The EMBO Journal (2005) 24, 3166–3177 | © 2005 European Molecular Biology Organization | All Rights Reserved 0261-4189/05
www.embojournal.org

THE EMBO JOURNAL

Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans

Isabel Astner[1,3], Jörg O Schulze[1,3], Joop van den Heuvel[1], Dieter Jahn[2], Wolf-Dieter Schubert[1,*] and Dirk W Heinz[1]

[1]Division of Structural Biology, German Research Centre for Biotechnology, Braunschweig, Germany and [2]Institute of Microbiology, Technical University Braunschweig, Braunschweig, Germany

5-Aminolevulinate synthase (ALAS) is the first and rate-limiting enzyme of heme biosynthesis in humans, animals, other non-plant eukaryotes, and α-proteobacteria. It catalyzes the synthesis of 5-aminolevulinic acid, the first common precursor of all tetrapyrroles, from glycine and succinyl-coenzyme A (sCoA) in a pyridoxal 5′-phosphate (PLP)-dependent manner. X-linked sideroblastic anemias (XLSAs), a group of severe disorders in humans characterized by inadequate formation of heme in erythroblast mitochondria, are caused by mutations in the gene for erythroid eALAS, one of two human genes for ALAS. We present the first crystal structure of homodimeric ALAS from *Rhodobacter capsulatus* ($ALAS_{Rc}$) binding its cofactor PLP. We, furthermore, present structures of $ALAS_{Rc}$ in complex with the substrates glycine or sCoA. The sequence identity of ALAS from *R. capsulatus* and human eALAS is 49%. XLSA-causing mutations may thus be mapped, revealing the molecular basis of XLSA in humans. Mutations are found to obstruct substrate binding, disrupt the dimer interface, or hamper the correct folding. The structure of ALAS completes the structural analysis of enzymes in heme biosynthesis.

The EMBO Journal (2005) **24,** 3166–3177. doi:10.1038/sj.emboj.7600792; Published online 25 August 2005
Subject Categories: structural biology
Keywords: aminolevulinic acid synthase; bacteriochlorophyll; PLP; tetrapyrrole biosynthesis; X-linked sideroblastic anemia

Introduction

Tetrapyrroles, such as hemes and chlorophylls, are arguably among the most important pigments in the biosphere and the most versatile of cofactors. They are required for molecular, energy, and electron transport, O_2 and CO_2 sensing, as well as in countless enzymatic reactions (Dailey, 1997; Frankenberg *et al*, 2003).

*Corresponding author. Division of Structural Biology, German Research Centre for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany. Tel.: +49 531 6181 764; Fax: +49 531 6181 763; E-mail: wds@gbf.de or dih@gbf.de
[3]These authors contributed equally to this work

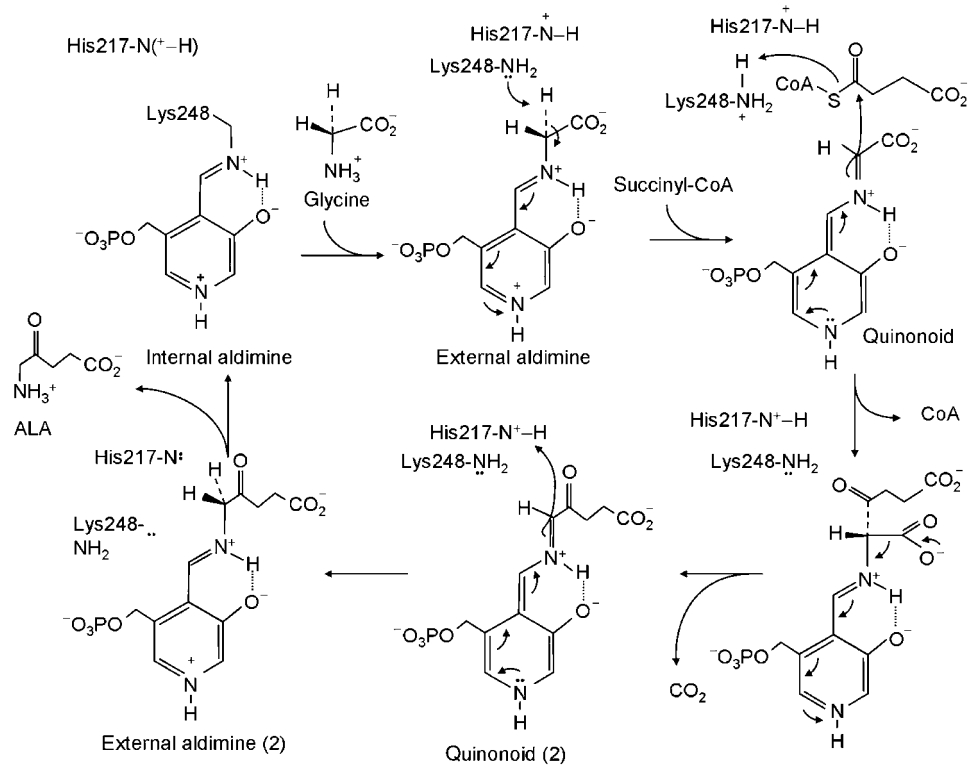
Received: 10 May 2005; accepted: 29 July 2005; published online: 25 August 2005

The initial enzyme of tetrapyrrole biosynthesis, 5-aminolevulinate synthase (ALAS; EC 2.3.1.37), was discovered by Shemin (Shemin and Kikuchi, 1958) and Neuberger (Gibson *et al*, 1958). ALAS catalyzes the decarboxylative condensation of glycine and succinyl-coenzyme A (sCoA) yielding 5-aminolevulinic acid (ALA), the universal precursor of all tetrapyrroles. Much later, ALAS was found to be restricted to non-plant eukaryotes and the α-group of proteobacteria (Jordan, 1990). In plants and most bacteria, ALA is instead derived from tRNA-bound glutamate in two enzymatic steps via the C_5 pathway involving glutamyl-tRNA reductase (Hungerer *et al*, 1996; Moser *et al*, 1999, 2001) and glutamate-1-semialdehyde-2,1-aminomutase (Hennig *et al*, 1997).

Humans and other mammals produce two isoforms of ALAS: housekeeping hALAS, ubiquitously expressed in all cell types, and erythroid-specific eALAS, exclusively produced in developing erythrocytes (Riddle *et al*, 1989). eALAS produces ~90% of the total body heme. ALAS is highly conserved from α-proteobacteria to humans, the sequence identity ranging from 40 to 80%. Thus, for example, the sequence identity of human eALAS and *Rhodobacter capsulatus* ALAS ($ALAS_{Rc}$) is 49%, rising to 70% if similar residues are included.

ALAS is a pyridoxal 5′-phosphate (PLP)-dependent homodimer, composed of 44–65 kDa monomers. It belongs to the oxoamine synthase subgroup of α-division PLP-dependent enzymes (Mehta and Christen, 1994; Christen and Mehta, 2001). The catalytic mechanism of the ALAS reaction involves binding of glycine to PLP to form a PLP–glycine complex (external aldimine; Scheme 1). Following proton abstraction, the resonance-stabilized quinonoid intermediate nucleophilically attacks sCoA, the second substrate. Release of CoA and decarboxylation of the resulting intermediate yields ALA. The final step, release of ALA, is the slowest and hence rate-limiting step of the reaction (Ferreira *et al*, 1993; Zhang and Ferreira, 2002; Cheltsov *et al*, 2003).

Mutations in the human gene for eALAS (*ALAS2*) cause X-linked sideroblastic anemia (XLSA), a group of severe disorders characterized by inadequate formation of heme and accumulation of iron in erythroblast mitochondria (May and Bishop, 1998). The physiological concentration of iron is normally tightly regulated to prevent nonspecific iron binding (Andrews, 1999). Iron overloads, as in XLSA, disturb the cellular reduction–oxidation state and reduce cellular lifetime, making eALAS critical for erythroid cell differentiation and survival (Harigae *et al*, 2003). A biological half-life of a mere hour means that mammalian eALAS is notoriously difficult to purify from natural sources (Furuyama and Sassa, 2002). We have therefore selected bacterial ALAS for the structural elucidation, especially as the high degree of sequence identity between bacterial and the core human ALAS should allow the former to serve as a suitable structural model for the latter.



Scheme 1 The reaction mechanism of ALAS. In the substrate-free state, the cofactor PLP is bound by Lys248 ($ALAS_{Rc}$), the internal aldimine. Incoming glycine induces transaldimination, leading to PLP binding glycine rather than Lys248 (external aldimine). The substrate sCoA is nucleophilically attacked by the PLP-activated glycine, leading to the addition of succinic acid to glycine and concomitant loss of CoA. Decarboxylation (carboxylate of the glycine moiety) of this intermediate yields the PLP-bound product, released in the rate-limiting step of the reaction.

