Articles

X-ray Structure of 5-Aminolevulinic Acid Dehydratase from *Escherichia coli* Complexed with the Inhibitor Levulinic Acid at 2.0 Å Resolution†

Peter T. Erskine,‡ Ed Norton,‡ Jon B. Cooper,*,‡ Richard Lambert,‡ Alun Coker,‡ Gareth Lewis,‡ Paul Spencer,‡,§ Mohammed Sarwar,‡ Steve P. Wood,‡ Martin J. Warren,|| and Peter M. Shoolingin-Jordan‡

Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K., and Department of Molecular Genetics, Institute of Ophthalmology, University College London, Bath Street, London EC1V 9EL, U.K.

Received September 3, 1998; Revised Manuscript Received December 28, 1998

ABSTRACT: 5-Aminolevulinic acid dehydratase (ALAD), an early enzyme of the tetrapyrrole biosynthesis pathway, catalyzes the dimerization of 5-aminolevulinic acid to form the pyrrole, porphobilinogen. ALAD from *Escherichia coli* is shown to form a homo-octameric structure with 422 symmetry in which each subunit adopts the TIM barrel fold with a 30-residue N-terminal arm. Pairs of monomers associate with their arms wrapped around each other. Four of these dimers interact, principally via their arm regions, to form octamers in which each active site is located on the surface. The active site contains two lysine residues (195 and 247), one of which (Lys 247) forms a Schiff base link with the bound substrate analogue, levulinic acid. Of the two substrate binding sites (referred to as A and P), our analysis defines the residues forming the P-site, which is where the first ALA molecule to associate with the enzyme binds. The carboxyl group of the levulinic acid moiety forms hydrogen bonds with the side chains of Ser 273 and Tyr 312. In proximity to the levulinic acid is a zinc binding site formed by three cysteines (Cys 120, 122, and 130) and a solvent molecule. We infer that the second substrate binding site (or A-site) is located between the triple-cysteine zinc site and the bound levulinic acid moiety. Two invariant arginine residues in a loop covering the active site (Arg 205 and Arg 216) appear to be appropriately placed to bind the carboxylate of the A-site substrate. Another metal binding site, close to the active site flap, in which a putative zinc ion is coordinated by a carboxyl and five solvent molecules may account for the activating properties of magnesium ions.

5-Aminolevulinate dehydratase (ALAD,[1] EC 4.2.1.24), also referred to as porphobilinogen synthase, catalyzes one of the initial steps in the biosynthesis of tetrapyrroles involving the condensation of two 5-aminolevulinic acid (ALA) molecules to form the pyrrole porphobilinogen (PBG) (Figure 1). Four porphobilinogen molecules are then condensed in a reaction catalyzed by porphobilinogen deaminase to form the linear tetrapyrrole, preuroporphyrinogen, which is cyclized and rearranged by uroporphyrinogen synthase to give uroporphyrinogen III, the first macrocyclic tetrapyrrole in the pathway (see refs *1*−*4* for reviews). The three steps from ALA to uroporphyrinogen III are common to the biosynthesis of heme, chlorophyll, cobalamins, and all other tetrapyrroles.

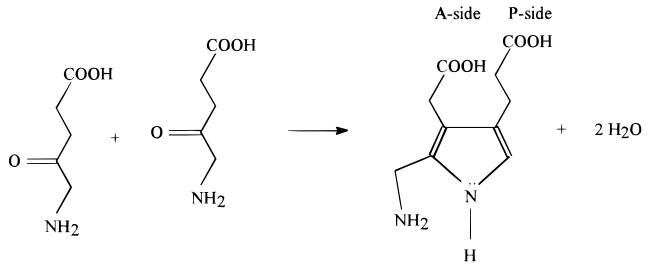

FIGURE 1: Knorr-type condensation reaction catalyzed by 5-aminolevulinic acid deydratase (ALAD) indicating the A- and P-sides of the product porphobilinogen.

ALADs have been purified from a variety of sources including bovine liver (*5*), human erythrocytes (*6*), bacteria such as *Escherichia coli* (*7*), and plants such as spinach (*8*, *9*). The cDNA/gene sequences are available for ALADs from well over a dozen sources. All ALADs share a high degree of sequence identity, contain about 350 amino acids per subunit, and are usually octameric. There are differences between the ALAD enzymes in terms of their metal

† We acknowledge the BBSRC, EPSRC, and Wellcome Trust for financial support.

* Corresponding author. E-mail: jbc2@soton.ac.uk. Fax: (44)-(0)-1703-594459. Tel: (44)-(0)1703-595381.

‡ University of Southampton.

§ Present address: Center for Biological NMR, Department of Chemistry, Texas A&M University, College Station, TX 77843.

|| University College London.

[1] Abbreviations: ALA, 5-aminolevulinic acid; ALAD, 5-aminolevulinic acid dehydratase; PBG, porphobilinogen; rms, root mean square; TIM, triose phosphate isomerase.

10.1021/bi982137w CCC: $18.00
Published on Web 03/17/1999

requirements, kinetic parameters, pH dependence, inactivation by inhibitors, and susceptibility to oxidation. All ALADs require a divalent cation for activity, with animal enzymes using zinc but with plant enzymes requiring magnesium. Representatives of both classes exist in bacteria.

ALAD catalyzes the synthesis of an asymmetric pyrrole from two identical substrate molecules. Single turnover experiments have shown that the first substrate molecule to bind to the enzyme ultimately forms the "propionate" half of the product PBG, while the second substrate molecule forms the "acetate" half of PBG (*10*). This has led to the terminology of the "P" and "A" binding sites in the enzyme which bind the substrates forming the propionate and acetate halves of the product, respectively. Studies on bovine and human ALAD showed that catalysis proceeds by the formation of a Schiff base link at the P-site between the 4-keto group of a substrate and an invariant lysine residue equivalent to Lys 247 in *E. coli* ALAD (*11*). Comparison of ALAD primary sequences (*4, 12*) reveals a strong degree of similarity which is most notable in the vicinity of the active site lysine. Another conserved region of the ALAD sequence has been implicated in metal ion binding.

Most ALADs can bind two zinc ions per subunit, and these sites have been classified variously as A and B or α and β (*7, 13, 14*) where the metal at the α site is essential for catalytic activity and the metal at the β site is thought to have more of a structural role. In *E. coli* ALAD, substitution of magnesium at the β-site and/or the presence of exogenous magnesium leads to a stimulation of activity (*7, 14*). However, in yeast ALAD, although magnesium can bind to the enzyme, it does not lead to any enhancement of enzyme activity (*15*). Zinc ions in ALAD are displaced upon addition of lead, a potent inhibitor of the zinc-dependent enzymes. Zinc is also lost on oxidation of the metal binding thiol groups, which is accompanied by loss of activity (*7, 16, 17*). In contrast, plant dehydratases lack several of the characteristic zinc binding residues and have an absolute requirement for magnesium ions. For example, in the pea, moss, and spinach enzymes two of the conserved cysteines and one histidine are replaced by aspartates, which are more appropriate for coordination of the essential Mg^{2+} ions (*18, 19*). Accordingly, the plant enzymes are less sensitive to oxidation than the mammalian or prokaryotic ALADs. It has been shown in mammalian ALADs that formation of the Schiff base complex between the P-site substrate and the active site lysine does not require Zn^{2+} or enzyme sulfydryl groups (*17, 20*), suggesting that either the zinc ion interacts with the A-site substrate molecule or is important for the structure of the A-site. The evidence for ordered binding has been used to propose several mechanisms. These differ in the precise order of events during the pyrrole ring formation.

The hereditary deficiency of functional dehydratase in humans is associated with the genetic disease Doss or ALAD porphyria (*21*), a disease with severe neurological symptoms due to the accumulation of 5-aminolevulinate which structurally resembles the neurotransmitter GABA, and may have pharmacologically significant properties (*22*). Inhibition of ALAD by lead ions is one of the major manifestations of acute lead poisoning which often leads to neurological and psychotic disturbances. It has been shown that the population has a very wide range of ALAD activity with 2% of individuals having less than 50% of the normal level (*23*).

