mutants complex

Mass spectrometer analysis of *E. coli* ALAD modified with 5-CLA can be seen in Fig. 6.14. The spectra shows mono- (M r= 35620 ± 0.96.), bi- (Mr = 35735 ± 1.03.), tri- (Mr = 35847 ± 0.84.) and tetra- (Mr = 35962 ± 1.23.) alkylation of the native enzyme (35517 Da.). Notably the mass of the mono-alklyated species is less than would be expected if normal cysteine alkylation had occurred. This corresponds to the elimination of an oxygen atom on formation of a Schiff base between the carbonyl group of 5-CLA and Lys 247 and the subsequent alkylation of Lys 195 with the loss of Cl^-. The absence of any species without the loss of H_2O indicates that the preferential reaction is the alkylation of K195 rather than of cysteines.

The pre-incubated ALAD-5-CLA sample was buffer exchanged into 10 mM Tris/HCl buffer, pH 7.5, containing 50 μM $ZnCl_2$ and 10 mM β-mercaptoethanol and the concentration of the enzyme was adjusted to 1 mg/ml. An aliquot (250 μl) of protein was added to an equal volume of 50% acetonitrile, 48% water and 2% formic acid was added to the protein solutions. A volume of 20 μl was injected into an electrospray triple quadrupole mass spectrometer in positive ion mode. Data were acquired at a cone voltage of 40 V, capillary voltage of 3.5 kV and HV lens voltage of 0.30 kV between masses of 500 m/z and 2000 m/z. The raw data obtained from the above values were then subjected to maximum entropy analysis according to the Micromass schedule.

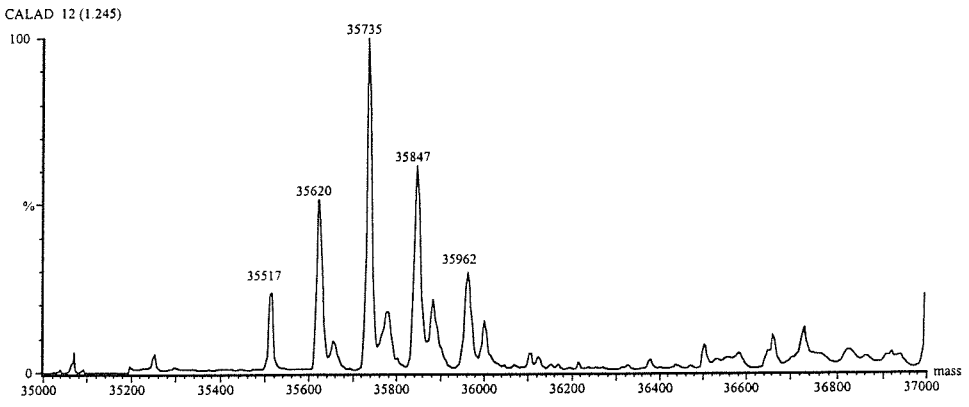


Fig 5.14: Fo-Fc and 2Fo-Fc maps, contoured at 2sigma, showing position of 5-CLA at the active site of *E. coli* ALAD. Map A shows difference density into which the inhibitor5-CLA was modelled and refined as a Schiff base with K247. Map B shows 5-CLA at active site where the chlorine atom of the inhibitor molecule and the amino group of Lys 195 are clearly too close together to be correct.

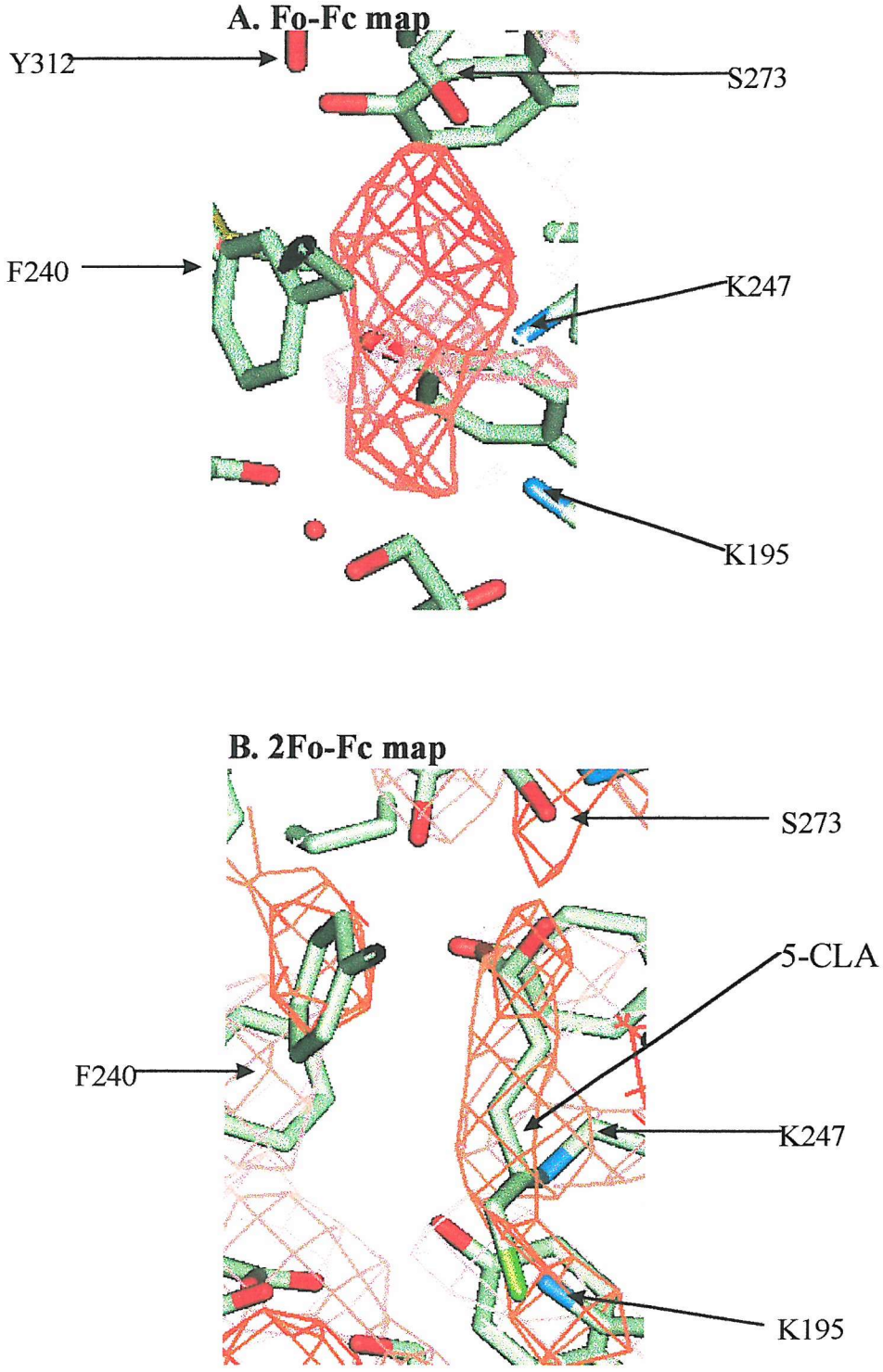
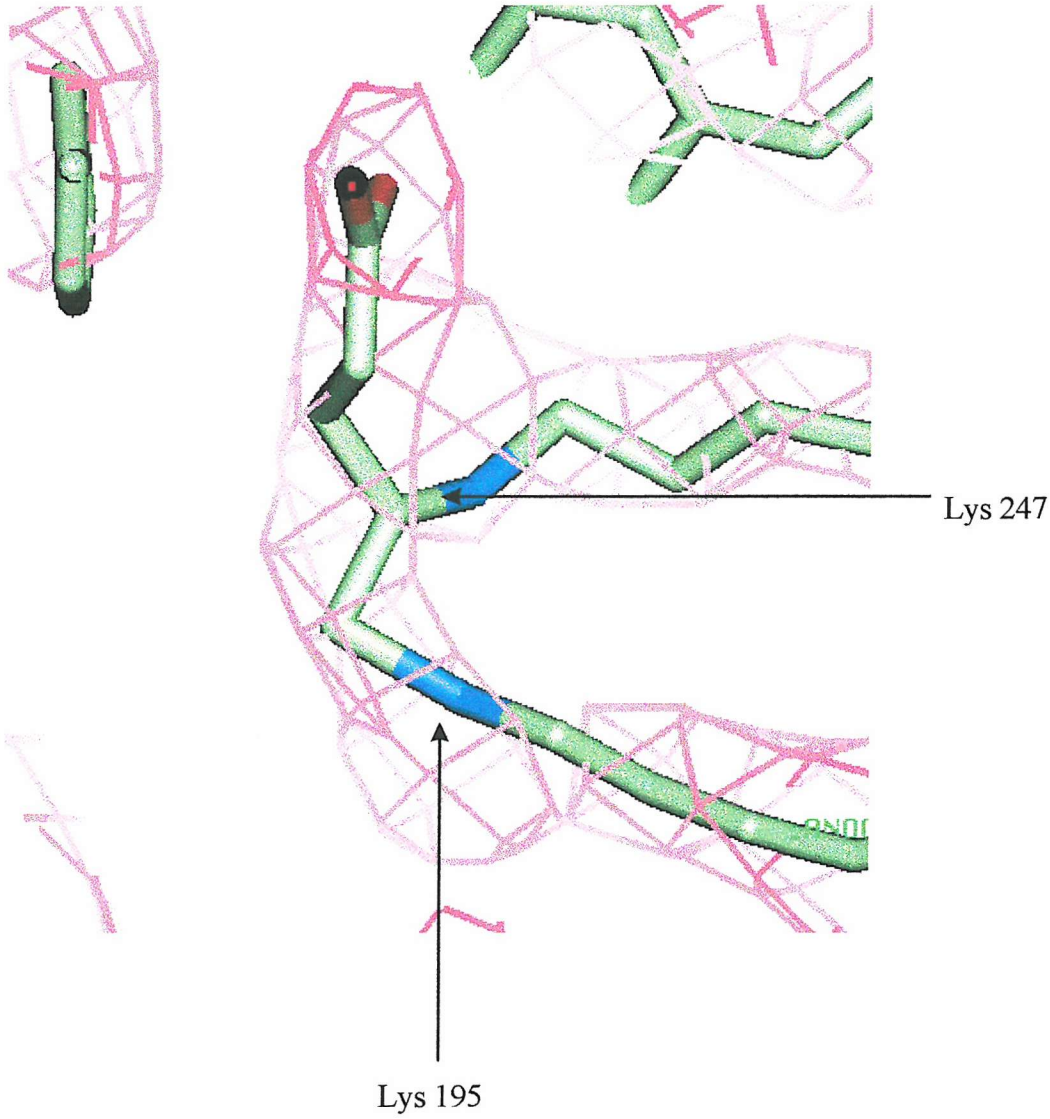


fig. 5.15 2Fo-Fc map, contoured at 2 sigma, showing that Lys 195 has been alkylated by the 5-CLA with the elimination of Cl^- . Both Lys 247 and Lys 195 are now bound covalently to the inhibitor molecule, Lys 247 as a Schiff base and Lys 195 as an alkylated species leaving the enzyme irreversibly inactivated