Here we describe the first crystal structure of full-length ALAS from *R. capsulatus* ($ALAS_{Rc}$), including the cofactor PLP, at 2.1 Å resolution. In addition, we present the structures of $ALAS_{Rc}$, respectively, binding glycine and sCoA at 2.7 and 2.8 Å. The high degree of conservation between $ALAS_{Rc}$ and human eALAS allows us to derive a model for the core of human eALAS, elucidating the structural and functional implications of naturally occurring XLSA-inducing mutations. In addition, we can now elucidate the structural aspects of catalysis of the oxoamine synthase subfamily of PLP-dependent enzymes in unprecedented detail.

Results and discussion

The structure of ALAS

ALAS has previously been identified as a member of the oxoamine synthase subgroup of α-division PLP-dependent enzymes (Grishin *et al*, 1995; Mehta and Christen, 2000; Schneider *et al*, 2000). The structure of ALAS from *R. capsulatus* ($ALAS_{Rc}$) was correspondingly solved by molecular replacement, using a model consisting of the superimposed structures of 8-amino-7-oxononanoate synthase (AONS; PDB code 1BS0; Alexeev *et al*, 1998) and 2-amino-3-ketobutyrate CoA ligase (KBL; PDB code 1FC4; Schmidt *et al*, 2001) of this subgroup. The structure was analyzed to a maximum resolution of 2.1 Å, with two dimers in the asymmetric unit. Data collection and refinement statistics are summarized in Table I.

$ALAS_{Rc}$ is a tightly interlocked homodimer (Figure 1). Each monomer consists of three domains, all of which participate in dimerization: an N-terminal domain (residues 1–52, yellow in Figure 1), a central, catalytic domain (residues 53–296, orange) that contributes most of the dimer interface, and a C-terminal domain (residues 297–401, red). The N-terminal domain consists of an α-helix, remote from the rest of the monomer, and a three-stranded, antiparallel β-sheet. The large central or catalytic domain bears a seven-stranded, largely parallel β-sheet (β-strand order: 324567′1, ′ implying an antiparallel orientation) covered on either side by nine α-helices in repeated β/α motifs. The C-terminal domain includes a three-stranded antiparallel β-sheet contacting both N-terminal and catalytic domains of the same monomer and three α-helices contacting the N-terminal domain.

PLP-binding site and active site pocket

An $ALAS_{Rc}$ dimer symmetrically binds two PLP molecules (green in Figure 1). In the absence of substrate or product, PLP is covalently bound as a Schiff base to Lys248 (Hunter and Ferreira, 1999a). This lysine, conserved in all ALAS, is located in loop β9–β10 of the catalytic domain. The catalytic domain of the second monomer participates in PLP binding by providing three of six hydrogen bonds to the phosphate group: Ser277*-O_γ and Thr278*-O_γ and -N (where * denotes the second monomer), in addition to Ala115-N, Tyr116-N, and Thr245-O_γ (Figures 2 and 3A). Recognition of the pyridinium ring of PLP, by contrast, exclusively involves one monomer: Val216 and His142 stack on either side of

Table I Data collection and structure determination statistics

	ALAS/PLP	ALAS/PLP/Gly	ALAS/PLP/sCoA
Data collection			
Unit cell dimensions, *a*, *b*, *c* (Å)	67.9, 92.9, 250.1	68.0, 92.0, 248.6	67.9, 91.1, 247.2
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Wavelength (Å)	1.05	0.95	0.95
Number of unique reflections	78265	42365	38554
Resolution range (Å)	50–2.1 (2.18–2.10)	30–2.7 (2.8–2.7)	30–2.8 (2.9–2.8)
Completeness of data (%)	83.8 (82.7)	96.4 (92.7)	99.6 (99.7)
Redundancy	4.7 (4.7)	4.1 (3.7)	3.6 (3.5)
R_{merge} (%)	7.7 (34.0)	6.1 (37.2)	7.6 (36.2)
I/σ	13.7 (3.8)	16.8 (2.9)	13.5 (2.8)
Refinement statistics			
Maximal resolution (Å)	2.1 (2.18–2.10)	2.7 (2.77–2.70)	2.8 (2.88–2.80)
No. of atoms: protein, cofactor, water	12202, 30, 512	12190, 80, 144	12220, 284, 142
Monomers per asymmetric unit	4	4	4
R-factor (%)	15.6 (17.0)	17.6 (29.6)	16.5 (26.8)
R_{free} (%)	21.7 (27.0)	23.6 (38.2)	23.6 (38.3)
Average *B*-factor (Å^2)	23.9	39.6	34.2
R.m.s.d. bond length (Å)	0.02	0.02	0.02
R.m.s.d. bond angles (deg)	1.61	1.68	1.69
Ramachandran plot[a]	91.1/8.2/0.7/0	91.6/7.8/0.6/0	90.3/8.8/0.7/0.1

Values in parentheses refer to the shell of highest resolution.
[a]Procheck: most favored/additionally allowed/generously allowed/disallowed regions.

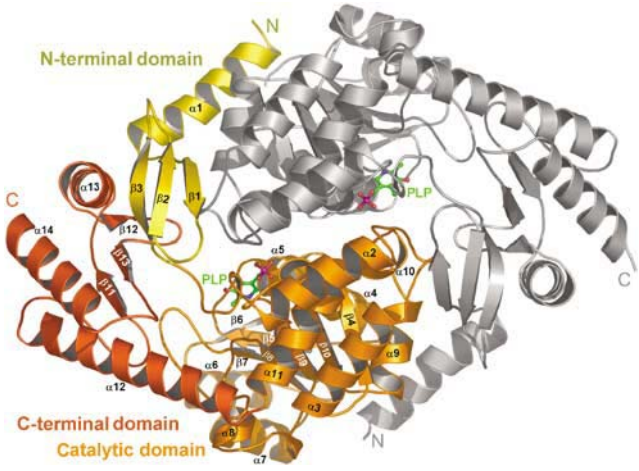


Figure 1 The structure of the ALAS homodimer from *R. capsulatus* ($ALAS_{Rc}$) in ribbon representation. The N-terminal domain is rendered in yellow, the catalytic domain in orange, and the C-terminal domain in red. The cofactor PLP (green) is depicted in ball-and-stick representation, carbon in green, oxygen in red, nitrogen in blue, and phosphorus in magenta. This figure and most others were produced using PyMol (DeLano, 2002).

the pyridinium ring, stabilizing it through hydrophobic interactions. Specificity is achieved through the hydrogen bond His217-$N_{\varepsilon 2}$–PLP-O3 and the salt bridge Asp214-$O_{\delta 2}$–PLP-$N_{\delta 2}$ (the protonated nitrogen of the pyridinium ring). The latter interaction, found in most PLP-dependent enzymes, increases the electron-withdrawing potential of the pyridinium ring (Hunter and Ferreira, 1999b). The tight binding allows PLP to remain bound during catalysis, despite losing its covalent bond to Lys248. PLP, instead, binds the substrate glycine through a Schiff base bond.

Alongside each PLP, both catalytic domains and the C-terminal domain of the PLP-binding monomer delimit the substrate-binding pocket. This pocket is connected to the enzyme surface by a channel (Figure 4B) created by the same domains involved in forming the active site. This channel is clearly involved in sCoA binding (see below).

Binding of substrates glycine and sCoA

To elucidate the binding mode of both substrates, we soaked crystals of $ALAS_{Rc}$/PLP with either glycine or sCoA for 24 h. The resulting complexes were analyzed crystallographically to maximum resolutions of 2.7 and 2.8 Å, respectively. Unit cell parameters and space group essentially remained unchanged (Table I), allowing glycine and sCoA to be located by difference Fourier techniques, following rigid-body and simulated-annealing refinement using the $ALAS_{Rc}$/PLP structure.