Although the structure of the yeast enzyme has recently been reported from our laboratories (*24*), this gave little definite indication of the key binding groups at the A- and P-sites apart from the two lysine residues. Furthermore, a large loop extending across the opening of the active site was poorly resolved. This is a highly conserved segment that is likely to be of major importance in the enzyme reaction. In this paper we present the first description of the high-resolution structure of an ALAD complexed with the inhibitor levulinic acid. Our structure provides much information on previously undefined details of the active site, especially residues of the loop (197−220) that appears to dock into the active site cleft on binding of inhibitor and presumably substrate. In the context of earlier work on yeast ALAD (*24*), the structure reported here defines more clearly the residues involved in substrate binding, metal coordination, and catalysis. The structure also represents the first high-resolution description of a prokaryotic ALAD.

EXPERIMENTAL PROCEDURES

Crystallization and Data Collection. The purification and crystallization of wild-type *E. coli* ALAD have been described in detail elsewhere (*7, 25*). Crystals were obtained by the hanging drop method using 2−5% saturated ammonium sulfate as precipitant. The enzyme, at a concentration of 7 mg/mL, was buffered in the pH range 8.1−8.4 using a 0.2 M Tris-HCl solution which included 15 mM levulinic acid, 40 μM zinc sulfate, and 4 mM mercaptoethanol. Crystallization occurred over a 2−4 week period and yielded crystals with typical dimensions of 0.3 mm × 0.3 mm × 0.2 mm. Cryoprotection of the crystals was achieved by rapid mixing of a 10-μL drop of mother liquor (containing the crystal) with 3.5 μL of a 50:50 mixture of glycerol and di-2-propanol. The crystal was then rapidly removed in a mohair loop and cryocooled by dipping it in liquid ethane prior to storage under liquid nitrogen. Crystals treated in this way nearly always diffracted to around 3.0 Å resolution without fading perceptibly during data collection. However, eventually one crystal was obtained which yielded data to 2.0 Å resolution, and 52.5° of data were collected from it at station 9.6 at Daresbury SRS (Warrington, U.K.). This particular crystal was grown in the presence of 1 mM bismuth nitrate in an early effort to make a heavy atom cocrystal. However, no bismuth sites were ever located, and there do not appear to be any in the fully refined structure. The wavelength for this data collection was 0.88 Å, and a 30 cm Marresearch image plate detector was used with an Oxford Cryosystems cryostat to maintain the temperature of the crystal at 100 K. The data were processed using MOSFLM (*26*) and other programs in the CCP4 suite of programs (*27*), and relevant statistics are shown in Table 1.

Molecular Replacement and Refinement. From estimations of the solvent content (*25*) it was initially unclear whether the crystals had one or two subunits per asymmetric unit. A Patterson function was calculated, and this was found to contain no peaks of comparable height to the origin peak, implying that there was only one subunit in the asymmetric unit. Molecular replacement calculations were performed using the programs ALMN and TFFC within the CCP4 suite (*27*). Cross-rotation and translation functions were performed using the yeast ALAD structure as the search model with its N-terminal arm removed since this region of the molecule

Table 1: Data and Refinement Parameters for *E. coli* ALAD[a]

unit cell and symmetry	
space group	*I*422
a, *b* (Å)	126.7
c (Å)	141.6
estimated solvent content (%)	71
subunits per asymmetric unit	1
entire data set	
resolution range (Å)	7.7−2.0
R_{merge} (%)	6.3
completeness (%)	85.8
overall $\langle I/\sigma(I)\rangle$	8.8
fraction $I > 3\sigma(I)$ (%)	80.5
multiplicity	2.5
outer shell	
resolution range (Å)	2.1−2.0
R_{merge} (%)	29.7
completeness (%)	70.8
overall $\langle I/\sigma(I)\rangle$	2.1
fraction $I > 3\sigma(I)$	51.3
multiplicity	1.6
refinement	
R-factor (%)	18.8
free *R*-factor (%)	25.5
correlation coefficient (%)	94.6
resolution range (Å)	7.7−2.0
cutoff $\sigma(F)$	0.0
total no. of reflections	33275
no. of parameters refined	11575
no. of restraints	6651
rms deviation of bond lengths (Å)	0.016
rms deviation of angle distances (Å)	0.032
rms deviation for van der Waals contacts (Å)	0.054
rms deviation of main chain planes (Å)	0.019
rms deviation of side chain planes (Å)	0.012
rms deviation of chiral tetrahedra (Å)	0.018
mean isotropic *B*-factor for 1292 main chain atoms (Å^2)	16.9
mean isotropic *B*-factor for 1194 side chain atoms (Å^2)	23.1
mean isotropic *B*-factor for 374 solvents (Å^2)	40.8

[a] The data were collected at a wavelength of 0.88 Å and were processed using the program MOSFLM (*26*). The structure was refined using X-PLOR (*28*) and RESTRAIN (*29*).

was expected to differ in the two structures due to poor homology. Overall, yeast and *E. coli* ALAD have 38% sequence identity. Inspection of the published alignment (*4*) indicated that there were no extensive insertions or deletions in *E. coli* ALAD relative to the yeast enzyme except for truncation of the N-terminal region. Apart from removal of this arm region (residues 1−39) the yeast ALAD search model was not modified in any other way, leaving the entire TIM barrel domain intact (88% of the molecule). The following searches confirmed that the crystals have only one subunit per asymmetric unit. The cross-rotation function yielded a peak of 3.5 rms when calculated with a 25 Å radius of integration and resolution limits of 20.0−3.0 Å. The noise level in the rotation function was high, but this peak appeared consistently when the rotation function was calculated with different parameters. The translation function yielded a much clearer solution of 13.1 rms with the highest noise peak at 7.7 rms (resolution = 3.0 Å).

Refinement of the solution obtained was then performed using the programs X-PLOR (*28*) and RESTRAIN (*29*) interspersed with graphics rebuilding using O (*30*) and TURBO−FRODO (Bio-Graphics, Marseille). Five percent of the data were reserved for free *R*-factor calculations. The maps were of very high quality and eventually allowed most of the discrepancies in the deposited *E. coli* ALAD sequences

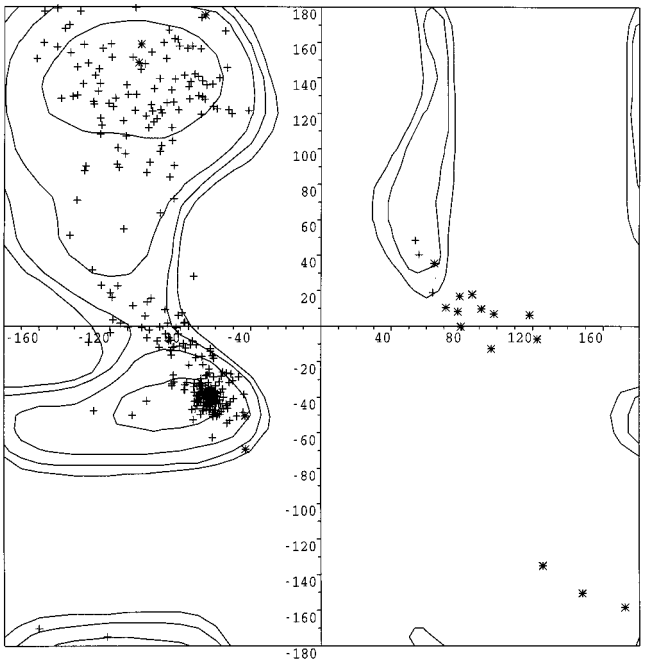

FIGURE 2: Ramachandran plot with polyalanine isoenergy contours. Glycine residues are indicated by star symbols (*) and all other residues as crosses (+).

to be resolved. All side chains were fitted and refined, and there are no regions with significant disorder except for the occasional side chain on the surface of the enzyme. At three places in the sequence, residues which differ from the published sequences of *E. coli* ALAD were incorporated. In the case of Arg 324 (the C-terminal residue) there was no electron density for the side chain presumably due to disorder so it was modeled as an alanine. Likewise with Ile 5 there was little satisfactory electron density for the side chain of this residue, which is in a highly solvent-exposed position. However, the electron density for it could be modeled very well as a serine residue, which indicates that there may be an error in the sequence at this point. In the case of Cys 134 the side chain electron density appeared to be much more consistent with that of a lysine residue. Since a lysine at this position refines well and can make a good side chain hydrogen bond, it was accepted for the final model.