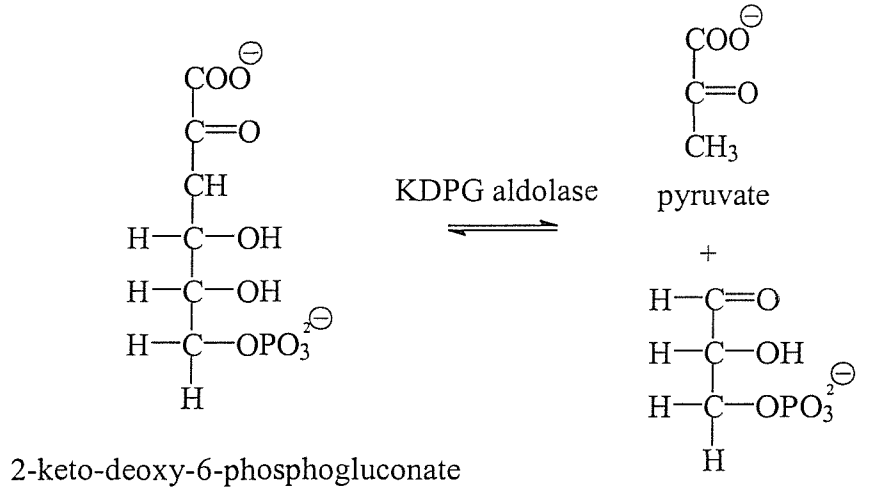
5.4.8 Irreversible inhibition of *E. coli* ALAD by 5-CLA

The evidence suggesting the covalent modification of K195 by 5-CLA is consistent with its action as an inactivating agent. Once bound *via* a Schiff base to K247, the C-5 carbon of 5-CLA is made δ^+ by the dual effects of electron withdrawal of the protonated Schiff base as well as the Cl atom and is susceptible to nucleophilic attack by the $-NH_2$ group of K195 which is in close proximity. This eliminates the Cl atom and irreversibly modifies the active site of the enzyme (scheme 5.6).

5.4.9 Inhibition of 2-keto-3-deoxy-6-phosphogluconate aldolase by bromopyruvate

2-Keto-3-deoxy-6-phosphogluconate (KDPG) aldolase catalyses an aldol cleavage of KDGP *via* an initial Schiff base formation between a substrate carbonyl an lysine residue at the active site (78). The cleavage products of this reaction are pyruvate and glyceraldehyde-3-P (Scheme 5.7).

Scheme 5.7 KDPG aldolase catalysed cleavage of 2-Keto-3-deoxy-6-phosphogluconate. This reaction produces pyruvate and glyceraldehyde-3-P as cleavage products.



Scheme 5.6 Inhibition of *E. coli* ALAD by 5-chlorolaevulinic acid. Normal Schiff base formation occurs between Lys 247 and the carbonyl carbon of the inhibitor. The positive charge of the Schiff base depresses the pK_a of Lys 195, leaving it unprotonated and able to act as a nucleophile. Once the Schiff base is formed alkylation of Lys 195 takes place (2). This leaves the inhibitor molecule irreversibly bound to the active site of the enzyme (3).

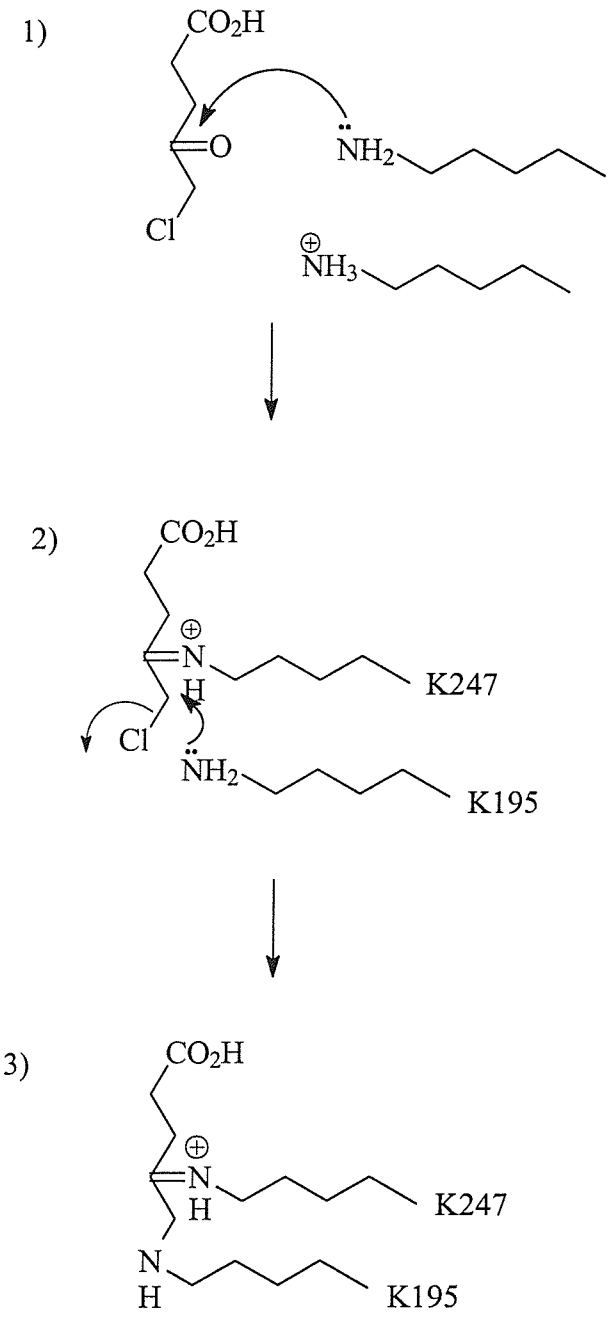
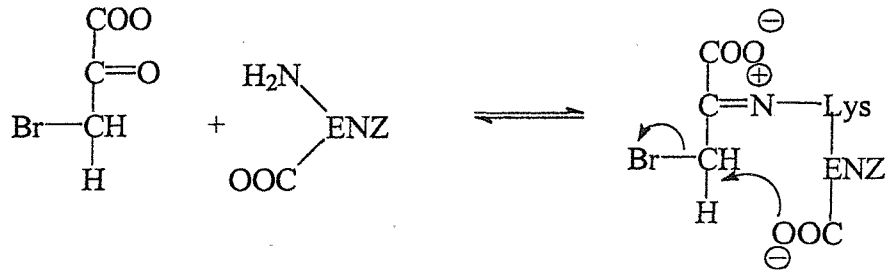
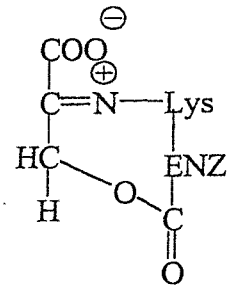


Fig 5.8 Inactivation of 2-Keto-3-deoxy-6-phosphogluconate (KDPG) aldolase by bromopyruvate: After initial Schiff base formation between a carbonyl group of the inhibitor and a lysine residue at the active site of KDPG aldolase, alkylation of an adjacent carboxyl group occurs. The action of bromopyruvate is analogous to the inactivation of *E coli* ALAD by 5-CLA.



bromopyruvate



alkylated enzyme

Meloche *et al.*, (79) carried out a number of inhibition experiments on KDPG aldolase using the pyruvate analogue β-halo-α-keto acid bromopyruvate and discovered that bromopyruvate inactivates KDPG aldolase

Meloche suggests that bromopyruvate is initially bound in imine linkage to the active site lysine residue of KDPG aldolase and once in this position undergoes attack by a carboxylate anion of the enzyme (fig 5.8), a theory substantiated by experiments with chiral 3S-[^{3}H]-bromopyruvate.

The initial formation of a Schiff base between KDGP aldolase and subsequent alkylation of an adjacent residue in the active site, resulting in enzyme incativation, is analogous to the proposed inhibition of *E. coli* ALAD by 5-CLA. Whereas in the example of KDPG aldolase the nucleophile trapped by alkylation is an enzyme carbonyl group, the group alkylated by 5-CLA inactivation of *E. coli* ALAD is invariant Lys 195.