The difference map of glycine-soaked crystals reveals unambiguous residual electron density alongside PLP in all four monomers in the asymmetric unit (Figure 3B). Glycine is thus covalently bound to PLP through a Schiff base bond, forming the so-called 'external aldimine'. The occupancy of glycine, however, appears to vary, underlining the dynamic nature of the Schiff base and the known fact that the

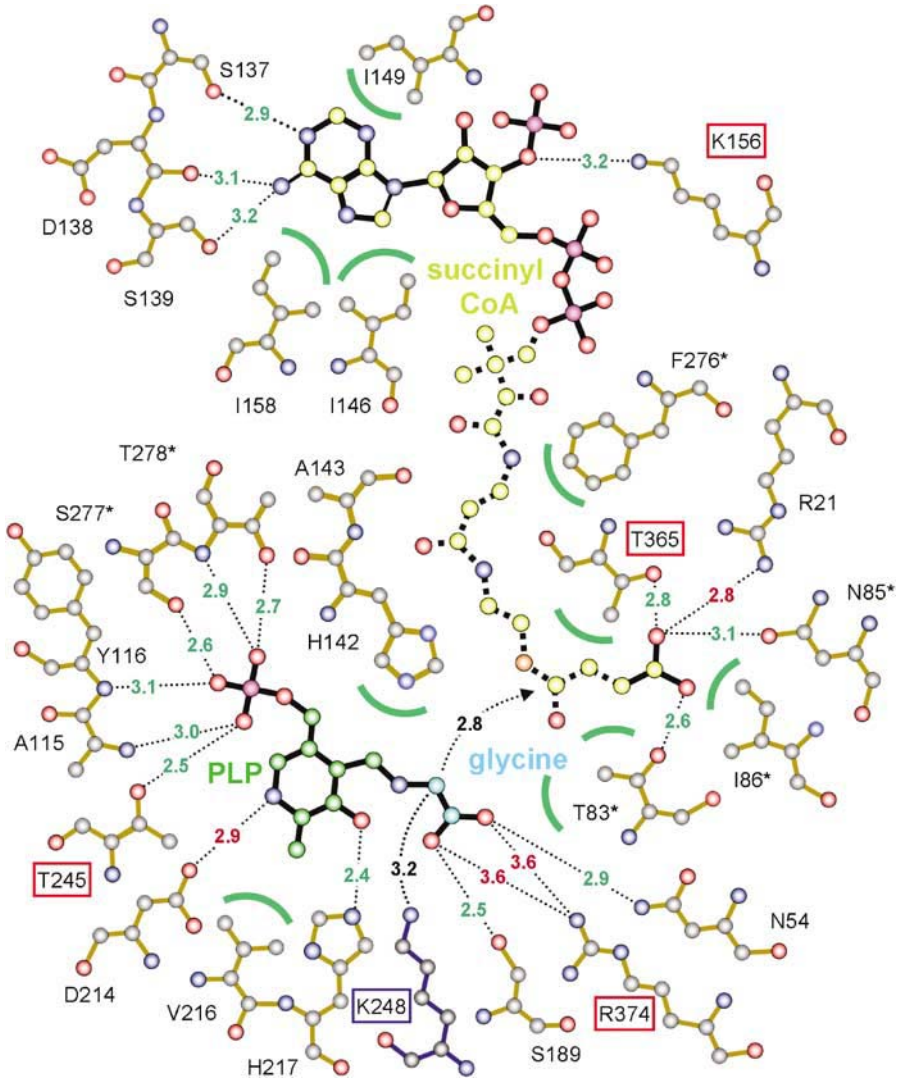


Figure 2 Schematic representation of the active site. Glycine-bound PLP and sCoA are highlighted by bonds in black; bonds of surrounding residues are shown in dark orange. Dotted lines indicate hydrogen bonds (green numbers) and salt bridges (red) between substrates or cofactors and ALAS; green semicircles indicate hydrophobic interactions. An asterisk marks residues from the second monomer. Lys248, involved in PLP binding and catalysis, is marked by a blue box, residues affected by mutations in human eALAS by a red box.

dissociation constant of ALAS for glycine is surprisingly high at 1.8–30 mM (Bolt *et al*, 1999). The glycine α-carboxylate group forms a weak salt bridge (3.8 Å) with the guanidinium side chain of Arg374 (confirming its involvement; Tan *et al*, 1998), and hydrogen bonds to Ser189-O_γ and Asn54-$N_{\delta 2}$.

In sCoA-soaked crystals, significant difference electron density is seen near the entrance of the active-site channel. In particular, the 3′-phosphate ADP moiety of sCoA is well defined (Figures 3D and E)—the first time sCoA and hence its binding pocket have been located in any member of the oxoamine synthase subfamily of PLP-dependent enzymes. It is bound within a hydrophobic pocket on the surface of the enzyme. Residues Ile146 and Ile158 lie on either side of the adenine moiety, while Ile149 delimits the bottom of the pocket (Figures 2, 3D and E). Specificity is achieved through hydrogen bonds Ser139-O–N6, Asp138-O–N6, and Ser137-O_γ–N1, where N1 and N6 belong to the sCoA adenine moiety. The ribose moiety of sCoA, in turn, is recognized by a hydrogen bond from Lys156-N_ζ to O3′. The binding pocket is located immediately alongside the entrance to the active-site channel (Figure 4B). At the end of the active-site channel, a planar, three-pronged electron density opposite the guanidinium group of Arg21 is visible (Figure 3E). We have interpreted this to mark the position of the succinate carboxylate group of sCoA, forming a salt bridge to Arg21-$N_{\eta 2}$. This carboxylate oxygen also forms hydrogen bonds to Thr365-O_γ and Asn85*-N_δ, while the second oxygen forms hydrogen bonds to Thr83*-O_γ and a buried water molecule (Figure 2). Lateral van der Waals interactions with Ile86* further serve to stabilize this interaction. The carboxylate group is also seen in the absence of sCoA—indicating the high affinity for acidic groups. Only weak electron density is visible for the intermediate region better modeled as discrete water molecules. The succinyl-pantetheine moiety of sCoA is thus disordered in the absence of the second substrate glycine. Nevertheless, the positions of the 3′-phosphate ADP and carboxylate moieties allow the entire sCoA to be modeled into the active site tunnel in an elongated conformation (Figure 3E and dashed black lines in Figure 2). The modeled positions of hydrophilic atoms of sCoA (OP1, NP1, OS1)

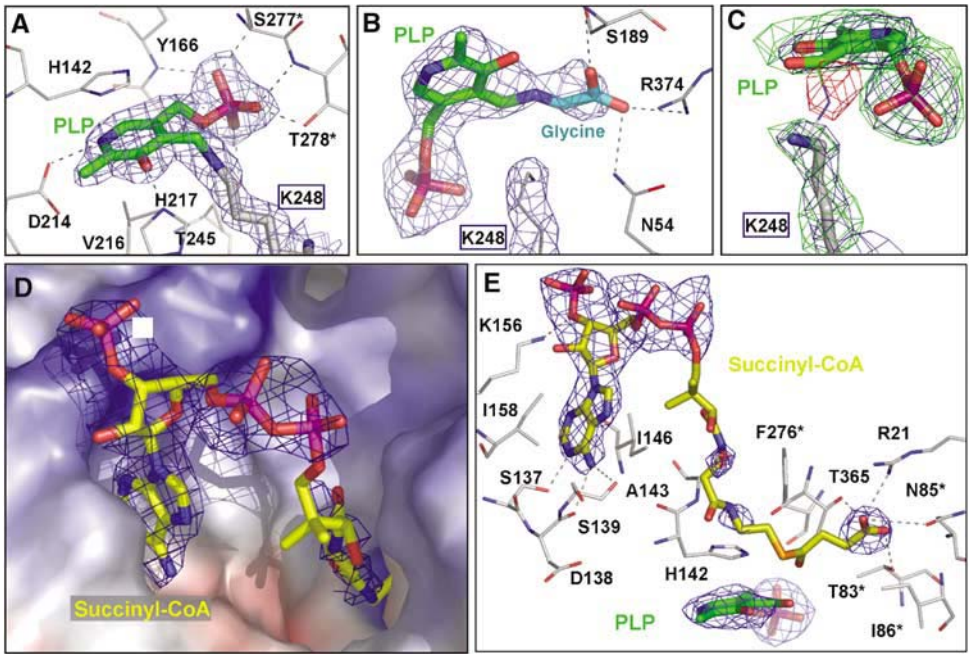


Figure 3 Detailed views of cofactor PLP and substrates in combination with their electron densities. (**A**) Electron density of the internal aldimine consisting of PLP (green) covalently bound to Lys248. Only residues in contact with PLP are shown. Residues of the second subunit are marked by an asterisk. (**B**) Electron density of PLP–glycine intermediate (cyan). The view is rotated compared to (A) to indicate the residues involved in glycine binding. Arg374 appears vital for glycine recognition. The interaction is, however, weakened by covalent bond formation—possibly aiding later decarboxylation. (**C–E**) The binding of sCoA to ALAS: (C) In three of four monomers of the sCoA/ALAS complex, PLP is not covalently bound to Lys248. This is documented by positive difference electron density in the absence of PLP and side chain of Lys248 (green, contoured at 3σ), the negative difference density in 'enforced' Schiff-base bond (red, contoured at −3σ, structure in narrow bonds), and the refined electron density of unbound PLP (blue, contoured at 1σ, thick bonds). (D) The adenine and ribose moieties of sCoA bind in a hydrophobic (white) pocket with positively charged rim (blue) adjacent to the active-site channel. (E) The electron density is well defined for the 3′-phosphate ADP moiety and for the carboxylate group of sCoA, indicating that the central portion is not rigidly bound in the absence of the second substrate glycine.