The modeling and refinement lowered the *R*-factor and free *R*-factor to values of 18.8% and 25.5%, respectively, at 2.0 Å resolution. The model has a correlation coefficient (F_o versus F_c) of 94.6% and an estimated coordinate error of 0.28 Å (*31*). Stereochemical and other refinement parameters are given in Table 1, and a Ramachandran plot is shown in Figure 2. By the PROCHECK (*32*) criteria the model has 92.6% of the main chain torsion angles within the "allowed regions" of the Ramachandran plot and an additional 7.4% within the "additional allowed regions". The refined structure contains a number of groups which were identified as sulfate anions. These were initially refined as water molecules, but large features of positive difference density persisted at each site, indicating the presence of more electron dense groups. The use of $(NH_4)_2SO_4$ in the crystallization medium and the general disposition of basic side chains around these sites indicated that they were most probably sulfate anions. The coordinates of the final model and structure factors have been deposited in the Brookhaven Protein Databank (code 1b4e).

RESULTS AND DISCUSSION

The monomers of the ALAD octamer are organized with 422 or D_4 point group symmetry with their active sites oriented toward the solvent region. The octamer has overall dimensions of 104 × 104 × 83 Å^3, consistent with those deduced from EM and solution scattering studies of bovine ALAD (*33*, *34*). Each subunit adopts the $(\alpha/\beta)_8$ or TIM barrel fold with an N-terminal arm approximately 30 residues in length which forms extensive intersubunit interactions. The TIM barrel fold is shared by a number of aldolases, a class of enzymes that are functionally related to ALAD, including the recently solved fructose-1,6-bisphosphate metalloaldolase (*35*, *36*). In ALAD the most extensive of the quaternary contacts are those between pairs of monomers which associate with their arms wrapped around each other to form compact dimers. These dimers further associate to form the octamer which possesses a solvent-filled channel of 15−20 Å diameter passing through the center.

Structure of the Monomer. The first 30 amino acids of the protein form an extended arm-like structure pointing away from the compact α/β domain (Figure 3a). The arm includes a region of distorted 3_{10}-helix (residues 8−12) followed by an α-helix (residues 15−21) denoted α1. The remaining tertiary structure of the ALAD monomer is dominated by the $(\alpha/\beta)_8$ or TIM barrel formed by an eight-membered cylindrical β-sheet surrounded by eight α-helices. In all enzymes adopting the $(\alpha/\beta)_8$ fold, 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 (*37*). In *E. coli* ALAD the TIM barrel is formed from the following elements of secondary structure from the N- to C-terminus: β1, α2, β4, α3, β5, α4, β6, α5, β7, α6, β8, α7, β9, α9, β10, and α10, using the nomenclature adopted for yeast ALAD (*24*). The loop regions between these α and β segments are elaborated extensively at the active site end of the barrel where the regular alternation of helices and strands is broken by the insertion of several extra secondary structure elements. One example of this is the β-hairpin formed by residues 43−57 which lies between β1 and α2. The hydrogen bonds made by the strands of this hairpin (β2 and β3) mean that they are part of the main β-sheet of the molecule and provide an extension to it at the active site end of the barrel. Another elaboration on the basic TIM barrel fold is the loop covering the active site (197−220) which includes a region of α-helix involving residues 202−208 denoted α_{act}. There was very little electron density for this region of the yeast ALAD molecule when crystallized in the absence of inhibitor (*24*), but it is very well defined in the *E. coli* ALAD−levulinic acid complex. Finally, there is also a helix (α8) that lies between β9 and α9 in the primary sequence. This helix is important for quaternary interactions about a 2-fold axis within the dimer and will be discussed later.

The ALADs from yeast and *E. coli* share 38% sequence identity. The longest stretches of conserved sequence in ALADs are all involved in maintaining the structure of the active site which is formed mainly by the loop regions at the exposed end of the β-barrel. Several of these loops make quaternary contacts with neighboring subunits, implying that they may play a role in transmitting conformational change throughout the octamer. The residues in the loop covering the active site (197−220) are also very strongly conserved, indicating that they have an important role in substrate binding during catalysis.

Structure of the Dimer. The N-terminal arm of ALAD (1−30) adopts a conformation in which it partly wraps around the $(\alpha/\beta)_8$ barrel domain of a neighboring monomer to form a dimer (Figure 3b). These interactions occur between monomers related by an intervening 2-fold axis so as to generate dimers resembling the number 69. The $(\alpha/\beta)_8$ domains of each monomer interact extensively about the 2-fold axis so that the two active sites are pointing approximately perpendicular to one another. Our observation of these dimers in the crystal structure is in accord with cross-linking studies of bovine ALAD (*34*), which suggested that the strongest oligomeric interaction is dimer formation.

The N-terminus of the arm region of each monomer is partly exposed on the solvent side of the octamer, but its inner surface interacts with helices α4 and α5 of the second subunit in the dimer. The arm adopts helical conformations between residues 8−12 and 15−21, and these regions interact with helix α6 of the adjacent subunit in the dimer. The arm region in *E. coli* ALAD is quite basic and forms several salt bridges; e.g., Arg 7 interacts with Asp 141 (in α4), Arg 12 interacts with Glu 232 (in α6), and Arg 9 makes extensive interactions with neutral polar groups. There is also an interaction between Arg 18 and Lys 217 of the neighboring monomer mediated by a well-defined sulfate ion which is hydrogen-bonded to both side chains. The latter residue is adjacent to Arg 216, which forms part of the putative binding site for the A-site substrate. This implies that dibasic anions, such as sulfate and phosphate, may have a stabilizing role on the active site structure.

The principal barrel−barrel contacts in the dimer involve the helical segments α7 (252−263), α8 (273−284), and α9 (290−304). Helix α8 is approximately perpendicular to the 2-fold axis of the dimer and pairs in an antiparallel manner with the equivalent helix in the neighboring monomer. Helix α8 is rich in the helix-forming amino acid alanine and appears to be an elaboration of the classical $(\alpha/\beta)_8$ barrel, as mentioned above, perhaps to allow extensive nonpolar intersubunit contacts within the ALAD dimers. The same helix contains an invariant methionine (278) which interacts with the equivalent residue across the subunit interface. Helix α7 interacts with the equivalent helix in the neighboring monomer in a parallel manner as they are roughly aligned with the intervening 2-fold axis. Residues in helix α7 form several intra- and intermolecular electrostatic interactions. For example, the guanidinium of Arg 257 interacts with the side chain of Glu 261, the side chain carboxyl of Glu 258 interacts with arginine 226, and Arg 260 interacts with Asp 307. These interactions are exclusively intramolecular. In contrast, the side chains of Asp 254 and Arg 303 provide a bidentate intersubunit salt bridge. In addition, the guanidinium of Arg 303 forms three hydrogen bonds, two of which are intermolecular, involving nearby main chain carbonyls. The two residues making the above salt bridge (Asp 254 and Arg 303) are invariant, and many of the others mentioned above are completely or highly conserved or replaced with amino acids of the same charge, indicating that these largely buried ionic interactions may be important for conserving the quaternary structures of ALADs. The helix α9 also contains several invariant residues including Arg 303 (men-

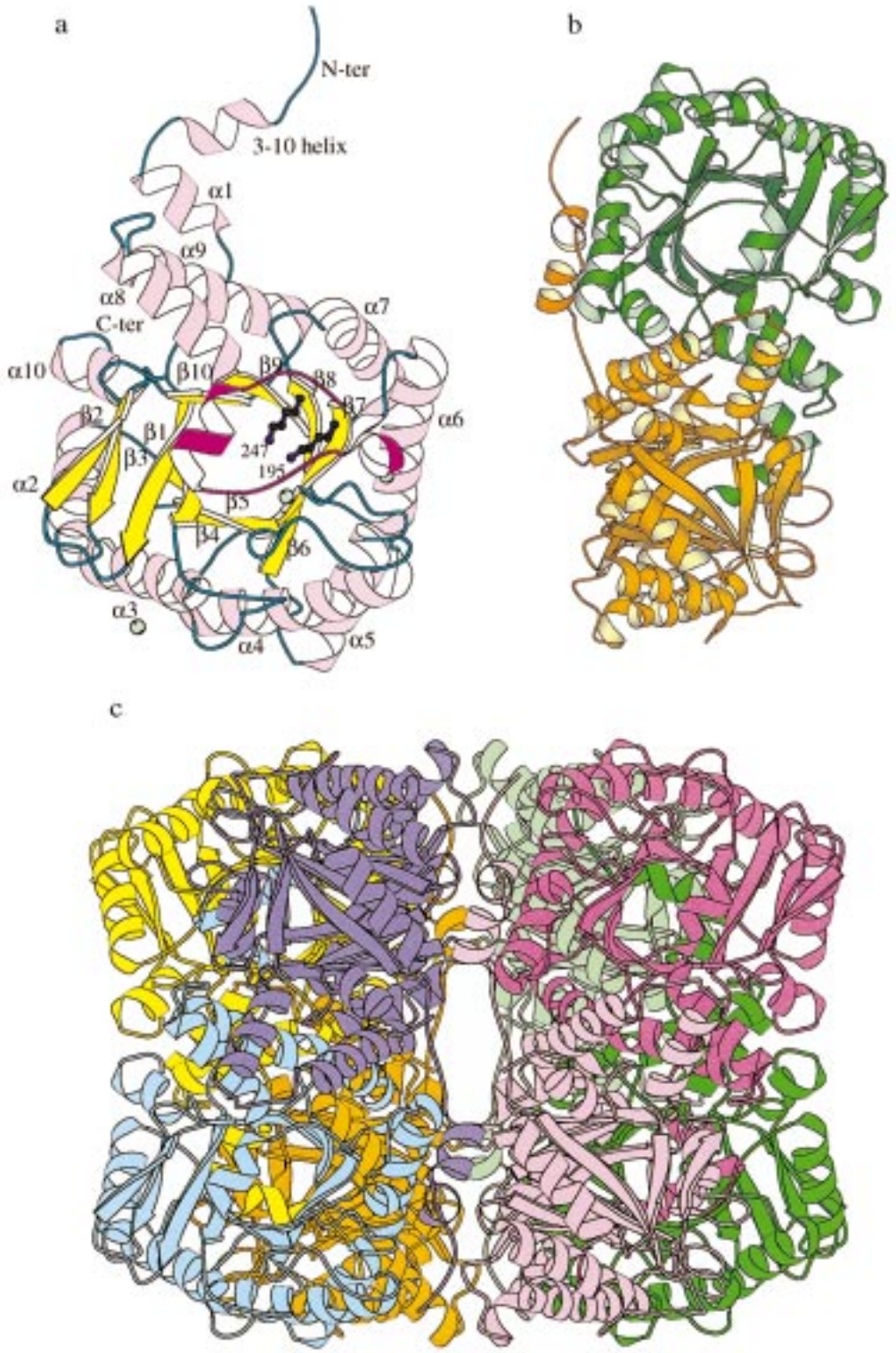

FIGURE 3: (a) Ribbon diagram of the fold of the ALAD monomer. The dominant feature of the tertiary structure is the closed $(\alpha/\beta)_8$ or TIM barrel with the active site located in a pronounced cavity (facing reader). Helical segments are shown in lilac, strands in yellow, and the connecting loop regions in cyan. The active site flap (197−220) is colored magenta, and the enzyme's N-terminal arm region (residues 1−30) which forms extensive quaternary contacts is shown above the barrel. The two adjacent lysines implicated in catalysis (195 and 247) are shown in blue. Lys 247 forms a Schiff base link with P-site ALA. Green spheres indicate the locations of the zinc binding sites. The active site zinc ion held by three cysteines can be seen close to the two lysines in the center of the TIM barrel. The zinc ion on the right-hand side is involved in intersubunit contacts, and that on the lower left is involved in crystal contacts. (b) Formation of *E. coli* ALAD dimers. The two monomers (colored differently) associate extensively about an intervening 2-fold axis, and their arm regions wrap around the TIM barrel domain of the neighboring subunit. (c) Organization of the *E. coli* ALAD octamer. The dimers (colored differently) are oriented with their long axes slightly inclined with respect to the central 4-fold axis of the octamer (vertical). Interactions between adjacent dimers are mediated principally by the arm regions. For clarity, the subunits within each dimer have been given slightly different colors. This figure and others like it were prepared using the program MOLSCRIPT (*44*) or the BOBSCRIPT variant (*45*).

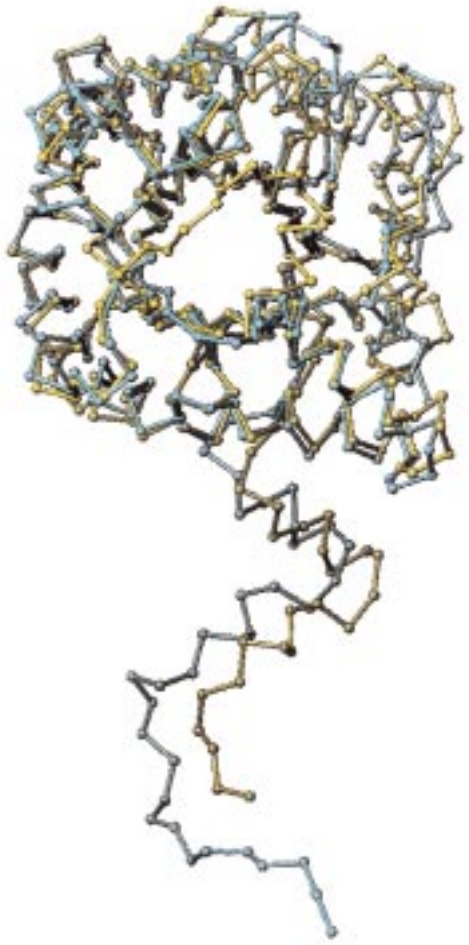

FIGURE 4: Least-squares superposition of *E. coli* and yeast ALAD shown in yellow and cyan lines, respectively. The TIM barrel domains of both structures superimpose well, but the N-terminal arm regions differ in length and orientation. The two structures superimpose with an rms deviation of 0.8 Å for 210 structurally equivalent Cα atoms closer than 1.5 Å.

tioned above due to its interaction with Asp 254) and Glu 296 which hydrogen-bonds to the hydroxyl of Tyr 252 (invariant) and two well-conserved serines (198 and 199) across the monomer−monomer interface.

Structure of the Octamer. The dimers in the ALAD octamer have the gross appearance of ellipsoids with their long axes inclined slightly with respect to the 4-fold symmetry axis. The overall appearance of the octamer is shown in Figure 3c, where the dimers are colored separately. All dimers within the octamer interact with each other in an identical manner, and these interactions are mediated principally by the arm regions that are on the surface of each dimer. In the dimer−dimer contacts, the two helical segments in the arm region (residues 8−12 and 15−21) associate with one end of a β-barrel in the neighboring dimer. This end of the β-barrel is essentially "capped" by the arm region of the neighboring subunit leaving the other end of the barrel, which forms the active site, exposed on the surface of the octamer. In this way the arm of *E. coli* ALAD forms two salt bridges with the neighboring subunit involving the following interactions: Arg 10−Asp 242 and Lys 13−Asp 186.

The barrel-capping interactions also involve a sulfate anion which lies between Lys 13 of the N-terminal arm and Lys 107 at the bottom of the neighboring TIM barrel. Another sulfate anion, linked by hydrogen bonds to the guanidinium of Arg 257, also participates in the interactions between adjacent dimers. This sulfate lies near a 2-fold axis in a channel between four monomers and hydrogen-bonds to its symmetry mate in the channel. This remarkable channel is formed by eight conserved arginine side chains which originate from residues 226, 227, 257, and 260 and their symmetry equivalents. The strong conservation of basic residues at these positions suggests that interaction with dibasic anions, such as phosphate, may be an important feature which stabilizes the octamer in vivo. The apparent role of dibasic anions in the structure may partly explain the effect which phosphate buffer has on prolonging the activity of the *E. coli* enzyme during storage (*7*). In the dimer−dimer contacts, the arm regions of neighboring dimers themselves associate briefly by van der Waals contacts in the vicinity of residue 25 close to the intervening 2-fold axis (which is coincident with the crystallographic a and b axes). In all of these side-by-side dimer−dimer contacts there are few direct interactions between the globular domains of adjacent dimers. However, the few residues which do form such direct interactions include the following ion pairs: Glu 264−Arg 262 and Arg 227−Asp 307. These residues are not particularly conserved, indicating that the most important dimer−dimer interactions are probably mediated by the arm regions.

Variability of the Arm Region. The N-terminal arm of *E. coli* ALAD is 11 residues shorter than that of the yeast enzyme and adopts a slightly different orientation (Figure 4). Nonetheless, this region of the monomer has the same function in both ALADs by forming extensive quaternary contacts. Despite the apparent importance of the arm region in the quaternary interactions, the arm sequences are poorly conserved and vary significantly in length across the family of ALADs. Accordingly, the arms of yeast and *E. coli* ALAD differ significantly in sequence and, to some extent, structure. However, residues of the first α-helix (α1) are present in all dehydratases, and nearly all of them possess residues corresponding to the preceding 3_{10}-helical segment. This implies that the general quaternary interactions made by the arm region in yeast and *E. coli* ALAD probably constitute a conserved feature of this enzyme family. The length of the arm region of *E. coli* ALAD is more typical of the other ALADs than the yeast enzyme.

The Active Site. The loops at the exposed end of the β-barrel form a pronounced cavity, the base of which is dominated by two spatially adjacent lysine side chains at positions 195 and 247 (Figure 5). Both of these residues occur at the C-terminal ends of adjacent β-strands in the barrel, namely, β7 and β8. One of the lysines (247) is followed by a proline residue which is invariant and adopts the *cis* conformation in both yeast and *E. coli* ALAD. The side chains of these two lysines emerge from a hydrophobic pocket formed by the side chains of tyrosines 192, 201, and 270 and other highly conserved residues such as Val 272 and Gln 220. The active site also contains a number of invariant polar residues such as Glu 123, Ser 273, Asp 118, and Ser 165 whose putative roles in substrate binding and catalysis are discussed later. Lys 247, which is known to form a Schiff base to the P-site substrate molecule, has a more hydrophobic environment than its neighbor (Lys 195).

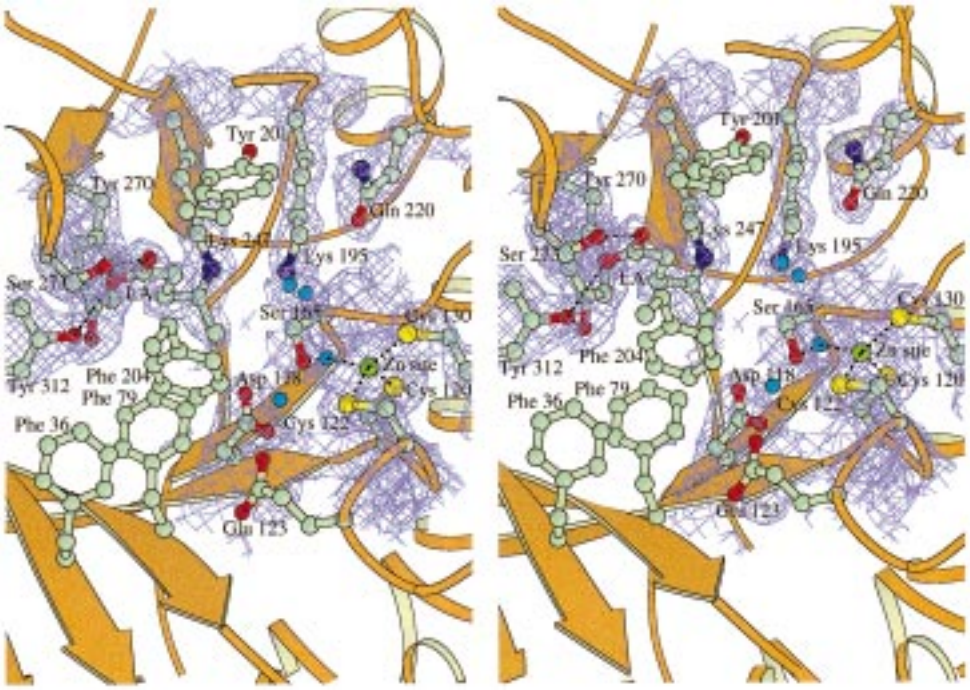

FIGURE 5: Stereoview of the active site showing the two lysines implicated in catalysis (247 and 195) in the top center. Lysine 247 has formed a Schiff base link to the inhibitor levulinic acid (LA), which is hydrogen bonded to Ser 273 and Tyr 312 (center left). The zinc binding site consists of cysteines 120, 122, and 130 and a solvent molecule which coordinate the metal ion (shown on the right). The $2F_o - F_c$ electron density at 2.0 Å resolution (contoured at 1.2σ) is shown for selected residues as purple lines, and solvent molecules are colored blue. Phe 204 (shown below the levulinic acid moiety) is notable since it was disordered in the native yeast ALAD structure; i.e., its interaction with the inhibitor may contribute to ordering of the active site flap (197−220). The feature of electron density containing two solvent molecules just above Ser 165 and in front of Lys 195 may originate from a disordered levulinic acid residue at the A-site.

The substrate binding cleft in ALAD has a number of aromatic residues which are predominantly tyrosines whose hydroxyl groups point toward the active site. These aromatic residues include Phe 36, Phe 79, Tyr 192, Tyr 201, Phe 204, Tyr 270, and Tyr 312 which are almost totally invariant. The two phenylalanine side chains at positions 36 and 79 appear to form a hydrophobic patch against which nonpolar residues in the helical segment (α_{act}) of the loop covering the active site are packed. The hydroxyl groups of tyrosines 270 and 312 are in hydrogen-bonding contact and are close to the side chain $-NH_2$ group of the Schiff base lysine (247). As will be discussed later, these tyrosine residues are likely to have an important role in substrate binding since they interact with the inhibitor, levulinic acid, covalently bound to Lys 247. The $-NH_2$ group of the other lysine, Lys 195, is approximately 5 Å from a well-defined zinc ion which is held by three cysteine side chains and a solvent molecule.

The Active Site Zinc Ion. Cysteine residues 120, 122, and 130 are located in the loop connecting β5 and α4 (Figure 5). This region of the molecule was previously identified as a potential metal binding region from primary sequence comparisons, since it possesses a number of likely zinc binding cysteine residues which are absent in the magnesium binding plant ALADs (*18, 19*). Two of the above three cysteine residues (122 and 130) are replaced by aspartate in plant ALADs, and Cys 120 becomes Ala. The three cysteine side chains in *E. coli* ALAD along with a water molecule coordinate the zinc ion with tetrahedral geometry. There is good electron density for this solvent ligand which lies between the metal ion and the two active site lysines. The zinc ion is close to the side chain of the invariant Ser 165 which itself hydrogen-bonds to the carboxyl group of an invariant aspartate (Asp 118). These two residues interact indirectly with the zinc ion through a network of hydrogen-bonded water molecules. The proximity of these interacting groups to the Schiff base lysine indicates that they most probably participate in substrate binding or catalysis. It is very unusual to find an active site zinc ion coordinated by three cysteine ligands since zinc shows a preponderance of imidazole and carboxyl ligands at the active sites of other enzymes such as the functionally related metalloaldolases. However, the active site of alcohol dehydrogenase is similar in that its zinc ion is held by two cysteines and a histidine. In this enzyme the zinc is presumed to coordinate the carbonyl of ethanal (acetaldehyde) to facilitate its reduction. In ALAD the active site zinc ion may perform a similar task by coordinating the carbonyl of the A-side substrate prior to its condensation with the P-side ALA.

The Second Zinc Site. Adjacent to the side chain of Lys 195 is a water-filled pocket which lies between the TIM barrel domain and the N-terminal arm of the neighboring subunit in the dimer. This pocket contains a zinc ion which is coordinated octahedrally by one side chain oxygen of Glu 232 (in helix α6) and five solvent molecules, presumed to be waters (Figure 6). While the nature of the groups forming this metal binding site cannot be determined unambiguously by X-ray analysis, all atoms at this site in the current model have refined with reasonable thermal parameters. The putative water molecules bound to the zinc ion are hydrogen-bonded to surrounding side chain atoms of Asp 169, Gln 171, and Asp 236. There have been many studies of the metal ion requirements of ALAD, and there is experimental evidence for the enzyme having at least two distinct metal binding sites which contribute to its catalytic function (*13*,

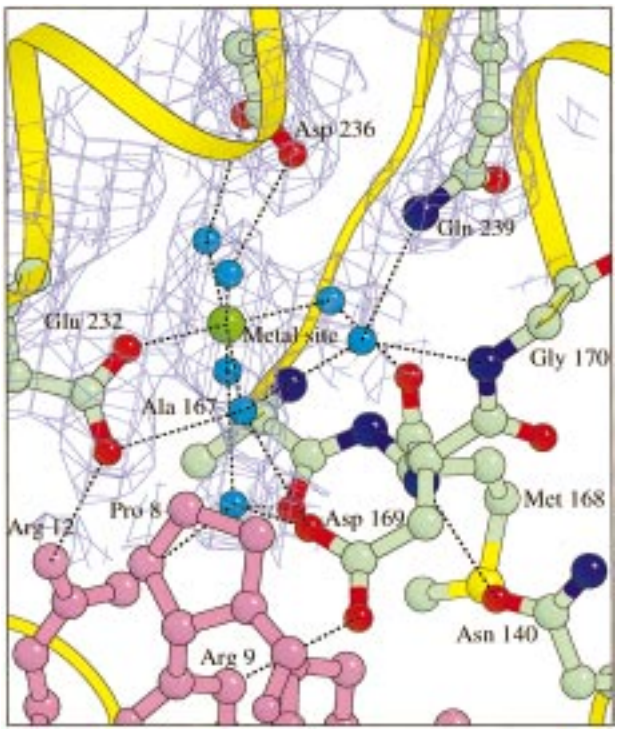

FIGURE 6: Pentahydrated zinc binding site at Glu 232. This glutamate side chain and five water molecules coordinate a zinc ion with approximate octahedral geometry. This site is at the subunit interface in the octamer. The lilac arginine and proline side chains (bottom left) originate from the N-terminal arm region of the adjacent monomer. The $2F_o - F_c$ electron density at 2.0 Å resolution (contoured at 1.2σ) is shown for selected residues as purple lines.

14). While the closest side chain atom of Lys 195 is less than 8 Å from this zinc ion, the lysine ϵ-NH_2 group is over 11 Å from the metal ion. Hence, it is difficult to envisage a direct catalytic role for this metal site. Nonetheless, it would be expected to contribute significantly to the stability of the octamer since the arm residue Arg 12 (from the neighboring subunit) forms a salt bridge with the residue ligating the metal ion (Glu 232). This would appear to be the most likely site at which Mg^{2+} could bind to ALAD. The apparent activating properties of magnesium may therefore stem from an effect on the active site flap which passes close to this metal ion or an effect on the enzyme structure close to the active site. In *E. coli* ALAD it has been shown that magnesium activates the enzyme even at optimal zinc concentrations (*7*). However, this effect is not observed with yeast ALAD, and in the latter enzyme this zinc/magnesium binding site is disrupted by an arginine side chain (251) which forms a five-membered salt-bridge interaction involving Asp 180, Arg 185, Arg 251, Asp 252, and Asp 183. In *E. coli* ALAD several of these residues are replaced by smaller side chains, thereby creating space for the second metal binding site. The absence of this site in yeast ALAD may account for the failure of magnesium ions to activate ALAD from this species at optimal zinc ion concentrations although other factors may contribute to this intriguing difference between the yeast and *E. coli* enzymes.

In the *E. coli* ALAD crystal structure, we have identified a third zinc binding site involving residues Glu 40 and His 84 from two neighboring subunits. The same two residues in each subunit coordinate the metal ion (which lies on a crystallographic 2-fold axis) with tetrahedral geometry. Both of these residues are in poorly conserved regions of the structure which differ greatly from the yeast homologue. Since this site resides between octamers in the crystal lattice, it is unlikely to be of functional significance although some ALADs are reported to form higher oligomeric states at elevated metal ion concentrations (*1*).

Levulinic Acid Binding. In the structure reported here, the bound P-site levulinic acid makes a Schiff base to Lys 247 and is extremely well defined by the electron density map (Figure 5). The carboxyl group of the levulinic acid forms three hydrogen bonds, one with the side chain of Tyr 312 and two with Ser 273 involving its hydroxyl and main chain nitrogen. These two residues are invariant in ALADs as are most of the residues which interact with the hydrophobic moiety of the levulinic acid. These include Phe 79, Tyr 201, Tyr 270, Val 272, and Phe 204. Phe 204, which packs very closely against the levulinic acid methylene groups and shields it from solvent, belongs to the loop covering the active site. This loop is completely invisible in the native yeast ALAD structure, presumably due to its flexibility in the absence of substrate or inhibitor. Hence the hydrophobic interaction between P-site levulinic acid or ALA and Phe 204 may be crucially important for ordering the remaining loop residues so that they can function in binding the second substrate.

Possible interactions made by the 5-amino group of the P-side substrate (ALA) can be analyzed by modeling an amino group onto the C-5 atom of the bound levulinic acid with the appropriate bond length and bond angle. Torsional rotation of the amino group allows all possible interactions to be explored, assuming that the remainder of the ALA molecule binds with interactions similar to those of levulinic acid. In one conformation the amino group makes reasonable hydrogen bonds with the $-NH_2$ of Lys 195, the −OH of Ser 165, the carboxyl of Asp 118, and the −OH of Tyr 192. All of these are invariant residues in ALADs, which suggests that at least a subset of them are important for binding the amino group of the P-site ALA.

The Location of the A-Site. The interactions that an ALA molecule could make at the A-site of the enzyme can be predicted with reasonable confidence from the levulinic acid complex reported here. Adjacent to the P-site levulinic acid is a solvent-filled pocket lined by the following highly conserved residues: Asp 118, Ser 165, Lys 195, Arg 205, Arg 216, and Gln 220 and the zinc−sulfur cluster involving cysteines 120, 122, and 130. Further evidence that this may be the A-site of the enzyme is provided by the presence of an unusual elongated feature of electron density which might represent an additional levulinic acid molecule bound with low occupancy (Figure 5).

Implications for the Catalytic Mechanism. The synthesis of PBG requires the formation of a C−N bond and a C−C bond with the loss of two water molecules. Single turnover experiments established that the substrate moiety to bind first forms a Schiff base link with the enzyme at the P-site (*10*). It has been suggested that the enzyme may function by forming an additional Schiff base linking the A-site substrate to the enzyme (*38*). This was proposed prior to determination of the X-ray structure, which showed that the side chain of the Schiff base lysine (247) is spatially adjacent to the side chain of another invariant lysine (195). Hence the possibility of a mechanism in which both substrate moieties are held to

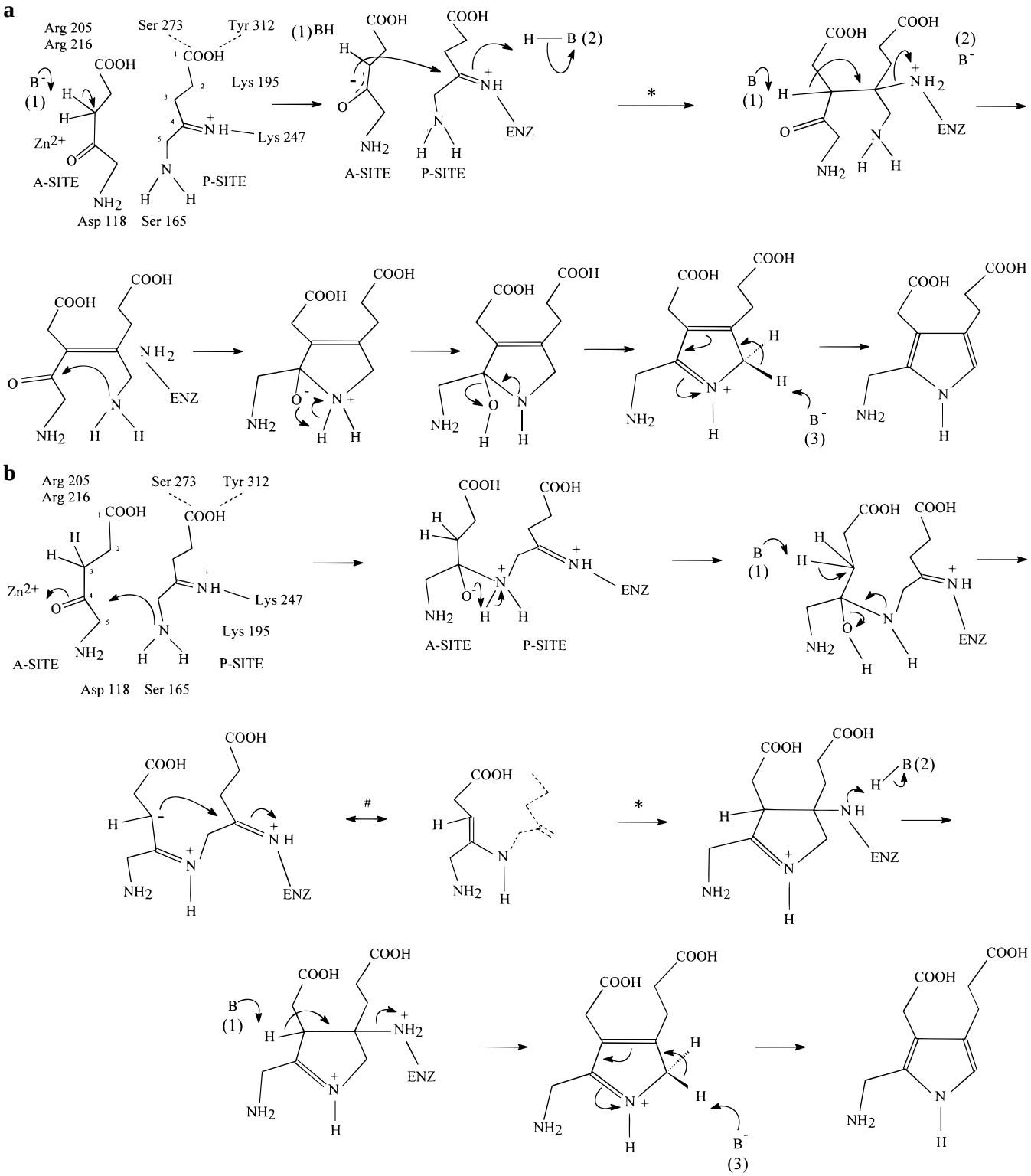

FIGURE 7: Details of the two putative ALAD catalytic mechanisms in which the order of the C−C and C−N bond forming reactions differs. In both schemes the first intermediate has one 5-aminolevulinic acid (ALA) moiety bound to the enzyme's P-site by a Schiff base link to the essential lysine residue (247 in *E. coli* ALAD). A second ALA molecule occupies the A-site where it can interact with the cysteine-bound zinc ion. Ultimately a Schiff base link is formed between the substrates, yielding one of the C−N bonds of the final pyrrole product after several deprotonations. Enzyme residues which form the substrate binding sites are indicated for the first intermediate of each scheme. The * symbols indicate the aldol condensation steps (C−C bond formation), and the # symbol in (b) indicates resonance with an enamine canonical form which stabilizes the carbanion at C-3.

the enzyme by Schiff bases cannot be ruled out totally although there is no direct experimental evidence for this.

In addition to the Schiff base formed between the enzyme and the P-site substrate, another Schiff base is formed between the two substrate moieties. This involves the C-4 keto group of the A-site substrate and the amino group of the P-site substrate, eventually forming the C−N bond linking both ALA molecules in the product. The C−C bond formed between the substrates results from nucleophilic attack by a stabilized carbanion at the C-3 position of the

A-site substrate, on the carbonyl carbon of the P-site substrate. The latter would be rendered highly electropositive by its Schiff base link to the enzyme.

The proximity of the two lysines at the active site is reminiscent of aspartic proteinases which have two aspartic acid side chains in close proximity at the catalytic center (*39*, *40*). These two aspartates are presumed to share a single negative charge in the optimal pH range. In contrast to aspartic proteinases where the two active site residues have a very symmetric environment in the absence of substrate, the two lysines of ALAD are in rather different environments. Prior to its condensation with the substrate, the ϵ-NH_2 group of Lys 247 would have to be neutral for it to act as a nucleophile. Lysine 195 is more likely to be positively charged since its environment is more polar. There are parallels with the enzyme acetoacetate decarboxylase which possesses two sequential lysine residues, one of which forms a Schiff base link with substrate. The other lysine is thought to depress the pK_a of the adjacent Schiff base lysine (*41*, *42*). Thus a positive charge on Lys 195 of ALAD would serve to lower the pK_a of Lys 247 and make it a more effective nucleophile for Schiff base formation. In addition, it could act similarly on the amino group of the P-side ALA to depress its pK_a and facilitate nucleophilic attack on the C-4 carbonyl of the A-side substrate leading to the formation of an initial C−N bond between the two substrate molecules.

In order for the A-site ALA to condense in the correct orientation with respect to the P-site substrate, it would have to bind with its 5-amino group close to the Asp 118−Ser 165 interacting pair. Its C-1 carboxyl group would also be oriented toward arginines 205 and 216, possibly making a salt bridge with one of them. The central region of the A-site ALA would then be in a position to interact with the zinc ion and its associated water molecule as well as the amino group of Lys 195. These groups could facilitate the deprotonation reactions and assist stabilization of the C-3 carbanion prior to aldol condensation. Thus these groups are thought to be essential for catalyzing the C−N and C−C bond formations which eventually link the two substrates and yield porphobilinogen. The C−C bond results from an aldol condensation, and the C−N bond results from formation of an intersubstrate Schiff base. Two reaction schemes have been deduced from chemical considerations and differ mainly in the order with which the C−C and C−N bonds are formed. These are shown in Figure 7. The scheme in which C−N bond formation occurs first (Figure 7b) may be more attractive from the mechanistic point of view since formation of the intersubstrate Schiff base increases the acidity of the A-site C-3 protons.

Both reaction schemes involve basic groups which perform the following roles: (1) abstraction of protons from the C-3 carbon of the A-site ALA to facilitate C−C bond formation between the substrates, (2) proton transfer to and from the nitrogen of the Schiff base lysine, and (3) deprotonation of the C-5 carbon of the P-site ALA in a process which has been shown to be stereospecific for the *pro-R* hydrogen (*43*). These basic groups are numbered accordingly in Figure 7. While it is not possible to discriminate with certainty between the two possible mechanisms shown, it is now possible to speculate, with reasonable confidence, on the nature of the basic groups which catalyze these reactions, disregarding the order of the events involved. We propose that the base which catalyzes the deprotonation of the P-side ALA at C-5 (base 3) is the amino group of Lys 247. The group which could facilitate proton transfer to and from the Schiff base (base 2) is most likely to be the amino group of Lys 195, and the strongly basic group required for deprotonation of the A-side C-3 (base 1) could be a zinc-bound hydroxide ion. The solvent molecule datively bound to the zinc ion is the most likely candidate for a hydroxyl species in the vicinity of the substrate. The zinc ion could also coordinate the carbonyl oxygen of the A-site substrate, thereby stabilizing the putative oxyanion intermediates in the reaction mechanism.

ACKNOWLEDGMENT

We thank Professor M. Akhtar FRS for his comments on the manuscript and congratulate him on the occasion of his retirement.

REFERENCES

1. Jordan, P. M. (1991) Biosynthesis of tetrapyrroles, *New Compr. Biochem. 19*, 1−65.
2. Jordan, P. M. (1994) Highlights in haem biosynthesis. *Curr. Opin. Struct. Biol. 4*, 902−911.
3. Warren, M. J., and Scott, A. I. (1990) Tetrapyrrole assembly and modification into the ligands of biologically functional cofactors, *Trends Biochem. Sci. 15*, 486-491.
4. Jaffe, E. K. (1995) Porphobilinogen synthase, the first source of heme's asymmetry, *J. Bioenerg. Biomemb. 27*, 169−179.
5. Gibson, K., Neuberger, A., and Scott, J. J. (1955) The purification and properties of δ-aminolevulinic acid dehydratase, *Biochem. J. 61*, 618−629.
6. Anderson, P. M., and Desnick, R. J. (1979) δ-aminolevulinate dehydrase from human erythrocytes, *J. Biol. Chem. 254*, 6924−6930.
7. Spencer, P., and Jordan, P. (1993) Purification and characterisation of 5-aminolevulinic acid dehydratase from *E. coli* and a study of reactive thiols at the metal-binding domain, *Biochem. J. 290*, 279−287.
8. Schneider, H. A. W., and Liedgens, W. (1981) An evolutionary tree based on monoclonal antibody recognized surface-features of a plastid enzyme (5-aminolevulinate dehydratase), *Z. Naturforsch. 36c*, 44−50.
9. Liedgens, W., Lutz, C., and Schneider, H. A. W. (1983) Molecular-properties of 5-aminolevulinic acid dehydratase from *Spinacia-oleracea*, *Eur. J. Biochem. 135*, 75−79.
10. Jordan, P. M., and Gibbs, P. N. B. (1985) Mechanism of action of 5-aminolevulinate dehydratase from human erythrocytes, *Biochem. J. 227*, 1015−1020.
11. Gibbs, P. N. B., and Jordan, P. M. (1986) Identification of a lysine at the active site of human 5-aminolevulinate dehydratase, *Biochem. J. 236*, 447−451.
12. Li, J. M., Russell, C. S., and Cosloy, S. D. (1989) The structure of the *E. coli Hem B* gene, *Gene 75*, 177−184.
13. Dent, A., Beyersmann, D., Block, C., and Hasnain, S. S. (1990) Two different zinc sites in bovine 5-aminolevulinate dehydratase distinguished by extended X-ray absorption fine structure, *Biochemistry 29*, 7822−7828.
14. Mitchell, L. W., and Jaffe, E. K. (1993) Porphobilinogen synthase from *Escherichia coli* is a Zn(II) metalloenzyme stimulated by Mg(II), *Arch. Biochem. Biophys. 300*, 169−177.
15. Senior N., Thomas, P. G., Cooper, J. B., Wood, S. P., Erskine, P. T., Shoolingin-Jordan, P. M., Warren, M. J. (1996) Comparative studies of the 5-aminolevulinic acid dehydratase from *P. sativum*, *E. coli* and *S. cerevisiae*, *Biochem. J. 320*, 401−412.
16. Tsukamoto, I., Yoshinaga, T., and Sano, S. (1979) The role of zinc with special reference to the essential thiol groups in δ-aminolevulinic acid dehydratase of bovine liver, *Biochim. Biophys. Acta 570*, 167−178.

17. Jaffe, E. K., and Hanes, D. (1986) Dissection of the early steps in the porphobilinogen synthase catalysed reaction. Requirements for Schiff's base formation, *J. Biol. Chem. 261*, 9348−9353.
18. Boese, Q. F., Spano, A. J., Li, J., and Timko, M. P. (1991) 5-Aminolevulinic acid dehydratase in pea. Identification of an unusual metal-binding domain in the plant enzyme, *J. Biol. Chem. 266*, 17060−17066.
19. Cheung, K.-M., Spencer, P., Timko, M. P., and Shoolingin-Jordan, P. M. (1997) Characterisation of a recombinant pea 5-aminolevulinic acid dehydratase and comparative inhibition studies with *Escherichia coli* dehydratase, *Biochemistry 36*, 1148−1156.
20. Jordan, P. M., Gore, M. G., and Chaudhry, A. G. (1976) Subunit modification of 5-aminolevulinate dehydratase involving cysteine residues, *Biochem. Soc. Trans. 4*, 762−763.
21. Doss, M., Von-Tieperman, R., Schneider, J., and Schmid, H. (1979) New types of hepatic porphyria with porphobilinogen synthase defect and intermittent acute clinical manifestation, *Klin. Wochenschr. 57*, 1123−1127.
22. Brennan, M. J. W., and Cantrill, R. C. (1979) δ-Aminolevulinic acid is a potent agonist for GABA autoreceptors, *Nature 280*, 514−515.
23. Thunell, S., Holmberg, L., and Lundgren J. (1987) Aminolevulinate dehydratase porphyria in infancy—a clinical and biochemical study, *J. Clin. Chem. Clin. Biochem. 25*, 5−14.
24. Erskine, P. T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, I. J., Sarwar, M., Spencer, P., Thomas, P., Warren, M., Shoolingin-Jordan, P. M., Wood, S. P., and Cooper, J. B. (1997) X-ray structure of 5-aminolevulinate dehydratase, a hybrid aldolase, *Nat. Struct. Biol. 4*, 1025−1031.
25. Erskine, P. T., Senior, N., Maignan, S., Cooper, J., Lambert, R., Lewis, G., Spencer, P., Awan, S., Warren, M., Tickle, I. J., Thomas, P., Wood, S. P., and Shoolingin-Jordan, P. M. (1997) Crystallisation of 5-aminolevulinic acid dehydratase from *Escherichia coli* and *Saccharomyces cerevisiae* and preliminary X-ray characterisation of the crystals, *Protein Sci. 6*, 1774−1776.
26. Leslie, A. G. W. (1991) Molecular data processing, in *Crystallographic computing* (Moras, D., Podjarny, A. D., and Thierry, J. C., Eds.) pp 50−61, IUC Oxford University Press, Oxford.
27. CCP4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
28. Brunger, A. T. (1987) *X-PLOR: A system for X-ray crystallography and NMR*, Yale University Press, New Haven and London.
29. Haneef, I., Moss, D. S., Stanford, M. J., and Borkakoti, N. (1985) Restrained structure-factor least-squares refinement of protein structures using a vector-processing computer, *Acta Crystallogr. A41*, 426−433.
30. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
31. Read, R. J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors, *Acta Crystallogr. A42*, 140−149.
32. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
33. Wu, W., Shemin, D., Richards, K. E., and Williams, R. C. (1974) The quaternary structure of δ-aminolevulinic acid dehydratase from bovine liver, *Proc. Natl. Acad. Sci. U.S.A. 71*, 1767−1770.
34. Pilz, I., Schwarz, E., Vuga, M., and Beyersmann, D. (1988) Small-angle X-ray scattering study of bovine porphobilinogen synthase, *Biol. Chem. Hoppe-Seyler 369*, 1099−1103.
35. Blom, N. S., Tetreault, S., Coulombe, R., and Sysguch, J. (1996) Novel active site in *E. coli* fructose-1,6-bisphosphate aldolase, *Nat. Struct. Biol. 3*, 856−862.
36. Cooper, S. J., Leonard, G. A., McSweeney, S. M., Thompson, A. W., Naismith, J. H., Qamar, S., Plater, A., Berry, A., and Hunter, W. (1996) The crystal structure of a class II fructose-1,6-bisphosphate aldolase shows a novel metal-binding active site embedded in a familiar fold, *Structure 4*, 1303−1315.
37. Branden, C., and Tooze, J. (1991) *Introduction to protein structure*, Garland, New York.
38. Neier, R. (1996) Chemical synthesis of porphobilinogen and studies of its biosynthesis, *Adv. Nitrogen Heterocycles 2*, 35−146.
39. Pearl, L. H., and Blundell, T. L. (1984) The active site of aspartic proteinases, *FEBS Lett. 174*, 96−101.
40. Veerapandian, B., Cooper, J. B. Sali, A., Blundell, T. L., Rosati, R. L., Dominy, B. W., Damon, D. B., and Hoover, D. J. (1992) Direct observation of the tetrahedral "intermediate" of aspartic proteinases, *Protein Sci. 1*, 322−328.
41. Highbarger, L. A., Gerlt, J. A., and Kenyon, G. L. (1996) Mechanism of the reaction catalyzed by acetoacetate decarboxylase—importance of lysine-116 in determining the pK_a of active-site lysine-115, *Biochemistry 35*, 41−46.
42. Schmidt, D. E., Jr., and Westheimer, F. H. (1971) p*K* of the lysine group at the active site of acetoacetate decarboxylase, *Biochemistry 10*, 1249−1253.
43. Chaudhry, A. G., and Jordan, P. M. (1976) Stereochemical studies on the formation of porphobilinogen, *Biochem. Soc. Trans. 4*, 760−761.
44. Kraulis, P. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
45. Esnouf, R. (1997) An extensively enhanced version of MolScript that includes greatly enhanced coloring capabilities, *J. Mol. Graphics 15*, 132.

BI982137W