5.5 Purification, characterisation and X-ray structure determination of *E. coli* ALAD mutant K247N

Mutation of key lysine residue 247 in *E. coli* ALAD has dramatic effects on enzyme activity with mutants K247A and K247C (section 5.1) being completely inactive. However, analysis of mutant K247N revealed a low but significant level of enzyme activity (0.01% of wild type) . It was hoped that the study of this mutant will give further insight to the contribution of the initial Schiff base formation between Lys 247 and substrate ALA to ALAD catalysed PBG formation. Details of the kinetic analysis of K247N can be seen in fig. 5.18

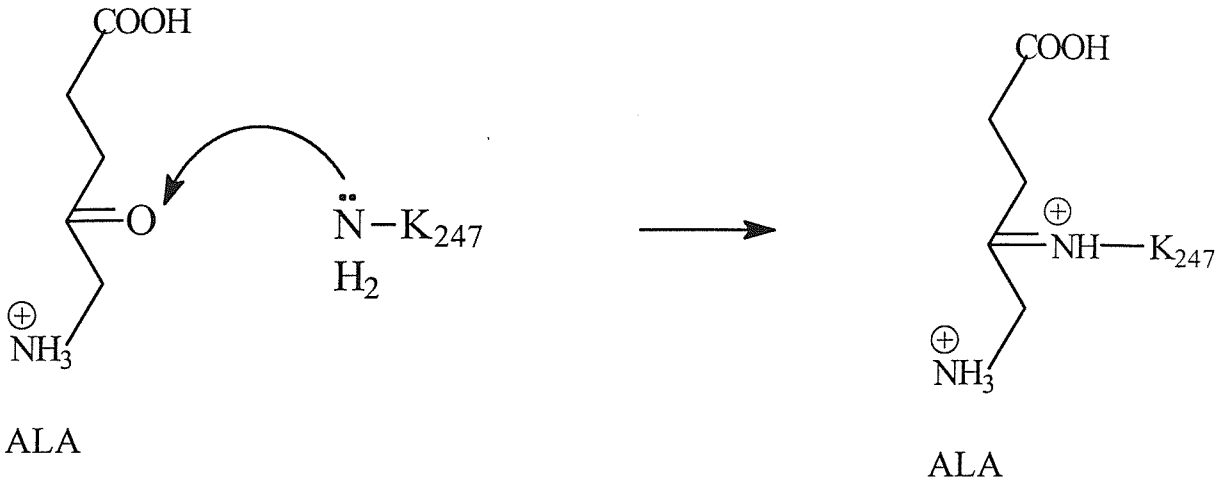
5.5.1 Schiff base formation

A Schiff base can be described as the imine arising from the condensation of an amine and a carbonyl containing compound. The initial step of the reaction is the nucleophilic attack by the amino nitrogen upon the electron-deficient carbonyl group to form an amino carbinol. This intermediate is unstable and usually dehydrates to the imine. This reaction is reversible with the imine existing in solution as an equilibrium mixture with the amino carbinol and its components. The resulting imine is highly susceptible to nucleophilic attack.

The formation of a Schiff base between enzyme and substrate is an example of covalent catalysis. This is an efficient method of catalysis because the transient chemical modification that occurs by forming a covalent bond gives an intermediate which facilitates the bond making and breaking process.

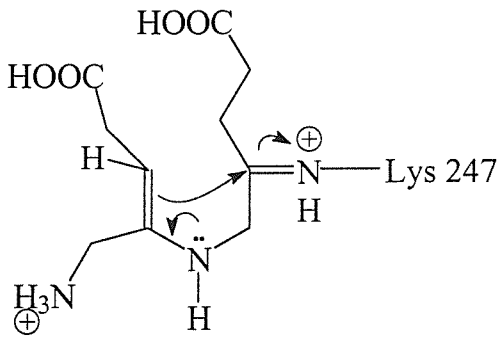
ALAD falls into the category of Schiff base forming enzymes that lack a carbonyl group but act *via* an intermediate imine formed between an amino group of the enzyme and a carbonyl group of the substrate. In the case of ALAD the initial binding of ALA to the active site occurs *via* Schiff base formation at invariant Lys 247. This involves the conversion of the carbonyl group of ALA into an imine. (scheme 5.9)

Scheme 5.9 Initial Schiff base formation between Lys 247 (K_{247}) and P-side ALA. This is the first step in ALAD catalysed PBG formation and is vital for enzyme activity.



This replacement of the carbonyl oxygen at the C-4 position of the ALA substrate molecule by the nitrogen of lysine 247 makes it much more basic. The pK_a of the imine is far higher than that of the carbonyl group, causing the nitrogen to exist in its protonated form. This protonated imine is many orders of magnitude more susceptible to nucleophilic attack than the carbonyl group (scheme 5.10). This in turn delivers a dramatic rate enhancement of the reactivity of the enzyme bound ALA as opposed to the substrate in its carbonyl form.

Scheme 5.10 The Schiff base formed between Lys 247 and ALA provides a good electrophile for an aldol condensation with C-3 of the second substrate molecule, a key reaction in PBG formation.



5.5.5 Mass spectrometric analysis of *E. coli* ALAD K247N mutant

The purified *E. coli* ALAD K247N mutant was buffer exchanged into 10mM Tris/HCl containing 50 μM and 10 mM β-mercaptoethanol, the concentration of the enzyme was adjusted to 1 mg/ml. An equal volume of 50% acetonitrile, 48% water and 2% formic acid was added to the protein solution. A volume of 20 μl was injected into an electrospray triple quandrapole mass spectrometer in positive ion mode. Data were acquired at a cone voltage of 40 V, capillary voltage of 3.5 kV and HV lens voltage of 0.30 kV between masses of 500 m/z and 2000 m/z. The raw data obtained from the above values was then subjected to maximum entropy analysis according to the Micromass schedule. The results of the mass spec. analysis showed the expected Mr. of 35480 for the mutant protein (fig 5.17).

5.5.6 Crystallisation of *E. coli* ALAD K247N mutant in complex with LA

Crystallisation was carried out by the vapour diffusion method using conditions similar to those previously described in chapter 2. The hanging drops contained 5 μl of concentrated protein solution (7 mg/ml) and 5 μl of well solution (200 mM Tris buffer, pH 8.3 containing 2% ammonium sulphate, 200 μM $ZnSO_4$, 6 mM β-mercaptoethanol, and 7 mM laevulinic acid)

5.5.7 Data collection and processing

The K247N ALAD mutant co-crystallised with LA in space group I422 (table 6.6). Cryoprotection was employed for data collection at 100 K by addition of the crystal to well solution containing 30% glycerol (v/v) and 40% saturated ammonium sulphate solution (v/v). The crystal was then flash cooled by immersion in liquid ethane, prior to storage under liquid nitrogen. Data from the K247N mutant ALAD (K247N-LA) complex were collected from cryocooled crystals on beamline 14 at the ESRF (Grenoble, France) on an 18 cm MAR image plate and were complete to 2.5 Å. Data were processed with the program MOSFLM (55) and the CCP4 suite (1994), for processing statistics see table 5.6. The position of the mutated residue and the LA molecule bound at the active site were located by calculation of difference Fourier maps phased with the native

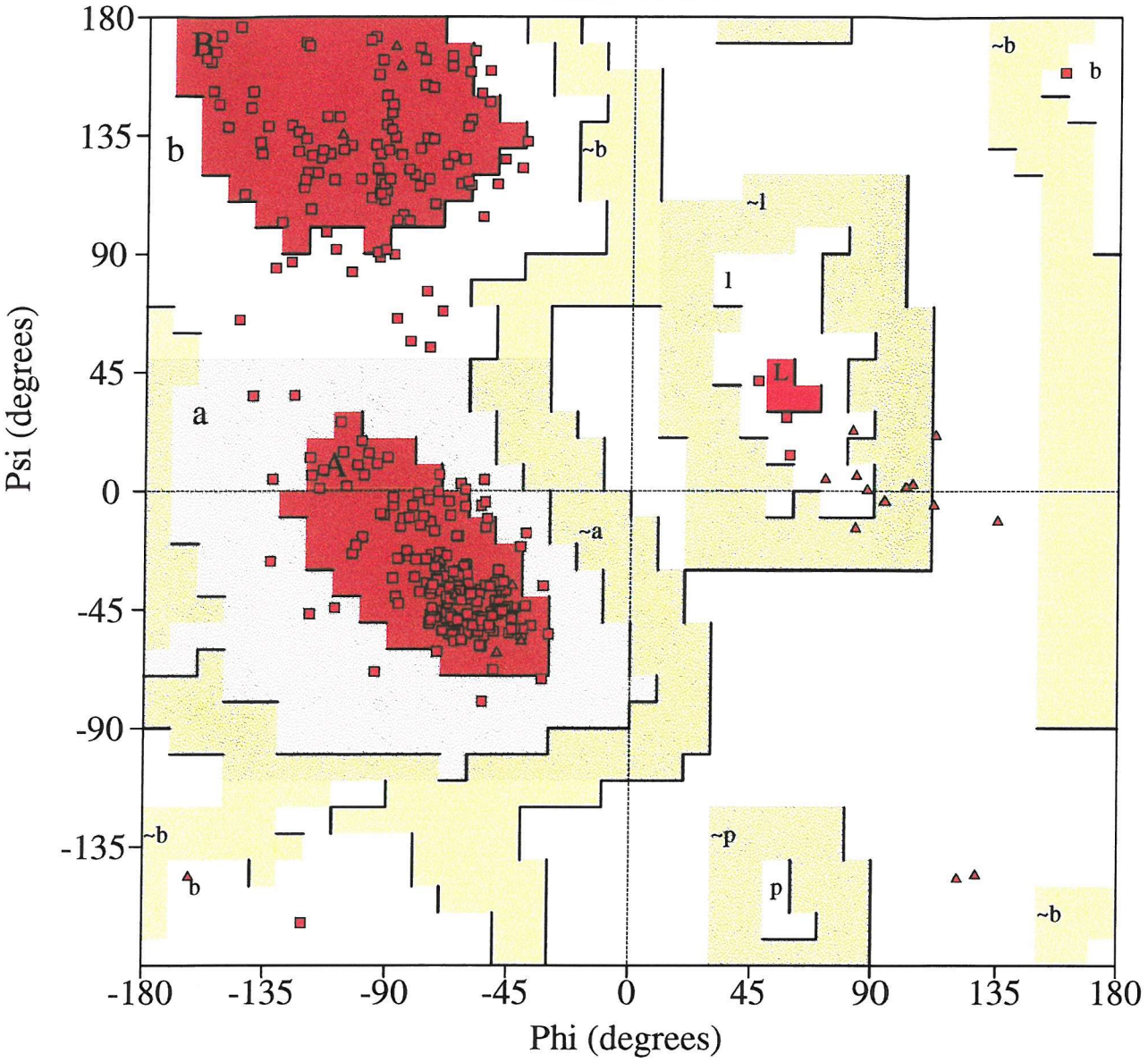
Table 5.6. Processing statistics for *E. coli* ALAD mutant K247N in complex with LA.

Space group	i422
a b Å	128.74
c Å	142.38
Entire data set	
Resolution range Å	33.15-2.5
R-merge (%)	7.2
Completeness (%)	99.5
Overall (I/Iσ)	7.4
Fraction I> 3σ (I)	99.5
Multiplicity	4.5
Outer Shell	
Resolution range	2.59-2.5
R merge (%)	11
I / σ I	97.1
Fraction I> 3σ (I)	6.1
Multiplicity	97.3
Refinement	
R factor (%)	24.2
Free R factor (%)	28.1
Resolution range	15-2.5
Cut off	2σ
Total number of reflections	20902
Number of atoms	3035

PROCHECK

Fig 5.17

Ramachandran Plot
K247N



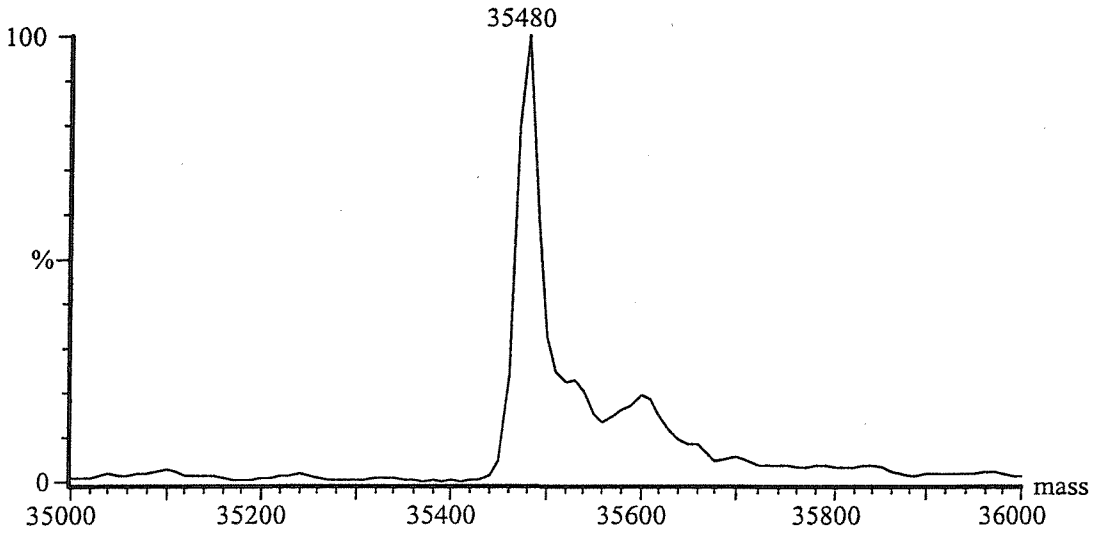
Plot statistics

Residues in most favoured regions [A,B,L]	247	85.8%
Residues in additional allowed regions [a,b,l,p]	41	14.2%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	288	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	11	
Total number of residues	323	

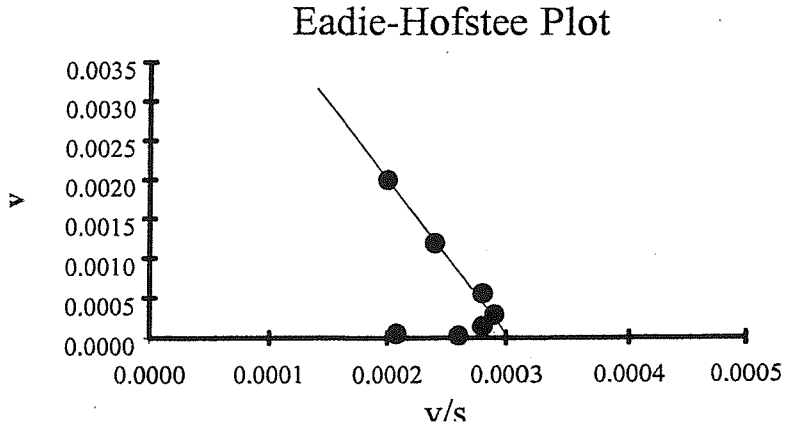
Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

247N_01.ps

Fig. 5.18 Mass spectrometric and kinetic analysis of E. coli ALAD mutant K247N
A. Mass spectrometric analysis of *E. coli* ALAD mutant K247N. Data shows Mr of 35480 ±1.89 The expected Mr = 35483.



B. Graph showing kinetic anaylsis of *E. coli* ALAD mutant K247N.
K_m = 19.4 ± 4.19 mM ; V_{max} = 0.003 ± 0.006μmoles PBG/mg/hr



E. coli ALAD structure (26). The structures were refined using X-PLOR (60) and were rebuilt using Quanta (62) running on Silicon-Graphics computers.

5.5.5 Quality of structure analysis

The structure of the *E. coli* ALAD K247N-LA complex was refined using data between 12.43 Å and 2.5 Å and had an R-factor of 24.2 % and a free R-factor of 28.1 % (see table 5.6). The structure evaluation program PROCHECK was used to asses the quality of the final model. A Ramachandran plot produced by the program showed 86 % of residues within the allowed regions and 100 % of residues within the additional allowed boundary (fig 5.17).

5.5.6 Location of the N247 side chain

The initial structural refinement of the K247N-LA complex was carried out using the native *E. coli* ALAD structure with residue 247 mutated to alanine, in order to prevent bias of difference Fourier map calculation.

On inspection of the Fo-Fc and 2Fo-Fc maps it was possible to locate a strong region of difference density in proximity to the α-carbon of the mutated A247 residue (fig 5.19). The residue was then substituted by asparagine and its side chain fitted to this density. After refinement, the N247 side chain was well defined by the resulting 2Fo-Fc map (fig 5.20).

5.5.7 Identification of the LA molecule bound at the P-site of K247N ALAD

After successful positioning of the N247 side chain, a separate region of difference density remained at the P-site of the active site (fig. 5.21). A molecule of LA was built into this density and the structure underwent further refinement. The inhibitor molecule is well defined by the final 2Fo-Fc map and is in close proximity to the N247 side chain.

5.5.8 Binding of carboxyl group of LA molecule at P-site

The carboxyl group of the bound LA molecule forms 3 hydrogen bonds, one with the hydroxyl oxygen of Tyr 312 and two with the hydroxyl group and main chain nitrogen of Ser 273, both invariant residues among known ALAD sequences. Several other invariant residues interact with the bound LA including Phe 204

Fig 5.19 Fo-Fc map showing P-site of K247N *E. coli* ALAD mutant contoured to 2 sigma. K247 has been mutated to alanine in order to prevent influence during refinement.

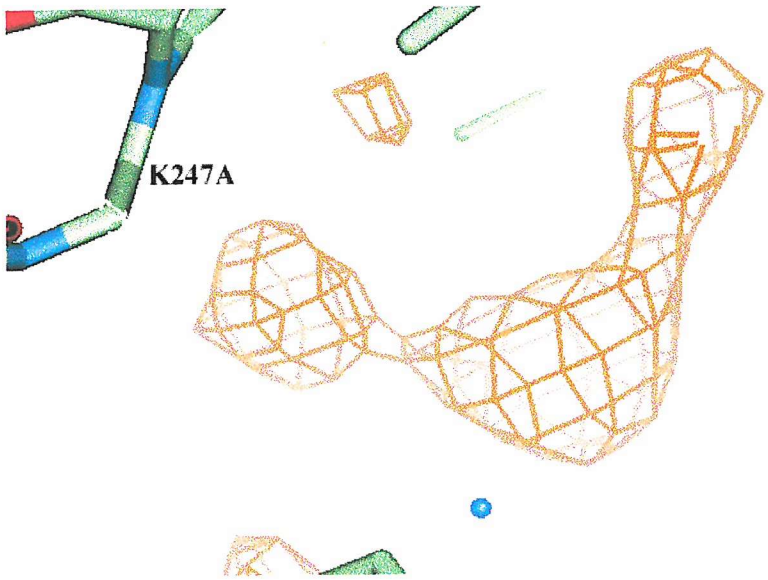
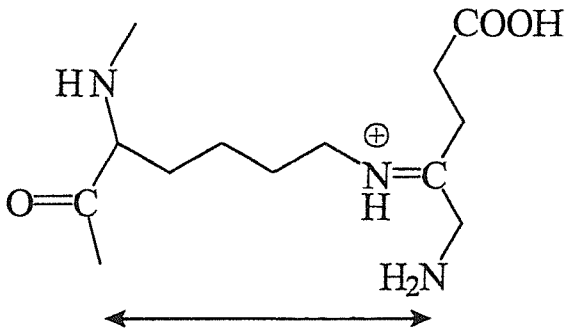


Fig 5.20 Fo-Fc and 2Fo-Fc maps (contoured to 2 sigma) showing P-site of *E. coli* ALAD mutant. Asn has now been inserted and is well defined by the 2Fo-Fc map.

Fig 5.21 2Fo-Fc map, contoured at 2 sigma, showing P-site of *E. coli* ALAD K247N mutant. LA has been inserted and is well defined by the map.

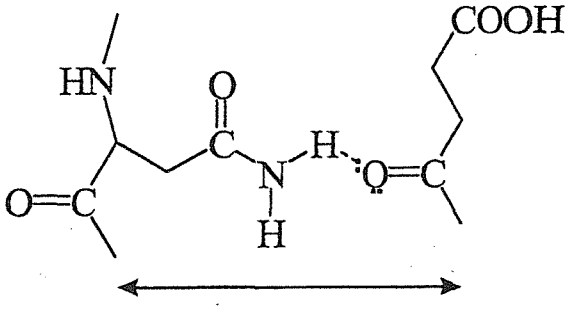
Scheme 5.11 Comparison of Schiff base formation between *E. coli* ALAD residue Lys 247 and binding of the ALA carbonyl group by the mutant Asn 247 side chain. Although the side chains of Asn and Lys are of different lengths, the distance between the C-4 of ALA and the α carbon of Lys observed in the native enzyme (A) is very similar to that between the C-4 carbon of LA and the α-carbon of Asn in the mutant (B).

A. Schiff base formation between Lys 247 and ALA



7.059 Å

B. LA molecule hydrogen bondingwith mutant residue K247N



7.332 Å

forms part of the active site loop and packs tightly against the inhibitor molecule. This binding is identical to that found in the native *E. coli* ALAD-LA complex (26).

5.5.9 Interaction between the N247 side chain and the LA molecule

The native *E. coli* ALAD-LA structure reveals a Schiff base formed between K247 and the C-4 carbonyl of the LA molecule. Although the N247 residue is unable to form a Schiff base with the LA, the carbonyl oxygen forms a strong H-bond with the N-H group of the mutant Asn side chain (scheme 5.11). The amino group of the adjacent Lys 195 also makes a reasonable H-bond with the carbonyl oxygen of the mutant Asn side chain.

As the binding of LA to the P-side of the active site of K247N is similar to that of the wild type ALAD-LA structure (26) and to that of the ALAD-ALA structure (chapter 3), it is deduced that ALA will bind to K247N in the same position as the LA molecule.

The complete inactivity of K247A and K247C (section 5.1) is in contrast to the measurable activity of K247N (fig. 5.18). This activity can be attributed to the activation of the C-4 carbonyl of ALA to act as an electrophile, presumably facilitated by binding *via* a H-bond with N247. Since the K247N mutant cannot form a Schiff base with the P-site ALA, the contribution to PBG formation made by the Schiff base formed between K247 and ALA may be equated with the contribution made with the carbonyl alone in the K247N mutant and be directly measured. Kinetic analysis shows that wild type ALAD is 10,000 times more active than the K247N mutant with a Km of 19.4 mM and a k_{cat}, of 0.003, revealing the importance of this initial imine formation to substrate binding and catalysis.

Chapter 6. Implications of research for catalytic mechanism of PBG formation

ALA is a highly reactive molecule and unless enzymatically controlled will spontaneously form products other than PBG such as a symmetric pyrazine. For this reason the correct "head to head" positioning of the two substrate molecules in the active site is equally important as the catalytic influence of ALAD.

Initial binding of ALA occurs at the P-site of *E. coli* ALAD. The carboxyl group of ALA forms hydrogen bonds with invariant residues Ser 273 and Tyr 312. The amino group of the substrate ALA forms a H-bond with Asp 118 which itself is closely associated with Ser 165. This positioning of ALA places the carbonyl oxygen of the ALA molecule in close proximity to the ε-nitrogen of Lys 247.

Once P-site ALA is bound in the correct position, Schiff base formation between Lys 247 and ALA takes place. The ε-NH_2 group of Lys 247 is unprotonated due to the positive charge of adjacent residue Lys 195 allowing it to act as a nucleophile and attack the δ+ C-4 carbonyl carbon of substrate ALA, forming a Schiff base and eliminating water. The positive charge carried by the newly formed Schiff base suppresses the pK_a of Lys 195 and the amino group of the substrate ALA which is agreement with observations made by Jaffe et al (64,65).

The binding of ALA at the P-site causes the hydrophobic side chain of Phe 204 to pack against the methylene groups of the substrate molecule which in turn leads to the docking of the flexible active site loop (residues 197-220) in place and brings residues Arg 205 and Arg 215 into the A-site ALA binding cavity. Once the A-site is formed, ALA can bind, with its carboxyl group forming a salt bridge with Arg 205 and the carbonyl group orientated towards the zinc ion, placing both bound A- and P-site ALA molecules in the essential "head to head" position.

The two key events in PBG catalysis are the formation of a C-C bond from an aldol condensation and a C-N bond resulting from an intersubstrate Schiff base. These bonds link the substrate molecules and complete the pyrrole ring. Although it is not possible to discriminate between the two possible mechanistic pathways, both are

shown in schemes 6.1a (Schiff base formation first) and 6.1.b (aldol condensation first), it is now possible to confidently speculate on the roles played by key active site residues in PBG catalysis.

The incoming amino group of A-site ALA is forced into close proximity to the amino group of P-site ALA causing the H-bond between Asp 118 and the amino group of P-site ALA to be broken and leaving the P-site substrate molecule able to change conformation. The new position of the P-site ALA places its nucleophilic - NH_2 group nearer to the C-4 carbon of A-site ALA, facilitating Schiff base formation. The carbonyl oxygen of A-site ALA is suitably placed to abstract a proton from a water molecule bound to the adjacent zinc ion and is eliminated.

The next step in PBG formation is the abstraction of protons from the C-3 position of A-site ALA. The zinc bound hydroxide ion could now act as a base to remove a proton in the first step of the aldol condensation between the A- and P-site substrate molecules (fig 6.5). The aldol condensation to form the C-C bond between the C-3 carbon of A-site ALA and the C-4 Schiff base carbon of P-site ALA can now occur, closing the five membered ring of PBG.

The ε-nitrogen of Lys 247 is now protonated by the adjacent Lys 195 and the zinc bound hydroxyl again removes a proton from the C-3 carbon of A-site ALA causing the ε-nitrogen of Lys 247 to be eliminated from the forming PBG molecule The regenerated Lys 247 acts as a base to remove a *pro-R* hydrogen stereospecifically from the C-5 carbon of P-site ALA (80) promoting the aromatisation of the pyrrole ring and the completion of PBG formation.

The two key events in PBG catalysis are the formation of a C-C bond from an aldol condensation and a C-N bond resulting from an intersubstrate Schiff base. These bonds link the substrate molecules and complete the 5-membered ring. Although this research has not clarified if it is the C-C bond or the C-N bond which is first formed, the roles of key catalytic and binding residues in the active site have been elucidated and their contribution to PBG formation described.

Lys 247

It has long been established that the initial reaction between *E. coli* ALAD and ALA is the formation of a Schiff base between the carbonyl of substrate ALA and Lys 247. The determination of the reduced ALAD-ALA structure confirms these observations and clearly defines the position of ALA in the P-site. A second role for Lys 247 became apparent when the ALAD-PBG structure was visualised. The proximity of Lys 247 to the C-2 carbon atom of PBG, originally the C-5 carbon atom of P-site ALA, identifies this residue as the base which is most likely to be able to remove the *pro-R* hydrogen in the final stages of PBG formation.

Although the most important catalytic residue in the active site, the mutation of Lys247 to Asn does not completely inactivate *E. coli* ALAD demonstrating the high reactivity of the substrate ALA molecules.

Lys 195

The primary function of Lys 195 in PBG production is to suppress the pk_a of the ε-nitrogen of Lys 247 allowing it to exist it an unprotonated state and hence to act as a nucleophile on formation of the Schiff base with substrate ALA. The pk_a of the amino group of Lys 247 has been directly measured as 7.5 (chapter 5), proving it to be deprotonated at physiological pH. Mutation of Lys195 dramatically reduces ALAD activity since Lys 247 is unable to form a Schiff base with P-site ALA. Conversely, Lys 195 acts to protonate the ε-nitrogen of Lys 247 during catalysis, making it a good leaving group which can be eliminated from the forming PBG molecule.

Ser 273 and Tyr 312

The carboxyl binding region of the P-site ALA binding site is formed by Ser 165 and Tyr 312. Although it does not appear that these residues play a direct catalytic role they are instrumental in orientating P-site ALA to the correct position for PBG formation.

Asp 118 and Ser 165

Asp 118 forms the amino group binding site for P-site ALA and is hydrogen bonded to the adjacent residue Ser 165. As with Ser 273 and Tyr 312 the role of Asp 118 is to bind P-site ALA in the correct orientation although the hydrogen bond formed with the amino group of P-site ALA is perturbed by the incoming amino group of A-site ALA.

Phe 204

Once P-site ALA has been bound, the hydrophobic side chain of Phe 204 packs against the methylene groups of the substrate molecule The dual effects of this interaction are to shield the forming PBG molecule from solvent water and to immobilise the flexible active site loop, of which Phe 204 is part, allowing key loop residues to form part of the A-site ALA binding site.

Arg 205 and Arg 215

The binding site of A-site ALA carboxyl group is formed by residues Arg205 and Arg 215 which are both located in the flexible active site loop region (residue 197-220). After binding P-site ALA the loop becomes less flexible and the two arginine residues come into the A-site active site to form a positive binding region for the carboxyl group of A-site substrate molecule. As with the P-site ALA binding region, the orientation of ALA is crucial to PBG formation, and the positioning of the A-site ALA carboxyl group binding site ensures both substrate ALA molecules are bound in a "head to head" formation.

Active site zinc ion

The active site zinc ion is positioned at the A-site ALA binding site and is co-ordinated by three highly conserved cysteine residues and a water molecule. The zinc ion is involved in both ALA binding, by co-ordinating the carbonyl of A-site ALA, and in catalysis, by acting as a base during deprotonation of the C-3 carbon of A-site ALA. A number of studies of the metal binding properties of ALAD have been carried out.

Studies carried out on human ALAD, also a zinc-dependent enzyme, show that the divalent ion cadmium causes a slight but significant increase in enzyme activity at low concentrations (< 3 μM) (81). It is probable that the cadmium ion is functioning as a metal hydroxide at the active site zinc binding site. The increased activity of cadmium substituted human ALAD is thought to be due to the decreased polarisation of the Cd-OH bond in comparison to the Zn-OH bond. The cadmium hydroxide is therefore more basic and may be more effective at abstracting protons from the C-3 carbon of A-site ALA, which is the proposed function for the zinc hydroxide at the active site.

Lead has been shown to be a potent inhibitor of ALAD (82) with a K_i of 28.1 μM and is a non competitive inhibitor. Co-crystallisation of yeast ALAD with lead (25) shows lead ions occupying the triple cysteine zinc binding site at the active site.

E. coli ALAD activity has also been shown to be stimulated by magnesium, even at optimal zinc concentrations (13). Analysis of the crystal structure of *E. coli* ALAD (26) showed the most likely place for magnesium to bind is at the second zinc binding site. Although too far away from the active site to be directly involved in catalysis, the metal ion may have an effect on the active site loop (residues 197-220) which passes close to the second metal binding site.

The amino group of A-site ALA

Although not part of the active site, the amino group of ALA seems to play an essential role in the mechanism of ALAD catalysed PBG formation. On binding at the A-site, the positively charged $-NH_3$ group of A-site ALA could disrupt the H-bonding between the amino group of P-site ALA and Asp 118 causing it to shift towards the C-4 carbonyl carbon of A-site ALA and facilitating the Schiff base formation between the two substrate molecules. The presence of the amino group of A-site ALA has been shown to be vital in ALAD catalysed PBG formation. Spencer *et al.*, (18) demonstrated that whilst the competitive inhibitor laevulinic acid is able to bind and form a Schiff base at the P-site of *E. coli* ALAD, it is unable to bind at the A-site.

Further work

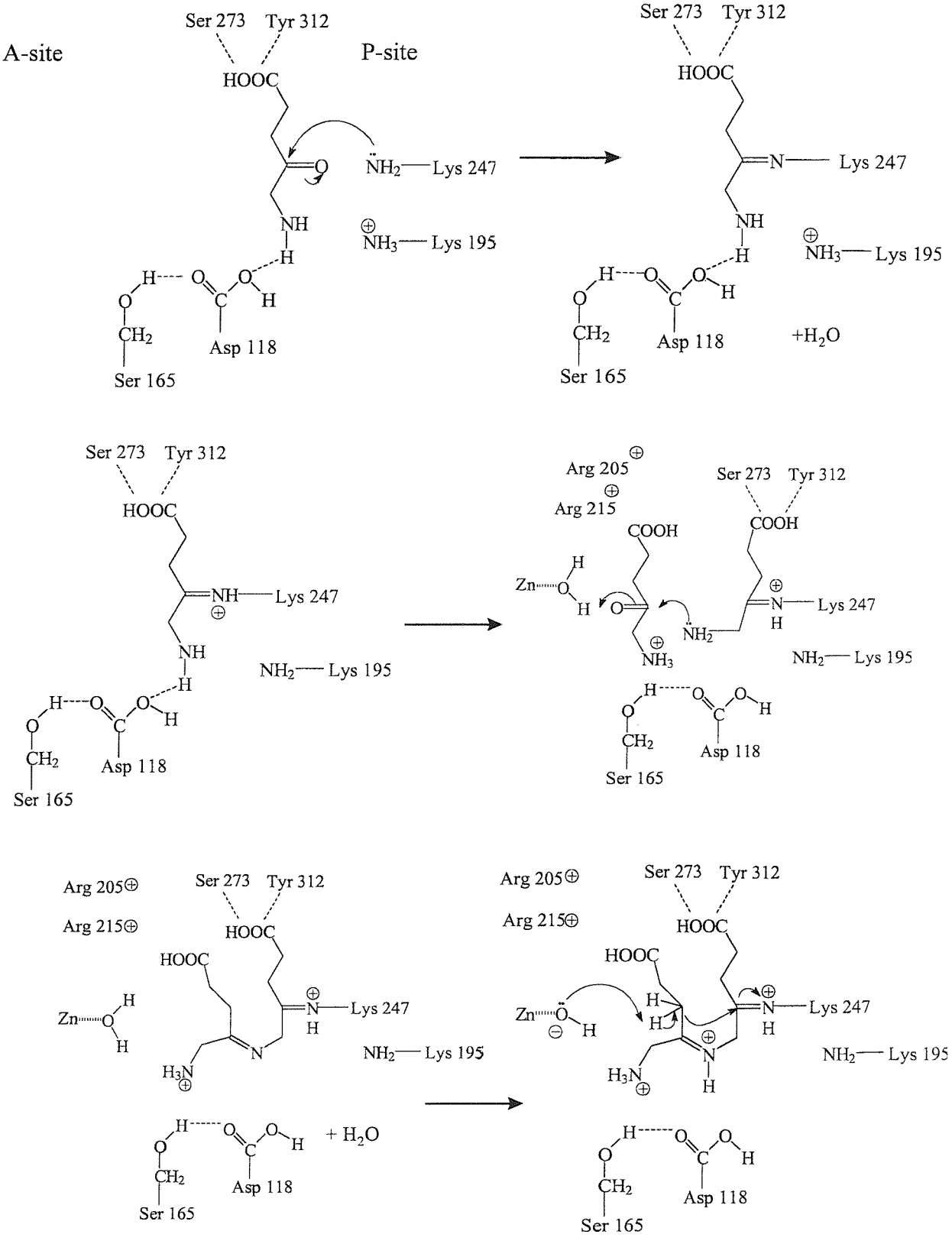
The main objective of any further investigation into the mechanism of ALAD catalysed PBG formation will be to determine the mechanistic pathway by which the reaction proceeds i.e. whether it is an aldol condensation or a Schiff base that initially links the two substrate molecules of ALA.

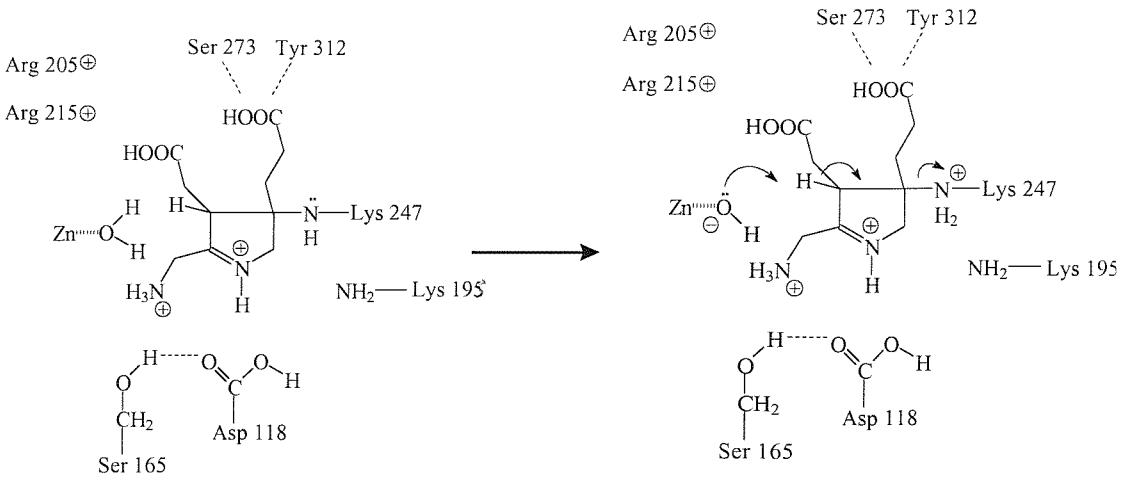
The possibility of trapping an intermediate *via* reduction of the Schiff base formed between the two substrate molecules and mass spectrometric analysis of a resulting protein-intermediate complex is a potential route to distinguishing between the two mechanisms, as is the use of specifically synthesised ALA and PBG analogues.

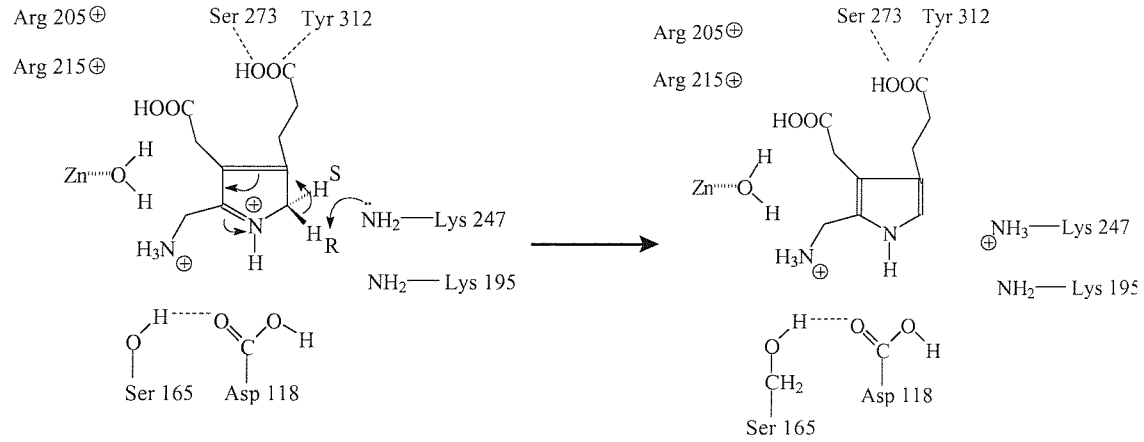
It is also possible that higher resolution data collected on the reduced ALAD-ALA + ALA complex would yield additional information on the exact atomic positioning of both P- and A-site substrate molecules.

Whichever method of study is used, the reactivity of ALA and the speed at which this reaction proceeds once both substrate molecules are bound in the correct orientation at the active site makes investigation into the mechanistic pathway very challenging.

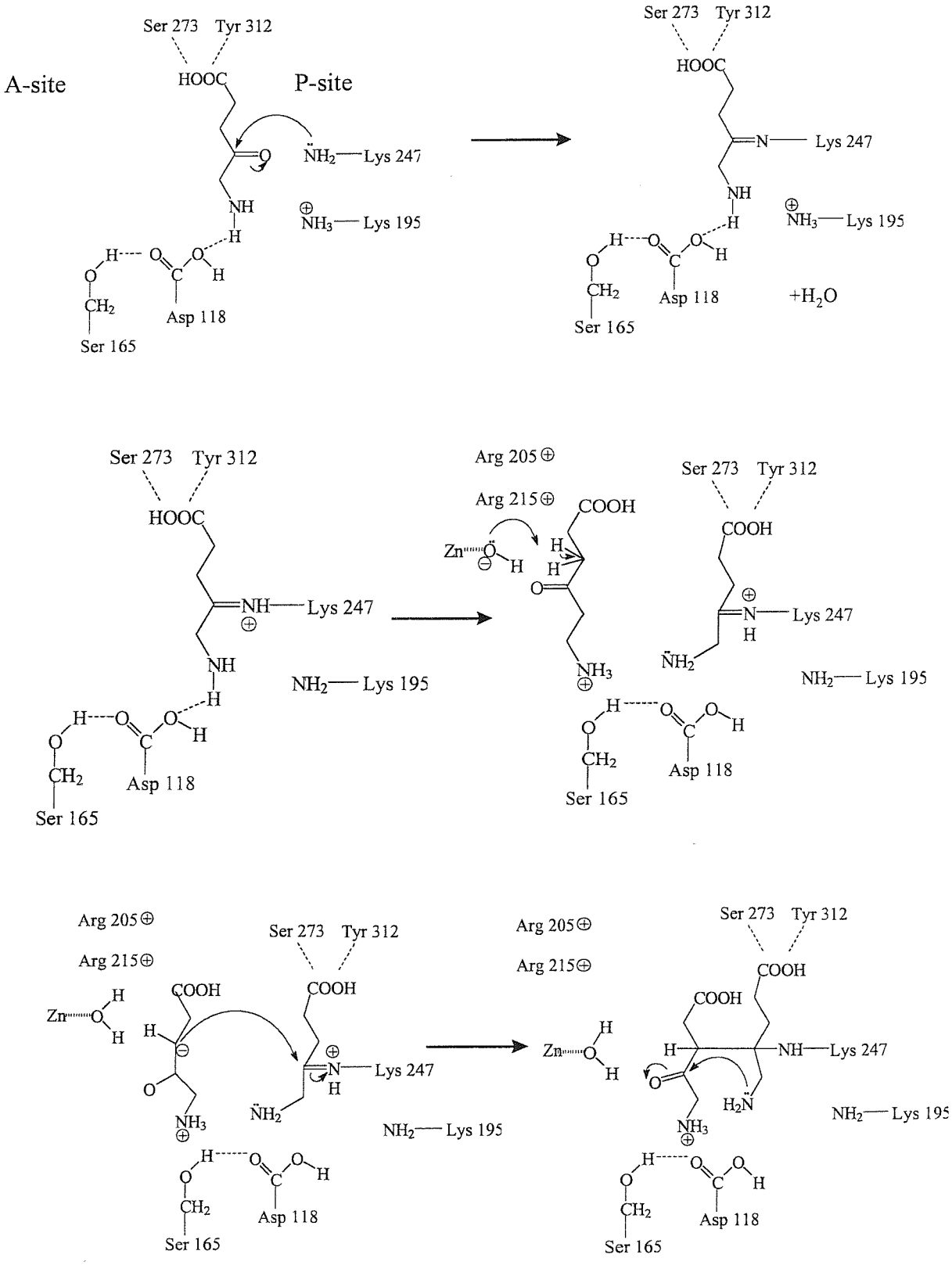
Scheme 6.1 a *E. coli* ALAD catalysed PBG formation. The reaction proceeds *via* an Schiff base formation between the amino group of P-site ALA and the C-4 carbonyl carbon of A-site ALA.

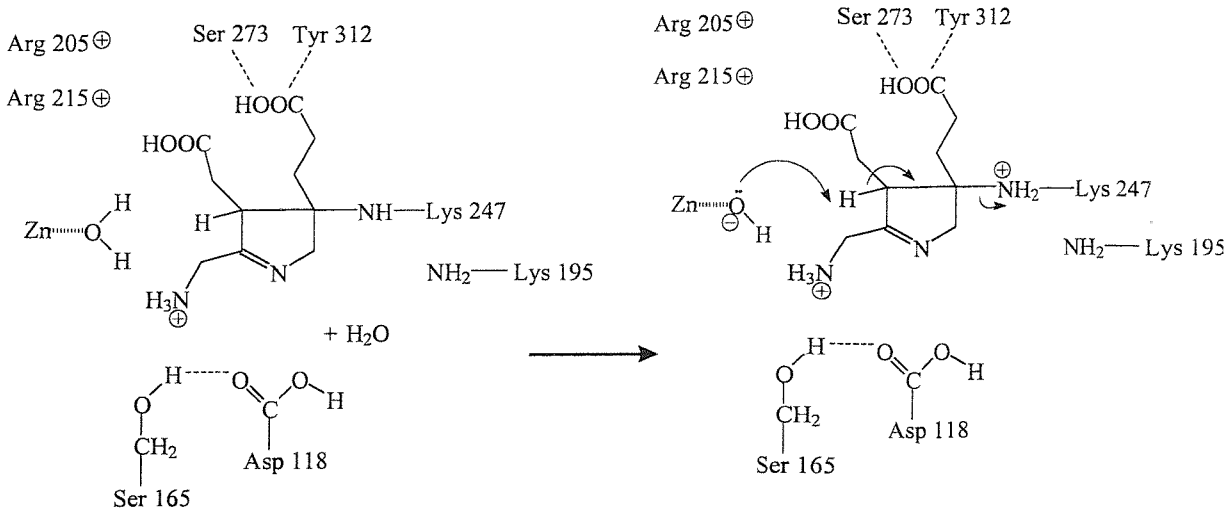


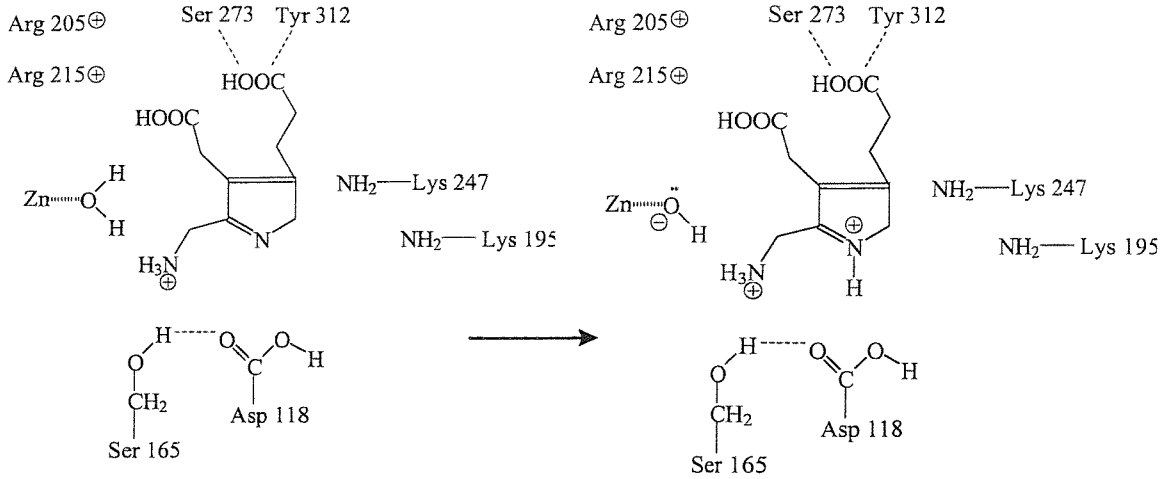

Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
O
H
H
H
N
Lys 247
H
H3N⊕
N
H
NH2— Lys 195
H
O
O
C
O
H
CH2
Asp 118
Ser 165
Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
O
H
H
H
N⊕
Lys 247
H2
H3N⊕
N
H
NH2— Lys 195
H
O
O
C
O
H
CH2
Asp 118
Ser 165

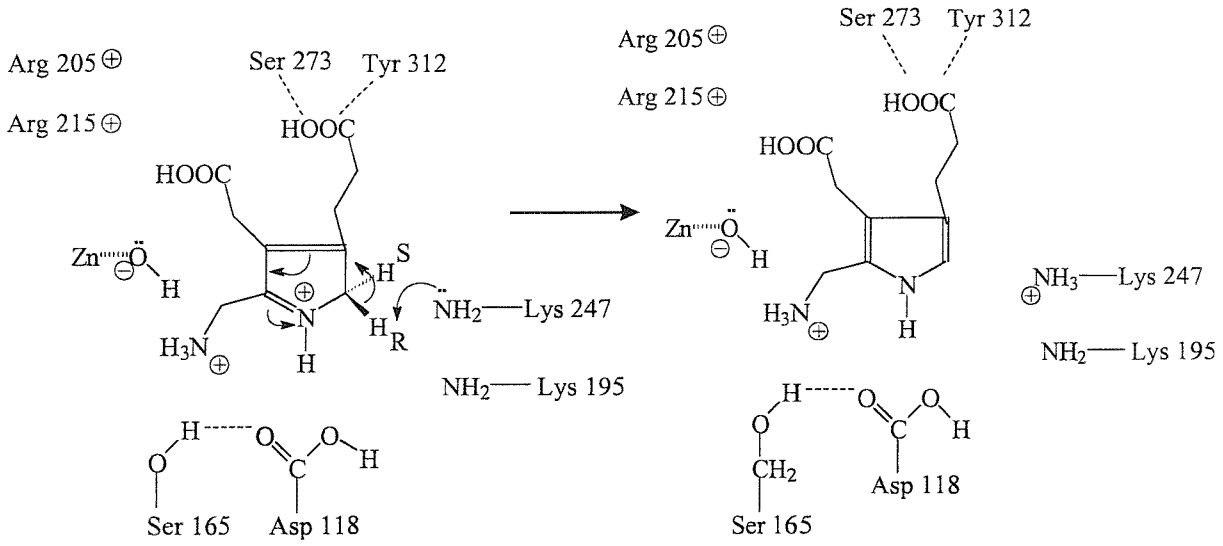

Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
O
H
H
H
S
H
N
H
H3N⊕
R
NH2—Lys 247
NH2— Lys 195
H
O
O
C
O
H
Ser 165
Asp 118
Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
O
H
H
N
H
H3N⊕
NH3⊕—Lys 247
NH2— Lys 195
H
O
O
C
O
H
CH2
Asp 118
Ser 165

Scheme 6.1 b *E. coli* ALAD catalysed PBG formation. An aldol condensation forms the initial bond between two substrate ALA molecules.




Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
O
H
H
NH—Lys 247
N
H3N⊕
NH2— Lys 195
+ H2O
H
O
CH2
Ser 165
Asp 118
Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
NH2—Lys 247
NH2— Lys 195
H3N⊕
N
H
O
CH2
Ser 165
Asp 118


Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
O
H
NH2—Lys 247
NH2— Lys 195
H3N⊕
N
H
O
CH2
Ser 165
Asp 118
Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
NH2—Lys 247
NH2— Lys 195
H3N⊕
N
H
O
CH2
Ser 165
Asp 118


Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
O
H
H
S
H
R
NH2—Lys 247
NH2— Lys 195
H3N⊕
N
H
O
CH2
Ser 165
Asp 118
Arg 205⊕
Arg 215⊕
Ser 273
Tyr 312
HOOC
HOOC
Zn
H
NH3⊕—Lys 247
NH2— Lys 195
H3N⊕
N
H
O
CH2
Ser 165
Asp 118

References

1. Beale S I and P.A. Castlefranco. Plant Physiology., 1974. **53**. 297

2. Wang W-Y., Gough S.P. and .Kannangara C.G. Carlsburg Research Communications, 1981, **46**, 243.

3. Bruyant, P and Kannangara C.G Carlsburg Research Communications, 1987, **52** 99-109.

4. Kannangara C.G., Gough S.P., Bruyant., P., Hoober, J.K., Khan, A., and Von Wettstein, D., Trends in Biomedical Science 1988, **13**, 39.

5. Urban-Grimal, D., Volland, C., Garnier, T.,and Labbe-Bois, R. European Journal of Biochemistry, **150**, 481-484

6. Jordan, P.M and Shemin D., (1972) in : The Enzymes (3rd edition) (P.D. Boyer Ed.) **5**, pp. 323-356, Academic Press

7. Lelean J. E. Lloyd, A. and Shoolingin-Jordan, P.M. Methods in Enzymology., 1997, **218.** 309.

8. Fanica-Gaignier, M., and Clement-Metral, J., European Journal of Biochemistry, 1973, **40**, 19-24.

9. Gibson, K.D., Neuberger, A., and Scott, J.J., Biochemical Journal 1955, **61**, 618-629.

10. Schmid, R., and Shemin, D., Journal of the American Chemical Society, 1955, **77**, 506-508.

11. Anderson, P.M., and Desnick, R.J., Journal of Biological Chemistry 1979 **254**, 6924-6930

12. Leidgens, W., Lutz, C. and Schneider, H.A.W. European Journal of Biochemistry, 1983, **135**, 75-70

13. Spencer, P and Jordan P.M., Biochemical Journal 1993, **290**, 279-287

14. Simmons, T.B.J. European Journal of Biochemistry, 1995, **234**, 178-183

15. Brennan, M.J.W. and Cantrill, R.C. Nature 1979, **280**, 514-515

16. Chaudry, A.G., Gore, M.G. and Jordan, P.M., Biochemical Society Transactions, 1976, **4(2)** 301-303

17. Jordan, P.M., and Seerha, J.S., FEBS Letters 1980, **114**, 283-286

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

19. Appleton, D., Duguid, A.B., Lee, S.K., Ha, Y-J., Ha, H.J., Leeper, F.J., 1998, Journal of the Chemical Society, Perkins Transactions, 1998 , **1**, 89-101

20. Cheung, K-M., Spencer, P., Timko, M.P., and Shoolingin-Jordan, P.M., Biochemistry, 1997, **36**, 1148-1156

21. Senior, N., Thomas, P.G., Cooper, J.B., Wood, S.P., Erskine, P.T., Shoolingin-Jordan, P.M. and Warren, M.J., Biochemical Journal, 1996, **320**, 401-412.

22. Banner, D.W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C.I., Wilson., I. A., Corran, P.H., Furth, A.J., Milma, J.D., Offord, R.E., Priddle, J.D., and Waley, S.G., Nature, 1975, **255**, 609-614

23. Xia, Z.X., Shamala, N., Bethge, P.H., Lim, L.W., Bellamy, H.D., Xuong, N.H., Lederer, and Mathews, F.S., Proceedings of the National Academy of Science, 1987, **84,** 2629-2633

24. Priestle, J.P., Grutter, M.G., White, J.L., Vincent, M.G., Kania, M., Wilson, E., Jardetzky, T.S., Kirschner, K., Jansonius, J.N., Proceedings of the National Academy of Science, 1987, **84(16),** 5690-5694

25. Erskine, P.T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, I.J., Sarwar, M., Spencer, P., Thomas, P., Warren, M.J., Shoolingin-Jordan, P.M., Wood, S.P., Cooper, J.B. (1997) *Nature Struct. Biol.* **4** 1025-1031

26. Erskine, P.T., Norton, E., Cooper, J.B., Lambert, R., Coker, A., Lewis, G., Spencer, P., Sarwar, M., Wood, S.P., Warren, M.J., Shoolingin-Jordan, P.M. (1999) *Biochemistry* **38** 4266-4276

27. Frankenberg, N., Erskine, P.T., Cooper, J.B., Shoolingin-Jordan, P.M., Jahn, D., Heinz, D.W. (1999) *J. Mol. Biol.* **259** 591- 602

28. Gang Guo, G., Gu, M., and Etlinger, J.D., Journal of Biological Chemistry, 1994, **269**, 12399-12402

29. Weissman, J.S., Sigler, P.B. and Horwich, A.L., Science 1995, 268, 523-524

30. Doss, M., von Tiepermann, R., Schneider, J., and Schmid, H., Klin. Wochenshcr.1979, **57**, 1123-1127

31. Jordan P.M. and Warren M.J., FEBS Letters., 1987, **225**, 87

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

33. Luke McNeil, personal communication

34. Shoolingin-Jordan., P.M., Journal of Bioenergy and Biomembranes, 1995, **27**, 181.

35. Lambert, R., Brownlie, P.D., Woodcock, S.C., Louie, G.V., Cooper, J.C., Warren, M.J., Jordan, P.M., Blundell, T.L. and Wood, S.P., Ciba Foundation Symposium 1994 pp 97-110

36. Mathewson, J.H. and Corwin, A.H., Journal of American Chemical Society, 1963, **83**,135-137

37. Leeper, F.J. Ciba Foundation Symposium 1994, pp 111-130

38. Alwan., F.A., Mgbeje., B.I.A and Jordan., P.M, Biochemical Journal, 1989, **264**, 397.

39. Phillips, J.D., Whitby, F.G., Kushner, J.P., Hill, C.P. (1997) *Protein Sci.* **6** 1343-1346

40. Tait., G.H Biochemical Journal., 1972, **128**, 1159

41. Xu, K., Delling, J., and Elliot, T., Journal of Bacteriology., 1992, **174**, 3953

42. Camadro, J.M., Chambon H., Jolles, J., and Labbe, P., European Journal of Biochemistry, 1986, **156**,579.

43. Poulson, R. and Polglase, W.J., Journal of biological chemistry., 1976, **250**, 1269.

44. Taketani, S. amd Tokunaga, R., Journal of Biological Chemistry 1981, **256**,12748

45. Camadro, J.M and Labbe, P., Journal of Biological Chemistry, 1988, **263**, 11675

46. Hansson, M. and Hederstedt, L. European Journal of Biochemistry, 1994, **220**, 201

47. Al-Karadaghi, S., Hannson, M., Nikolov, S., Jonsson, B., and Hederstedt, L., Structure 1997, **5**, 1501.

48. Innis, M.A., Gelfand, D.H., and Sninsky, J.J., PCR protocols: A guide to methods and applications, Academic Press, pp 177-179

49. Mauzerall, D. and Granick, S. Journal of Biological Chemistry, 1956, **219**, 435-446

50. Henderson, R., Proccedings of the Royal Society, London, B, p6.

51. Teng, T-Y., Journal of applied crystallography, 1990, **23**, p387

52. Wonacott, A.J., The Rotation Method in crystallography, North-Holland Publishers, Amsterdam. 1977.

53. Hendrix, J and Lentfer, A. EMBL research reports, 1988, p170.

54. Collaborative Computational project, Number 4, Acta Cryst., 1994, **D50**, p760

55. Leslie, A.G.W., MOSFLM user guide 1994

56. Rossman, M.G., and Blow, D.M. Acta Cryst. 1962, **15**, p24

57. Lifchitz, A. Acta. Cryst. 1983, **A39**, p130

58. Rao, S.N., Jyh-Hwang, J, and Hartsuck, J.A. Acta. Cryst., **A36**. p878

59. Crowther, R.A., and Blow, D.M., Acta Cryst., 1967, **23**, p544

60. Brunger, A.T., XPLOR Version 3.1: A system for X-Ray Crystallography and NMR. Yale University Press, New Haven, CT. 1992.

61. Gibbs, P.N.B., and Jordan, P.M., Biochemical Journal 1986, **236**, 447-451

62. QUANTA96 X-Ray Structure Analysis Users's Reference, June 1996. San Diego: Molecular Simulations (1996)

63. Erskine, P.T., Newbold, R, Coker, A., Warren, M.J., Shoolingin-Jordan, P.M., Wood, S.P., Cooper, J.B., Protein Science, 1999, **8**, No6. 1250-1256

64. Petrovich, R.M., and Jaffe, E.K., Biochemistry 1997, **36**, 13421-13427

65. Rajagopalan, J.S., Taylor, K.M., and Jaffe, E.K., Biochemistry 1993, **32**, 3965-3972

66. Woodcock, S.C., and Jordan, P.M., Biochemistry 1994, **33**, 2699-2695

67. Highbarger, L.A., and Gerlt, J.A., Biochemistry, 1996, **35**, 41-46

68. Toney, M.D., and Kirsch, F.J., Science, 1989, **243**, 1485-1488

69. Frillingos, S., and Kaback, H.R., Biochemistry, 1996, **35**, 13363-13367

70. Boehlein, S.K., Walworth. E.S., Richards, N.G.J., and Schuster, M.S., Journal of Biological Chemistry, 1997, 227, 12384-12392.

71. Westhiemer, F. H., Tetrahedron, 1995, 51,

72. Kokesh, F.C., Journal of the American Chemical Society, 1971, **93,** .7270

73. Frey, P.A., Journal of the American Chemical Society, 1971.93 p.7266

74. Pearl, P., and Blundell, T., FEBS Letters, 1984,**174**, No1, 96-101.

75. Davies, D.R., Annual Review of Biophysics and Biophysical Chemistry. 1990, **19**, 189-215.

76. Seerha, S. and Jordan, P.M., European Journal of Biochemistry, 1981, 113, 435-446.

77. Jaffe, E.K., Abrams, R.W., Kaempfen, H.X., and Harris, K.A., Biochemistry, 1992, **31**, 2113-2123.

78. Wood, W.A., The Enzymes, 3rd Edition, Editor P. Boyer (New York Academic Press), **Vol5**, p. 281

79. Meloche, H.P., and Mehler, L., Journal of Biological Chemistry, 1973, 284, 6333

80. Chaudhry, A.G., and Jordan, P.M. Biochemical Society Transactions, 1976, **4**, 760-761.

81. Gibbs, P.N., Ph.D. Thesis, 1984, Southampton University, UK.