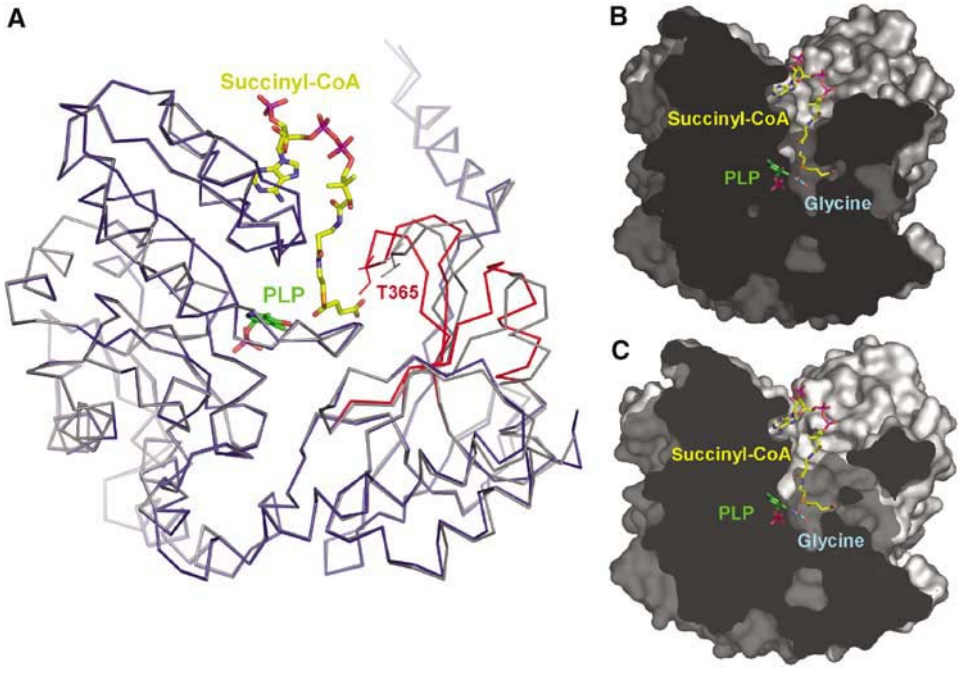


Figure 4 Open and closed conformations in ALAS. (**A**) Monomers A (blue) and B (gray) are overall quite similar, except for the loops β11–α14 and α14–β12 that adopt a closed conformation in monomer A (red). The open conformation may aid entry of sCoA to the active site, while the closed conformation clamps it into position prior to catalysis. In the second dimer in the asymmetric unit (and in all monomers of the substrate-bound complexes) only the closed conformation is observed. (**B, C**) The molecular surface of the ALAS dimer, showing a cross section through the active site in the closed (B) and in the open (C) conformation. The complete set of substrates and cofactors are included in (B) and (C) to indicate good fit in the closed conformation, allowing for high substrate specificity.

roughly correspond to the water molecules in the channel, indicating that these atoms would be stabilized by hydrogen bonds to the surrounding protein, while the remaining aliphatic parts are stabilized by hydrophobic van der Waals interactions.

In the substrate-free structure of ALAS, PLP is covalently bound by Lys248. The substrate glycine replaces Lys248 in binding to PLP, as expected (Figures 3B). However, though sCoA does not bind to PLP itself, PLP appears no longer to be bound to Lys248 in at least three (of four) monomers of this complex (Figure 3C). Instead, the pyridinium ring of PLP is tipped by $\sim 15^{\circ}$ around the C5–C5A bond, so that its O3 and C4A atoms move away from Lys248—comparable to the glycine-bound state and to other external aldimines of PLP-dependent enzymes. The rotation of the PLP ring is accompanied by a movement of His142 in all substrate-bound monomers, such that it remains stacked above the PLP ring. Lys248 forms a new hydrogen bond through its N_{ζ} to $N_{\gamma 2}$ of the conserved residue Asn54.

Superimposing the glycine- and sCoA-bound complexes of $ALAS_{Rc}$ indicates that both substrates can simultaneously be accommodated within the active site, without incurring steric clashes (Figure 4B). Overall, this leads to an arrangement in which the distance between the glycine C_{α} atom and CS1 of sCoA is ~ 2.8 Å. The nucleophilic attack of CS1 by the deprotonated C_{α} of the quinonoid, anionic glycyl-PLP to create the new C–C bond is thus feasible.

Glycine-rich stretch

A glycine-rich stretch comprising residues 77-GAGSGGTR NISGT-89, in loop α3–α4 of the catalytic domain, was earlier proposed to be involved in substrate binding (Gong *et al*, 1996). Especially three residues, Gly77, Gly79, and Arg84, were found to be intolerant of mutational replacement. Indeed, the structure indicates that this loop is crucial for the recognition of the succinyl-moiety of sCoA. In particular, residues Thr83, Asn85, and Ile86 are directly involved in positioning the terminal carboxylate group of sCoA (Figure 3E). Structurally, this intricately folded loop is furthermore surrounded by the catalytic domain of the same monomer and all three domains of the second monomer. It therefore occupies a particularly sensitive position that may easily be disrupted by any change in the amino-acid sequence. Arg84 forms a salt bridge to Glu97, helping to stabilize the loop. Gly77 adopts a conformation reserved for glycines. Mutated to any other residue, it would be forced into a different region of the Ramachandran plot, again disrupting this loop and upsetting the recognition of succinyl.

Open and closed ALAS conformations

The crystal packing in the complex $ALAS_{Rc}$/PLP places two ALAS dimers (or four monomers) into the asymmetric unit. Three of these monomers, A, D, and E, are structurally essentially indistinguishable. Monomer B, by contrast, adopts an 'open' conformation in which two loops (residues 332–348 between β11 and α14 and residues 358–374 between α14 and β12; red in Figure 4A) of the C-terminal domain move outwards, widening the active site, the active-site channel, and opening a second channel, both longer and narrower than the first (Figure 4C). Comparing the 'open' and 'closed' states of ALAS indicates that in particular Thr365, located at the outermost tip of the loop 358–375, moves inwards by 3.5 Å during this transition. In the closed conformation, Thr365 is firmly positioned among the residues of the active site stabilized by a hydrogen bond Thr365-$O_{\gamma 1}$–Arg21-$N_{\eta 2}$ and by van der Waals interactions Thr365-$C_{\gamma 2}$–Ile361-$C_{\gamma 1}$ and –Met190-C_{ε}. As a result, the narrow active-site channel winds past Thr365, ending opposite the side chain of Arg21. The carboxylate group of the substrate sCoA presumably fills this end of the channel, forming a hydrogen bond to Thr365-O_{γ} (see above). Thr365 is thus crucial in locking the succinyl group into position and is itself locked into the closed position when the substrate is present.

The open conformation is observed only in one monomer of the substrate-free crystal structure. In glycine- and sCoA-bound complexes, all monomers adopt the closed conformation. The open state presumably provides more facile access to the active site, especially for the extensive sCoA. The closed conformation in turn constricts the active site and active-site channel, clasping the pantetheine and succinate moieties of sCoA into position—presumably as a prelude to catalysis. An additional salt bridge Arg368–Glu17 further stabilizes the closed conformation. Arg368 is conserved in comparison with human eALAS, while Glu17 of $ALAS_{Rc}$ is replaced by aspartate (Asp159) in eALAS (Figure 5A).

Structural aspects of enzymatic reaction

Structurally and functionally, ALAS belongs to the α-division of PLP-dependent enzymes and along with AONS, serine palmitoyl transferase, and KBL constitutes the α-oxoamine synthase subfamily. These enzymes catalyze the decarboxylative condensation of a carboxylic acid CoA thioester and an amino acid (Mehta and Christen, 2000). Of the two substrate molecules, the amino acid (glycine in ALAS) is covalently bound through a Schiff base, while CoA is bound noncovalently.

The crystal structure of $ALAS_{Rc}$ indicates that the PLP–glycine intermediate forms a weak salt bridge (3.8 Å) to the conserved Arg374 (Figure 3B). Replacing the arginine by lysine in murine eALAS reduces its enzymatic activity, but only by 23% (Tan *et al*, 1998). Clearly, a positively charged residue is required at this position, presumably to aid initial recognition and orientation of glycine prior to PLP binding. Formation of the Schiff bond in the external aldimine (Scheme 1) would cause glycine to move towards PLP and away from Arg374, weakening the salt bridge interaction. Though overall stability of the glycine would be sacrificed, a weaker salt bridge may favor elimination of the α-carboxylate as CO_2 later in the reaction.

The side-chain methyl group of Thr83* and the C_{α} atom of PLP–glycine intermediate are in direct van der Waals contact (3.9 Å) with each other. Amino acids other than glycine could theoretically bind only if their side chains face away from Thr83*. This would, however, block the binding site of sCoA, indicating that Thr83* is pivotal in governing the specificity of ALAS for glycine. In AONS Thr83* is replaced by a serine (Alexeev *et al*, 1998), which lacks the terminal methyl group. Alanine, larger by exactly this methyl group, is hence the favored substrate of AONS. Accordingly, ALAS may be modified to accept threonine (and to a lesser extent serine, alanine and glycine) by engineering the replacement of Thr83 by serine, as shown for ALAS from *R. sphaeroides* (Shoolingin-Jordan *et al*, 2003).

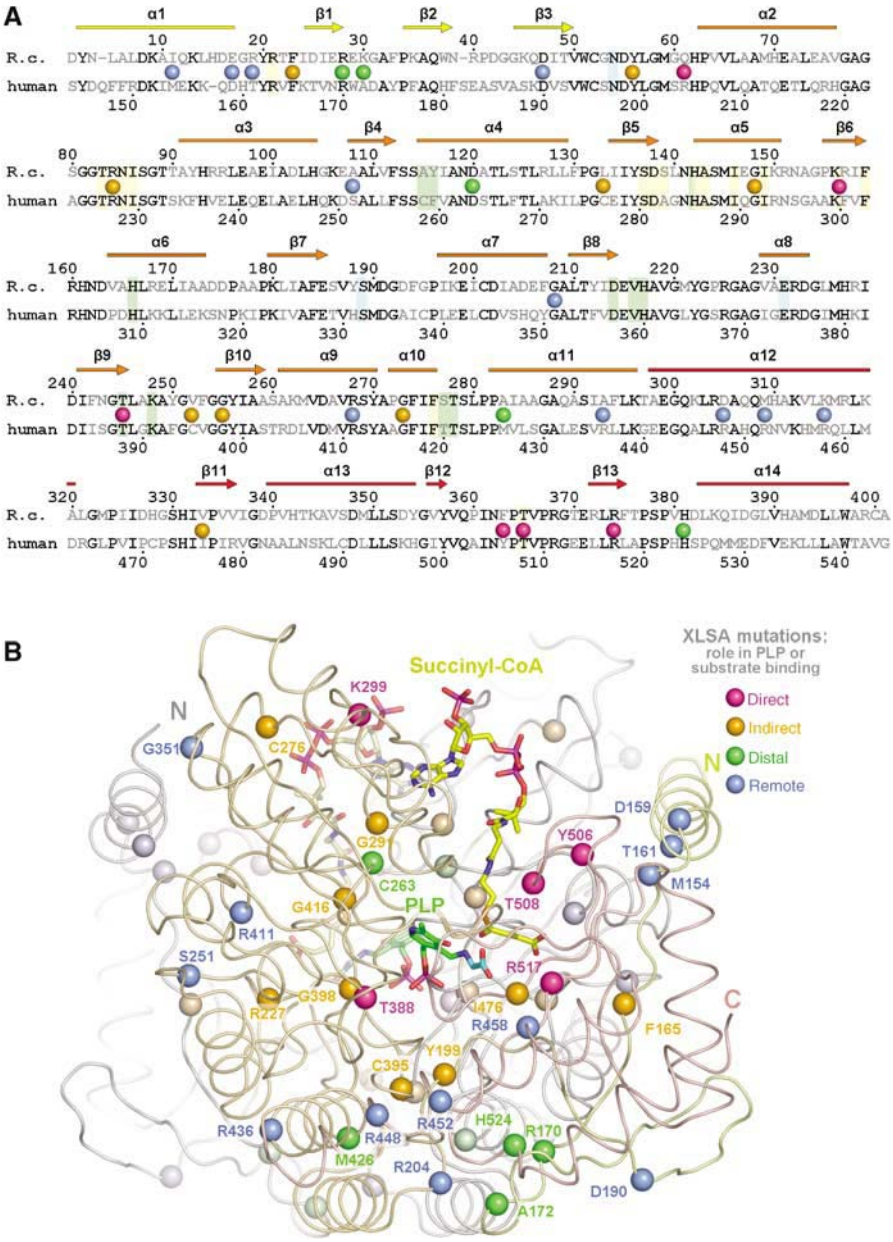


Figure 5 XLSA-causing mutations in human eALAS. (**A**) Sequence alignment of $ALAS_{Rc}$ and human eALAS. Conserved residues are shown in black, all others in gray. Rectangles and arrows (colors as in Figure 1) above the alignment indicate α-helices and β-strands, respectively. Colored spheres highlight residues replaced in XLSA-causing mutations. Magenta—cofactor and substrate binding; green—mutations affecting cofactor and substrate binding indirectly; orange—mutations in the hydrophobic core; blue—surface-exposed mutation sites. Colored boxes indicate residues directly involved in substrate/cofactor binding (see Figure 2)—green, cyan and yellow indicate residues involved in PLP, glycine, and sCoA binding. (**B**) A model of human eALAS. Residues replaced in XLSA mutants are indicated by colored spheres. Color-coding for spheres and domains as described in (A). Pale colors indicate the second monomer.

The sCoA carboxylate group is recognized by Arg21 at the far end of the substrate channel. Another two hydrogen bonds from Asn85* and Thr365 add to this recognition. A mutation of Thr365 (Thr508 in human eALAS) to serine induces XLSA (see below), confirming its crucial participation in substrate binding.

We observe that sCoA binds in the absence of glycine, indicating that binding is initiated by 3′-phosphate adenosine recognition. Only upon glycine binding would the succinyl-pantetheine 'tail', however, fill the remaining channel. The conformational equilibrium would then presumably be tipped towards the closed conformation, constricting the active-site tunnel and eliminating any excess water molecules, initiating the enzymatic reaction.

In vitro activity assays with several acyl-CoA derivatives indicate the importance of the acyl chain length and hydrophilicity (Shoolingin-Jordan *et al*, 1997). The crystal structure underscores these conclusions, as the optimal chain length is clearly dictated by the successful positioning of CS1 opposite glycine-C_α following succinyl carboxylate recognition at the far end of the channel. This also explains why increasing the acyl-group hydrophilicity by substituting the carboxylate by carbonyl, hydroxyl, or methyl groups should negatively influence enzymatic activity. Tight binding of the succinyl

carboxylate group furthermore slows the release of ALA, making it the rate-limiting step in the ALAS-catalyzed reaction (Hunter and Ferreira, 1999b).

Modeling human eALAS on ALAS from R. capsulatus

Before the present crystal structure of ALAS from *R. capsulatus*, structural information on ALAS was available only for the 49 amino-acid presequence of murine ALAS from an NMR study (Goodfellow *et al*, 2001), and by modeling the catalytic domain of ALAS on related enzymes such as AONS (Cheltsov *et al*, 2003; Shoolingin-Jordan *et al*, 2003).

Excluding the variable N-terminal regulatory and targeting domain (140 residues; Conboy *et al*, 1992), the catalytic core of human eALAS is 49% identical (70% similar) by sequence to $ALAS_{Rc}$ (Figure 5A). This is a particularly high degree of conservation between a human and bacterial protein, and is significantly higher than the 29% between AONS and $ALAS_{Rc}$. The sequence similarity is uniformly distributed throughout the length of 401 residues. It is, however, particularly pronounced in the active site, the active-site channel, the monomer–monomer interface, and the protein core (Figure 5A). Compared to $ALAS_{Rc}$, human eALAS bears two single-residue insertions and a single deletion in the first 40 residues of $ALAS_{Rc}$, in the least-conserved part of the sequence. The high degree of amino-acid conservation and the fact that the enzymatic activity of truncated human eALAS (K_M 2.6 mM glycine) is similar to that of $ALAS_{Rc}$ (2.8 mM), as well as ALAS from *R. sphaeroides* (5.8 mM), indicate that the structure of $ALAS_{Rc}$ represents a valid model for human eALAS.

Mutations of eALAS cause XLSA

Hereditary sideroblastic anemia or XLSA is caused by point mutations in *ALAS2*. In total, 42 such substitutions affecting 37 residues of human eALAS have been identified (Table II; reviewed by Bottomley, 2004). Some of these cases have been successfully treated by supplemental pyridoxine (vitamin B_6, converted to PLP on uptake), phlebotomy (bleeding), or iron-chelating drugs. The success of treatment, however, varies especially due to the differential responsiveness of mutated eALAS patients to pyridoxine. The fact that many of the XLSA-sensitive residues of eALAS are either strictly conserved in $ALAS_{Rc}$ or replaced by similar residues (Table II, columns 1 and 2) now permits the precise location of these mutation sites to be mapped and analyzed (Figure 5B).

In our discussion of the mutations, we have assigned four groups of XLSA mutations depending on their proximity to the active site.

XLSA mutations directly involved in substrate or PLP binding (magenta in Figure 5A and B and Table II)

R517C: Arg517 (Arg374 in $ALAS_{Rc}$) engages the substrate glycine carboxylate group in a salt bridge, and is thus crucial for substrate recognition and discrimination (see above; Tan *et al*, 1998). Replacing arginine by shorter, uncharged cysteine would severely reduce the substrate specificity. This effect cannot be corrected by higher concentrations of PLP, hence confirming the observation that individuals carrying this mutation do not respond to pyridoxine therapy.

T388S: Thr388 (Thr245) is part of the PLP recognition pattern by hydrogen bonding the PLP phosphate. Replacing threonine by serine allows for higher rotational freedom of serine, significantly reducing the affinity of eALAS for PLP. A lower affinity for PLP may be offset by increasing its concentration. Correspondingly, patients respond favorably to pyridoxine treatment.

T508S: Thr508 (Thr365) is central to the function of ALAS. It appears essential both for active-site closure (see above) and in stabilizing sCoA. This stabilization involves hydrophobic interactions to the succinyl-moiety and a hydrogen bond to the succinyl-carboxylate group. Removing the methyl group by replacing threonine by serine would impair the balance of interactions in the active site. In particular, conformational constraint on the succinyl-moiety of sCoA would be lost, affecting the position of atom CS1 of sCoA, the atom that attacks the deprotonated C_α of the PLP–glycine aldimine. The lower affinity for sCoA cannot be counteracted by an increase in the amount of PLP. Correspondingly, this mutation is not convincingly pyridoxine responsive in patients.

K299Q: Lys299 (Lys156), located on the surface of ALAS, binds the O3′ atom of sCoA (ribose moiety) through a hydrogen bond. Replacing lysine by glutamine causes this interaction to be lost, decreasing the binding affinity of eALAS for sCoA. Patients with this mutation are reported to respond to pyridoxine treatment (Table II). However, as PLP binding in ALAS is not affected by the mutation, the effect of pyridoxine supplementation may be indirect.

Y506 frame shift and R204 stop: Two mutations cause the polypeptide to be terminated prematurely. Termination at position 204 (61) removes the major part of the catalytic domain, while a frame shift after position 506 (363) leads, at the very least, to the loss of Arg517 (Arg374), required for glycine binding. Both polypeptides will thus not give rise to a functional enzyme, and increasing the concentration of PLP will not rescue their activity.

XLSA mutations indirectly affecting cofactor or ligand binding (orange in Figure 5A and B and Table II). The following XLSA mutations involve residues not directly in contact with the substrates glycine and sCoA, nor with PLP. They are, however, located within the core of the ALAS monomer, frequently in the immediate structural vicinity of the active site. Replacing them with physically incompatible residues upsets the active site or the binding affinity of cofactor or substrates.

F165L: Phe165 (Phe23) structurally stabilizes the neighboring Arg163 (Arg21, Figure 2), required for sCoA carboxylate group recognition, through hydrophobic interactions. Replacing Phe165 by leucine destabilizes Arg163, reducing the specificity of sCoA binding. As PLP binding is not affected by the mutation, response to pyridoxine treatment is marginal.

C276W: Cys276 (Leu133) is located at the N-terminal end of β-strand β5 (Figure 5A), filling a small hydrophobic pocket. It is physically linked to residues Ser280 (Ser137) and Asp281 (Asp138) involved in recognition of the sCoA adenine moiety. Replacing this residue with the much bulkier tryptophan will severely reduce sCoA binding. As PLP binding is not affected, patients are not responsive to pyridoxine supplementation.

G291S: Gly291 (Gly148) is intermediate to Ile289 (Ile146) and Ile292 (Ile149), both part of the sCoA adenine-binding pocket. Replacing glycine by serine will upset the pocket, decreasing the affinity for sCoA. Furthermore, Gly291 is one

Table II Overview of known XLSA mutations in human ALAS

Mutation in eALAS[d]	Corresponding residue in $ALAS_{Rc}$[a]	Occurs in	Severity[b]/ pyridoxine responsiveness[c]	Affects binding of			Structural impairment
				PLP	Gly–cine	sCoA	
M154I	$I11^{-}$	♂	?/?				Near Asp408 (C-term. domain) and Asp159, required for active-site closure
D159N	$D16^{+}$	♂	−−/+?			+	Loss of salt bridge (SB) to Arg511 (C-term. domain) involved in stabilizing closed active-site conformation
D159Y		♂	−−/+				
T161A	$R19^{-}$	♂	?/?			(+)	At interface of N-, C-term. and catalytic* domain; may affect Arg163 involved in sCoA binding
F165L	$F23^{++}$	♂	−−/+			+	Destabilizes Arg163 required for sCoA binding; affects interaction of N- and C-terminal domains
R170S	$R28^{++}$	♂	−−/+	+			Disrupts buried SB to Asp198 destabilizing glycine-rich stretch and hence PLP binding. Mutation to histidine more severe due to larger size
R170C		?	?/?				
R170H		♀	−/−				
R170L		♂	−−/+				
A172T	$K30^{-}$	♀	−−/++	+			Affects Gly220 of glycine-rich stretch, indirectly reducing PLP-binding affinity; at interface, destabilizing enzyme
D190V	$D47^{++}$	♂	−/−				Remote from active site on enzyme surface; may affect binding of sCoA synthetase (Furuyama and Sassa, 2000)
Y199H	$Y56^{++}$	♂	−/+	+	+		Destabilizes His360 (PLP binding), Arg517 (glycine recognition), and Lys391 (catalysis)
R204stop	$Q61^{++}$	♀	−−/−	+	+	+	Loss of enzyme function through deletion of catalytic and C-terminal domains
R204Q		♂	−−/+			+	Loss of SB to Asp173, weakening interaction of N-term. and catalytic domains; affects glycine-rich stretch
R227C	$R84^{++}$	♀	−/−		+	+	Disrupts adjacent Thr226 involved for sCoA and glycine binding by loss of SB to Glu240
S251P	$A108^{+}$	♀	−−/?	+			General destabilization by placing wide, hydrophobic residue into restrictive pocket on protein surface
D263N	$D120^{+}$	♂	−−/++?			+	Close to molecular symmetry axis, that is near Asp263*; possible loss of divalent metal binding
C276W	$L133^{+}$	♀	−−/−			+	Ser280 and Asp281 in binding pocket of sCoA-adenine moiety disrupted; binding affinity of sCoA reduced
G291S	$G148^{++}$	♂	−−/++	+		+	Disrupts binding pocket of sCoA-adenine moiety, reducing affinity for sCoA
K299Q	$K156^{++}$	♂	−−/++			+	Affinity for succinyl-CoA reduced through loss of hydrogen bond (HB) to ribose-O3'
G351R	$G208^{++}$	♂	−/++				May displace N-terminal extension of eALAS destabilizing enzyme; distance to PLP is 25 Å
T388S	$T245^{++}$	♂	−−/++	+			PLP affinity reduced by weakening HB to PO_4 group
C395Y	$V252^{+}$	♀	−−/++	+			Larger size disrupts Thr388 and Lys391 affecting PLP binding and catalysis
G398D	$G255^{++}$	♀	−−/−	+			Abrogates PLP binding by filling PLP–PO_4^{-} site
R411C	$R268^{++}$	♀♂	−−/+			(+)	Loss of SB to D408 may affect Thr420 and Thr421 in loop α9–α10 involved in PLP binding
R411H		♀	−/+				
G416D	$G273^{++}$	♀	−/+	+		+	Disrupts Phe419 (sCoA pantetheine binding) and Thr420 and Thr421 involved in PLP–PO_4^{-} binding
M426V	$A283^{-}$	♂	−−/++	(+)			Branched valine may affect neighboring Gly220 and Ala221 of glycine-rich stretch, affecting PLP binding
R436W	$A293^{-}$	♀	−−/−				Residues constitute a positively charged cluster on the surface of eALAS remote from active site; potentially involved in binding sCoA synthetase (Furuyama and Sassa, 2000)
R448Q	$D305^{-}$	♂	−/−+				
R452H	$M309^{-}$	♀♂	−(−)/−+				
R452C		♀	−/+				
R452S		♂	−/+				
R458H	$K315^{+}$	♂	?/?				
I476N	$V333^{+}$	♂	−−/++	+	+		Displaces His360 (PLP binding) and Arg517 (glycine recognition); hydrophilic residue in hydrophobic pocket
Y506–fs	$F363^{+}$	♀	−−/−		+	+	Frame shift disrupts C-terminal domain, catalysis abolished; sCoA and glycine binding severely affected
T508S	$T365^{++}$	♂	−/?			+	HB to sCoA carboxylate group weakened, residue required for closed enzyme conformation
R517C	$R374^{++}$	♀	−−/−		+		Arg517 is essential for glycine recognition, loss of substrate specificity
H524D	$H381^{++}$	♂	−−/−				Stacking of His524 (C-term. domain) and Trp194 of N-term. domain lost; contacts to catalytic domain disturbed
R559H	n/a	♀	?/?				n/a — not present in model
R560H	n/a	♂	−−/+				n/a — not present in model
S568G	n/a	♂	(−)/−				n/a — not present in model

♀/♂: observed in females/males.
[a]Conserved^{++}, similar^{+}, not conserved^{-} residue in $ALAS_{Rc}$ compared to eALAS (colors as in Figure 5).
[b]Severe symptoms −−, moderately severe −.
[c]Completely responsive ++, partially responsive +, not responsive −, unknown ? (according to Bottomley, 2004).
[d]For references, please refer to extended Table II in Supplementary data.

helical turn away from His285 (His142) involved in PLP binding (Figure 2). Disruption of α-helix α5 will thus affect PLP binding as well. As a result, this mutation may be counteracted by increasing levels of PLP.

C395Y: Cys395 (Val252) is spatially immediately adjacent to Leu389 (Leu246), which in turn is located sequentially between Thr388 (Thr245) and Lys391 (Lys248), respectively, involved in PLP–phosphate binding and in catalysis. Replacing Cys395 by significantly larger tyrosine will adversely affect PLP binding. Correspondingly, the activity of this mutant enzyme may be rescued by higher concentrations of PLP.

G398D: Gly398 (Gly255) is wedged in between Cys258 (A115) and Thr388 (Th245), itself involved in an XLSA

mutation and involved in PLP–phosphate binding. Adding a side chain by mutation to aspartate will sterically disrupt this pocket. By additionally introducing a negative charge, the negative phosphate will no longer bind. Correspondingly, pyridoxine treatment is ineffectual in this case.

I476N: The side chain of Ile476 (Val333) stabilizes that of His360 (His217), involved in PLP binding, and Arg517 (Arg374), required for glycine recognition. Replacing this residue by asparagine will thus likely affect both PLP and glycine binding. Patients with this mutation therefore respond well to pyridoxine.

Y199H: Like Ile476 (Val333), above, Tyr199 (Tyr56) stabilizes His360 and Arg517, as well as the catalytic Lys391 (Lys248) and Asn197 (Asn54), involved in glycine binding. Replacing tyrosine by histidine, though frequently a tolerable mutation, appears to destabilize PLP and glycine binding sufficiently, so that elevated intracellular levels of PLP only partially offset the mutation.

R227C: Arg227 (Arg84), located in the glycine-rich stretch and immediately adjacent to Thr226 (Thr83), itself involved in substrate glycine and sCoA carboxylate binding, is stabilized by a buried salt bridge to Glu240 (Glu97). Replacing Arg227 by cysteine abrogates this ionic interaction. Apparently, this destabilizes the active site to such an extent that PLP supplementation proves ineffectual.

G416D: Gly416 (Gly273) is one α-helical turn removed and thus physically adjacent to Phe419 (Phe276), which is crucial for interaction with the pantetheine moiety of CoA. Gly416 is also physically close to Thr420 (Ser277) and Thr421 (Thr278), both of which bind the PLP–phosphate. Substitution of Gly416 by aspartate would hence disrupt both PLP and sCoA binding. Surprisingly, the disruption is not severe, as PLP does impart a moderate improvement in ALAS activity.

Buried XLSA mutations physically distant from the active site (green in Figure 5A and B and Table II). Five mutational sites (accounting for eight described mutations) are quite distant with respect to the active site-, cofactor-, and substrate-binding pockets. Here, the influence of a substitution is less direct. Instead, a local change may radiate throughout the structure, influencing protein stability, substrate-binding affinity, or enzymatic catalysis. Buried residues are particularly sensitive in this regard, normally permitting only conservative substitutions if complete destabilization and hence inactivation is to be prevented.

Due to the large number of mutations, the indirect effect of the mutations, and hence the partly speculative nature of the structural explanation, the discussion of these mutations is provided in the Supplementary data.

XLSA mutations on the protein surface (highlighted in blue in Figure 5A and B and Table II). In all, 16 mutations affect the 12 separate surface-exposed residues in ALAS not located near the active-site channel. These residues are exposed to the surrounding aqueous medium and thus tend to be hydrophilic. Replacing them with hydrophobic residues or by a residue of opposite charge will generally not directly affect the binding of substrates or of cofactors. Instead, they may lead to protein destabilization and/or aggregation, affecting the amount of protein available to produce ALA. Alternatively, surface-exposed residues may affect the interaction of ALAS with other proteins. ALAS has, for example, previously been reported to bind the β-subunit of ATP-specific sCoA synthetase, the enzyme that generates sCoA for eALAS (Furuyama and Sassa, 2000). The replacement of Asp190 by valine, an XLSA-inducing mutant (Table II), was shown to prevent this interaction. At least six further mutations, affecting four positively charged amino acids (R436W, R448Q, R452C/H/S, R458H), cluster near Asp190, indicating that they, too, may induce XLSA by inhibiting the interaction of eALAS with sCoA synthatase.

The detailed discussion of each individual mutation is provided as Supplementary data to this publication.

Conclusions

The structure of ALAS from *R. capsulatus* is the first structure of this essential enzyme from tetrapyrrole biosynthesis. Coincidentally thus, the first enzyme required for heme biosynthesis in humans, animals, and α-proteobacteria is the last to be elucidated structurally. The high degree of conservation of this enzyme from bacteria to humans underlines the importance of this enzyme, the low evolutionary tolerance of changes, and allows a surprisingly detailed view of human eALAS. This will aid the understanding of the disease and allow the efficacy of pyridoxine treatment to be rationalized. It may furthermore allow sufferers of the disease and medical personnel involved in its treatment to lift the veil of uncertainty by providing firm structural evidence to understand the physical basis for the symptoms.

Materials and methods

Cloning, production, and purification of ALAS from R. capsulatus, activity assay

The *hemA* gene of *R. capsulatus* (DSM 1710) was amplified from genomic DNA using PCR and cloned into the *Bam*HI/*Hin*dIII sites of the *Escherichia coli* expression vector pQE30 (Qiagen). The protein was produced in *E. coli* M15 (pRep4) (Qiagen) cells at 20°C in TB medium after induction with 0.1 mM isopropylthio-β-D-galactoside. Cells were harvested by centrifugation, disrupted by French press, and the insoluble membrane and cell wall fraction were discarded following centrifugation. $ALAS_{Rc}$ was isolated from the soluble fraction by Ni-NTA affinity chromatography (Ni-NTA Agarose, Qiagen), followed by anion exchange chromatography (MonoQ, Amersham Biosciences) and gel filtration (Superdex 200, Amersham Biosciences). The purified protein was concentrated to 10 mg/ml. Protein integrity and purity were verified using SDS–PAGE, N-terminal sequencing, dynamic light scattering, and mass-spectrometric analysis.

The human *ALAS2* gene encoding eALAS was amplified by PCR to yield a truncated (135 amino acids N-terminal and 34 amino acids at the C-terminus) and N-terminally His-tagged protein. Expression and purification procedures were as described for the $ALAS_{Rc}$.

Enzymatic activity was analyzed by spectrophotometrically monitoring the increase of NADH ($\lambda = 340$ nm) produced by α-ketoglutarate dehydrogenase during the regeneration of sCoA consumed by ALAS (Hunter and Ferreira, 1995).

Crystallization

$ALAS_{Rc}$ was crystallized by hanging drop vapor diffusion at 20°C under anaerobic conditions (MACS1000 anaerobic workbench, Don Whitley Scientific). In all, 2 μl of protein in 20 mM Hepes, pH 7.3, 200 mM NaCl, 1 mM DTT, and 1 μM PLP was added to 2 μl of reservoir solution containing 100 mM Hepes, pH 7.5, 200 mM Na acetate, 8% isopropanol (v/v), and 20% PEG 4000. Rectangular crystals grew to a size of $50 \times 50 \times 50\,\mu m^3$ within 2 weeks. The crystals belong to space group $P2_12_12_1$ with cell constants $a = 67.3$ Å, $b = 90.6$ Å, and $c = 245.6$ Å. A Matthews coefficient of 2.24 $Å^3$/Da indicated the presence of two ALAS dimers per

asymmetric unit, corresponding to a solvent content of 44% (Matthews, 1968). For the substrate-bound structures, glycine or sCoA sodium salt was added to the reservoir solution (final concentration 10 mM) and the crystals were transferred into 4 µl of this solution and soaked for 24 h.

Data collection and structure determination

X-ray diffraction data of substrate-free and substrate-soaked crystals were, respectively, collected at the beamline BW6 (DESY, Hamburg) and beamline BL1 (BESSY II, Berlin). Data were processed and scaled using the HKL suite (Otwinowski and Minor, 1997). All further steps were performed using programs of the CCP4 suite (CCP4, 1994). The superimposed structures of AONS (PDB code 1BS0; Alexeev *et al*, 1998) and KBL (PDB code 1FC4; Schmidt *et al*, 2001) were used as a model for molecular replacement in EPMR (Kissinger *et al*, 1999). CNS (Brunger *et al*, 1998) was used for the rigid-body and simulated-annealing refinement of all three structures. REFMAC5 was used for subsequent refinement, including NCS-restraint and TLS-refinement protocols (Murshudov *et al*, 1997). The program O was used for manual model building and structural analysis (Jones *et al*, 1991). Structures were validated using PROCHECK (Laskowski *et al*, 1993) and WHAT IF (Vriend, 1990). Molecular depictions were prepared using PyMOL (DeLano, 2002). Sequence alignments were performed with ClustalW (Thompson *et al*, 1994), structure-based sequence alignments using DALI (Holm and Sander, 1996).

Accession number

The coordinates of the structures have been deposited in the Protein Data Bank (Code 2BWN, 2BWO and 2BWP).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

We thank S Alves, B Voedisch, S Gläsener, U Widow, G Layer, and D Gebauer for help in cloning, protein purification, and crystallization. We are indebted to Professor S Bottomley, for kindly sharing her list of all XLSA mutations with us. We gratefully acknowledge synchrotron beam time at BW6, DESY, Hamburg, Germany (substrate-free ALAS) and BL1, BESSY, Berlin, Germany (glycine-bound and succinyl-CoA-bound ALAS). This work was funded by the Deutsche Forschungsgemeinschaft (DFG).

References

Alexeev D, Alexeeva M, Baxter RL, Campopiano DJ, Webster SP, Sawyer L (1998) The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *J Mol Biol* **284:** 401–419

Andrews NC (1999) Disorders of iron metabolism. *N Engl J Med* **341:** 1986–1995

Bolt EL, Kryszak L, Zeilstra-Ryalls J, Shoolingin-Jordan PM, Warren MJ (1999) Characterization of the *Rhodobacter* sphaeroides 5-aminolaevulinic acid synthase isoenzymes, HemA and HemT, isolated from recombinant *Escherichia coli*. *Eur J Biochem* **265:** 290–299

Bottomley SS (2004) Sideroblastic anemias. In *Wintrobe's Clinical Hematology*, Greer J, Foerster J, Lukens JN, Rodgers GM, Paraskevas R, Glader Bertil (eds), pp 1012–1033. Philadelphia: Lippincott Williams & Wilkins

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54** (Part 5)**:** 905–921

CCP4, CCP, Number 4 (1994) The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr D* **50:** 760–763

Cheltsov AV, Guida WC, Ferreira GC (2003) Circular permutation of 5-aminolevulinate synthase: effect on folding, conformational stability, and structure. *J Biol Chem* **278:** 27945–27955

Christen P, Mehta PK (2001) From cofactor to enzymes. The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. *Chem Rec* **1:** 436–447

Conboy JG, Cox TC, Bottomley SS, Bawden MJ, May BK (1992) Human erythroid 5-aminolevulinate synthase. Gene structure and species-specific differences in alternative RNA splicing. *J Biol Chem* **267:** 18753–18758

Dailey HA (1997) Enzymes of heme biosynthesis. *J Biol Inorg Chem* **2:** 411–417

DeLano WL (2002) The PyMOL Molecular Graphics System. (www.pymol.org)

Ferreira GC, Neame PJ, Dailey HA (1993) Heme biosynthesis in mammalian systems: evidence of a Schiff base linkage between the pyridoxal 5′-phosphate cofactor and a lysine residue in 5-aminolevulinate synthase. *Prot Sci* **2:** 1959–1965

Frankenberg N, Moser J, Jahn D (2003) Bacterial heme biosynthesis and its biotechnological application. *Appl Microbiol Biotechnol* **63:** 115–127

Furuyama K, Sassa S (2000) Interaction between succinyl CoA synthetase and the heme-biosynthetic enzyme ALAS-E is disrupted in sideroblastic anemia. *J Clin Invest* **105:** 757–764

Furuyama K, Sassa S (2002) Multiple mechanisms for hereditary sideroblastic anemia. *Cell Mol Biol (Noisy-le-grand)* **48:** 5–10

Gibson KD, Laver WG, Neuberger A (1958) Initial stages in the biosynthesis of porphyrins. 2. The formation of delta-aminolaevulic acid from glycine and succinyl-coenzyme A by particles from chicken erythrocytes. *Biochem J* **70:** 71–81

Gong J, Kay CJ, Barber MJ, Ferreira GC (1996) Mutations at a glycine loop in aminolevulinate synthase affect pyridoxal phosphate cofactor binding and catalysis. *Biochemistry* **35:** 14109–14117

Goodfellow BJ, Dias JS, Ferreira GC, Henklein P, Wray V, Macedo AL (2001) The solution structure and heme binding of the presequence of murine 5-aminolevulinate synthase. *FEBS Lett* **505:** 325–331

Grishin NV, Phillips MA, Goldsmith EJ (1995) Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Prot Sci* **4:** 1291–1304

Harigae H, Nakajima O, Suwabe N, Yokoyama H, Furuyama K, Sasaki T, Kaku M, Yamamoto M, Sassa S (2003) Aberrant iron accumulation and oxidized status of erythroid-specific delta-aminolevulinate synthase (ALAS2)-deficient definitive erythroblasts. *Blood* **101:** 1188–1193

Hennig M, Grimm B, Contestabile R, John RA, Jansonius JN (1997) Crystal structure of glutamate-1-semialdehyde aminomutase: an alpha 2-dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. *Proc Natl Acad Sci USA* **94:** 4866–4871

Holm L, Sander C (1996) Mapping the protein universe. *Science* **273:** 595–603

Hungerer C, Weiss DS, Thauer RK, Jahn D (1996) The hemA gene encoding glutamyl-tRNA reductase from the archaeon *Methanobacterium thermoautotrophicum* strain Marburg. *Bioorg Med Chem* **4:** 1089–1095

Hunter GA, Ferreira GC (1995) A continuous spectrophotometric assay for 5-aminolevulinate synthase that utilizes substrate cycling. *Anal Biochem* **226:** 221–224

Hunter GA, Ferreira GC (1999a) Lysine-313 of 5-aminolevulinate synthase acts as a general base during formation of the quinonoid reaction intermediates. *Biochemistry* **38:** 12526

Hunter GA, Ferreira GC (1999b) Pre-steady-state reaction of 5-aminolevulinate synthase. Evidence for a rate-determining product release. *J Biol Chem* **274:** 12222–12228

Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr Sect A* **2:** 110–119

Jordan PM (1990) Biosynthesis of 5-aminolevulinic acid and its transformation into coproporphyrinogen in animals and bacteria. In *Biosynthesis of Heme and Chlorophylls*, Dailey HA (ed), pp 55–121. New York: McGraw-Hill Publishing Company

Kissinger CR, Gehlhaar DK, Fogel DB (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* **55:** 484–491

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26:** 283–291

Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* **33:** 491–497

May A, Bishop DF (1998) The molecular biology and pyridoxine responsiveness of X-linked sideroblastic anaemia. *Haematologica* **83:** 56–70

Mehta PK, Christen P (1994) Homology of 1-aminocyclopropane-1-carboxylate synthase, 8-amino-7-oxononanoate synthase, 2-amino-6-caprolactam racemase, 2,2-dialkylglycine decarboxylase, glutamate-1-semialdehyde 2,1-aminomutase and isopenicillin-*N*-epimerase with aminotransferases. *Biochem Biophys Res Commun* **198:** 138–143

Mehta PK, Christen P (2000) The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol* **74:** 129–184

Moser J, Lorenz S, Hubschwerlen C, Rompf A, Jahn D (1999) