X-ray structure of a putative reaction intermediate of 5-aminolaevulenic acid dehydratase

Peter T. ERSKINE*, Leighton COATES*, Danica BUTLER*, James H. YOUELL*, Amanda A. BRINDLEY†, Steve P. WOOD*, Martin J. WARREN†, Peter M. SHOOLINGIN-JORDAN* and Jonathan B. COOPER*

*Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton, SO16 7PX, U.K., and †School of Biological Sciences, Queen Mary, Mile End Road, London, E1 4NS, U.K.

The X-ray structure of yeast 5-aminolaevulenic acid dehydratase, in which the catalytic site of the enzyme is complexed with a putative cyclic intermediate composed of both substrate moieties, has been solved at 0.16 nm (1.6 Å) resolution. The cyclic intermediate is bound covalently to Lys\textsuperscript{263} with the amino group of the aminomethyl side chain ligated to the active-site zinc ion in a position normally occupied by a catalytic hydroxide ion. The cyclic intermediate is catalytically competent, as shown by its turnover in the presence of added substrate to form porphobilinogen. The findings, combined with those of previous studies, are consistent with a catalytic mechanism in which the C–C bond linking both substrates in the intermediate is formed before the C–N bond.

Key words: 5-aminolaevulenic acid dehydratase, catalytic mechanism, reaction intermediate, trapped intermediate, X-ray structure.

INTRODUCTION

The enzyme 5-aminolaevulenic acid dehydratase [ALAD; porphobilinogen (PBG) synthase, E.C. 4.2.1.24] catalyses an early step in the biosynthesis of tetrapyrroles involving the condensation of two 5-aminolaevulenic acid (ALA) molecules to form the pyrrole framework uroporphyrinogen III, which undergoes various modifications to form a variety of essential metallo-prosthetic groups, including haem, chlorophyll and the cobalamins. ALADs share a high degree of sequence identity, contain about 350 amino acids per subunit and are usually octameric. In humans, hereditary deficiencies in ALAD give rise to the rare disease Doss porphyria [6], which is replaced by two aspartate residues and an alanine residue in [5]. A more detailed characterization of the metal requirement of ALADs is given in [5].

Previously, the structures ALADs in which the P-site is occupied by a number of bound inhibitors have been analysed, and the structure of substrate ALA bound in this site has also been determined [16]. All ligands form a Schiff base with Lys\textsuperscript{263} and are held with their carboxy groups making hydrogen bonds with the conserved residues Ser\textsuperscript{260} and Tyr\textsuperscript{289}. Hydrophobic interactions are made with the side chains of residues Tyr\textsuperscript{216}, Phe\textsuperscript{217}, Tyr\textsuperscript{287} and Val\textsuperscript{290}, which effectively form four sides of this substrate-binding subsite. In yeast ALAD, ALA appears to be oriented in two conformations, but in both of these conformers the C-5 amino group interacts with the side chains of Asp\textsuperscript{31} and Ser\textsuperscript{79}. This implicates these two residues as having potential catalytic roles in the reaction mechanism. The A-site is a solvent-filled cavity bounded on one side by the P-site ALA molecule and on the other by two arginine residues (Arg\textsuperscript{230} and Arg\textsuperscript{232}) which are strongly conserved. The base of the A-site is formed by the zinc ion held by three cysteine residues. There is a solvent molecule bound to the zinc ion which is within H-bonding distance of the amino group of P-side ALA. This water molecule may be a zinc-bound hydroxide which abstracts a proton from C-3 of A-side ALA as a prerequisite for formation of the C–C bond between the two substrates.

Hitherto, mechanistic proposals for ALAD have been dominated by the experimental finding that the enzyme only forms a Schiff base with one of the two identical substrate molecules. This was based on labelling studies which showed only...
P. T. Erskine and others

Scheme 1 Reaction catalysed by ALAD

Figure 1 Tertiary structure of a monomer of yeast ALAD showing the active-site Lys263 along with the bound intermediate in ball-and-stick representation. The zinc ion is also shown (the larger, separate, black sphere). Residues 213–233 over the active site are omitted in order to show the position of the putative intermediate more clearly.

This inhibitor forms two Schiff bases at the active site, involving Lys263 as well as Lys210. One half of the molecule resides in the P-site and the other half fills the putative A-site with its carboxylic group interacting with invariant arginine residues Arg220 and Arg232. The structure of Pseudomonas aeruginosa ALAD complexed with 5-fluorolaevulinic acid [19] shows that both A- and P-sites are occupied by an inhibitor molecule covalently bound through a Schiff base with Lys253 and Lys263 respectively. The structural evidence that both invariant lysine residues may form Schiff bases with substrate at the active site is suggestive of a double-Schiff-base mechanism such as that proposed by Neier [20].

Previously, we found that ALA was covalently bound as a Schiff base to Lys263 at the P-site of yeast ALAD, despite the fact that ALA had not been added to the enzyme [16]. It has also been observed that human ALAD, highly purified from erythrocytes, possesses a ligand resembling PBG at the active site, despite the fact that no ligands had been externally added [12]. In the present study, we have found that crystallizing the purified yeast dehydratase in the presence of added 5-aminolaevulinic acid results in the trapping of what appears to be a similar cyclic intermediate complex, arising from the reaction of two substrate molecules at the active site. The intermediate is catalytically competent and can turn over to form the product, PBG. The complex, refined at 0.16 nm (1.6 Å) resolution, shows a covalent link between Lys253 and the intermediate, a direct link with the zinc ion, and provides important stereochemical information giving an unprecedented view of active-site interactions. A catalytic mechanism taking into account the structure of the enzyme–intermediate complex is proposed.

EXPERIMENTAL

Recombinant yeast ALAD was purified as described previously [9]. Crystals of yeast ALAD with the bound intermediate were obtained by co-crystallizing the enzyme in the presence of substrate (ALA). The hanging-drop vapour-diffusion method was used with ALA present at a concentration of 0.2 mM and enzyme present at a concentration of 2 mg/ml. Apart from the presence of ALA, the crystallization conditions [2–10 % (w/v) poly(ethylene glycol) of molecular mass 6000 Da/200 mM Tris/HCl, pH range 7.0–8.0] are otherwise identical with those used to crystallize the native enzyme [21]. Crystals of the complex, which appeared within 3–4 weeks, belong to space group I422 with unit cell parameters of a = b = 10.32 nm (103.2 Å), c = 16.7 nm (167.7 Å). The crystals were flash-cooled in liquid ethane and X-ray data were collected at the European Synchrotron Radiation Facility (ESRF; Grenoble, France) using the ID-29 beam line with the crystal maintained at a temperature of 100 K using an Oxford Cryosystems (Long Hanborough, Oxford, U.K.) cooler. The data were processed using MOSFLM [22] and the CCP4 suite [23]. The structure was refined using SHELX [24] and the program RESTRAIN [25] was used for refinement of translation-libration-screw (TLS) tensors. Graphical rebuilding was done using TURBO-FRODO (Bio-Graphics, Marseille, France) running on Silicon-Graphics (SGI) computers. The co-ordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1ohl and r1ohlsf respectively.

RESULTS AND DISCUSSION

Interpretation of the electron-density difference map

The initial difference map for yeast ALAD co-crystallized with ALA showed strong electron density for a five-membered ring...
Figure 2  Stereo view of the electron-density map of the putative intermediate (labelled *) covalently bound to Lys263 at 0.16 nm (1.6 Å) resolution

The map was weighted as described by Read [32], and is contoured at 1.0 root mean square. The P-site is on the right-hand side and the A-site is formed by residues shown on the left-hand side including the zinc ion. The other invariant lysine (Lys321) is in the foreground just to the left of Lys263, but it has been omitted for clarity. The side chains of the two lysine residues lie approximately parallel with each other.

The final structure, which has been refined using data between 4.5 nm (45 Å) and 0.16 nm (1.6 Å), has an R-factor of 18.8% and an R-free of 24.1% (see Table 1). Refinement of the TLS tensors for the TIM barrel and arm domains gave a significant improvement in R-free of over 7%. The structure has 90.8% of its residues within the ‘most favoured’ regions of the Ramachandran plot by the PROCHECK criteria [26] and 8.6% of residues within the so-called ‘additional allowed’ boundary. No residues are in the disallowed regions. The temperature factors are reasonable, except for a few residues in the active-site flap (most notably 228–230), where the electron density remains poor despite extensive efforts to rebuild these residues throughout the refinement process.

Table 1  X-ray statistics for the complex of yeast ALAD with the putative intermediate after TLS refinement

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics for entire dataset</td>
<td></td>
</tr>
<tr>
<td>Resolution range [nm (Å)]</td>
<td>4.47–0.16 (44.7–1.6)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>5.5</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.2</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>7.2</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>21.7</td>
</tr>
<tr>
<td>%I &gt; 3σ(I)</td>
<td>77.5</td>
</tr>
<tr>
<td>Statistics for outer shell</td>
<td></td>
</tr>
<tr>
<td>Resolution range [nm (Å)]</td>
<td>0.17–0.16 (1.7–1.6)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>38.8</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.2</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>2.6</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>2.5</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>R-factor (%)</td>
<td>18.8</td>
</tr>
<tr>
<td>R-free (%)</td>
<td>24.1</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>57 545</td>
</tr>
<tr>
<td>RMSD bond lengths [nm (Å)]</td>
<td>0.0005 (0.005)</td>
</tr>
<tr>
<td>RMSD 1–3 distances [nm (Å)]</td>
<td>0.0013 (0.013)</td>
</tr>
<tr>
<td>RMSD bumps [nm (Å)]</td>
<td>0.0015 (0.015)</td>
</tr>
<tr>
<td>RMSD chiral tetrahedra [nm (Å)]</td>
<td>0.0009 (0.009)</td>
</tr>
<tr>
<td>RMSD main-chain planes [nm (Å)]</td>
<td>0.0009 (0.009)</td>
</tr>
<tr>
<td>RMSD side-chain planes [nm (Å)]</td>
<td>0.0007 (0.007)</td>
</tr>
</tbody>
</table>
Nonetheless the electron density for the active-site loop in this complex is substantially better than that of the native enzyme.

**Interactions made by the putative intermediate**

The active site of ALAD is located in a large cavity at the C-terminal end of the eight-stranded all-parallel $\beta$-barrel (Figure 1). It involves two invariant lysine residues (Lys$^{263}$ and Lys$^{210}$ in yeast ALAD) as well as a zinc ion which is co-ordinated by three cysteine residues (Cys$^{133}$, Cys$^{135}$ and Cys$^{143}$) in the yeast enzyme.

The trapped intermediate is held by a covalent bond to the invariant lysine residue in the P-site, namely Lys$^{263}$ (Figure 2). The P-side of the intermediate is surrounded by the hydrophobic side chains of Tyr$^{216}$, Phe$^{219}$, Tyr$^{237}$ and Val$^{269}$. The P-side carboxy group forms hydrogen bonds with the side chains of Ser$^{279}$ and Tyr$^{128}$ and the P-side amino group is close to the side chains of Ser$^{179}$ and Asp$^{131}$. The side chains of Phe$^{89}$ and Tyr$^{207}$ are also involved in forming this binding site. All of the above residues are invariant or strongly conserved, implying an important role for them in substrate binding to ALADs from all species. The putative intermediate defined in this analysis makes essentially the same interactions on its P-side as a number of inhibitors which occupy the P-site only [10,11,16,27]. The analysis confirms that the intermediate is not bound covalently to the other active-site lysine residue (Lys$^{210}$), even though this has been shown to make a Schiff base in complexes with inhibitors that occupy the A-site [18].

In previous studies of inhibitor and substrate binding at the P-site alone [10,11,16,27], the location of the A-site of ALAD was inferred from the presence of a solvent-filled cavity adjacent to the P-site ligand. This cavity is formed by the zinc ion and a number of invariant residues, including Arg$^{220}$, Arg$^{252}$ and Gln$^{236}$. These predictions are largely confirmed by the currently determined structure in which it was found that the arginine side chains make electrostatic interactions with the A-side carboxy group (Figure 2). This was also found in the bound structure of the inhibitor 4,7-dioxosebacic acid, which occupies the A- and P-sites [17–18]. However, in the structure of the putative intermediate there is an additional interaction between the A-side carboxyl group and the side-chain amino group of Lys$^{210}$ – one of the two invariant catalytic lysine residues. It was speculated that the zinc ion plays an important mechanistic role by interacting with the A-side ALA via a solvent molecule which is bound dative to the metal ion. The requirement for a base in the mechanism to abstract hydrogen atoms from the C-3 of A-side ALA [shown as base B(1) in Scheme 2]. Although it is possible that the C–N bond could form first, thus yielding a Schiff base linking both substrates, this would imply that the Schiff base between A-side ALA and Lys$^{210}$ observed in recent studies [17–19] has no role in catalysis other than perhaps to anchor the A-side substrate prior to catalysis. If instead it is assumed that the Schiff base linking A-side ALA with Lys$^{210}$ has a catalytic function (rather than just a passive binding role), then it is likely that inter-substrate C–C bond formation would occur before the C–N bond formation, as shown in Scheme 2.

In the catalytic reaction, the putative zinc-bound hydroxide is likely to deprotonate the C-3 of A-side ALA [shown as base B(1) in Scheme 2]. The fact that this hydroxide is apparently displaced by the C-5 amino group of A-side ALA, which binds to the zinc instead, may contribute to the stability of the trapped intermediate. Previous NMR studies using isotopically labelled PBG have suggested that the amino group of the reaction product binds to the active-site zinc ion [28]. Two invariant residues, Asp$^{131}$ and Ser$^{179}$, which are likely to catalyse various proton transfers in the reaction, are shown in Figure 2 below the putative intermediate. Two potential catalytic roles for these residues [shown as bases B(2) and B(3) in Scheme 2] could include deprotonation of the P-side C-5 amino group and proton transfer to the amino group of Lys$^{210}$. The fact that the side chains of Lys$^{210}$ and Lys$^{263}$ are adjacent would allow one to transfer a proton to the other during the reaction. For example, the base [labelled B(4) in Scheme 2] which protonates Lys$^{263}$ to facilitate breakdown of the putative intermediate could be Lys$^{210}$.

The apparent stability of the postulated intermediate, under conditions where no particular attempts have been made to trap it, begs the question of how it is stabilized in the crystal structure. Although it is possible that crystal packing contributes in some way to its stabilization, we believe that the intermediate may actually represent a bona fide step in the catalytic cycle that is awaiting the completion of the final stages of the catalytic mechanism (last three stages in Scheme 2). This premise is supported by preliminary observations that enzyme, allowed to turn over in the presence of [4,14C]ALA, results in 14C-labelled enzyme in which two substrate-equivalents are bound per subunit [29]. Addition of unlabelled ALA to the labelled enzyme results in the release of [14C]PBG and generates unlabelled enzyme, presumably with the same (although unlabelled) intermediate at the active site. This is further supported by the observation from MS (N. Mills-Davies, D. Butler and P. M. Shoolingin-Jordan, unpublished work) that the human enzyme forms a stable complex in solution with a molecule of mass close to that of enzyme-bound PBG. In all cases, the enzyme intermediate complex is Ehrlich's reagent-negative [30], indicating that the bound intermediate does not have the same chemical properties as PBG. The X-ray crystallography and biochemical studies are thus mutually reinforcing.

A further interesting observation arising from the X-ray analysis of the enzyme–ligand complex is that the high resolution of
Scheme 2 Proposed mechanism for ALAD

It is assumed that the Schiff base made by Lys210 with A-side ALA facilitates proton abstraction at its C-3 position. This leads to nucleophilic attack on P-side ALA forming the first C–C bond linking the substrates. The inter-substrate C–N bond then forms, yielding the putative intermediate found in the crystal structure (labelled *). Base (1)BH is likely to be the putative zinc-bound hydroxide.

We gratefully acknowledge the ESRF for access to synchrotron beam time and associated travel funds. We thank Professor Muhummad Akhtar, FRS, and Dr M. Montgomery (both of the University of Southampton) for many useful discussions. We thank the Biotechnology and Biological Sciences Research Council (BBSRC) and the Engineering and Physical Sciences Research Council (EPSRC) for financial support.

REFERENCES


Received 4 April 2003; accepted 2 June 2003
Published as BJ Immediate Publication 2 June 2003, DOI 10.1042/BJ20030513

© 2003 Biochemical Society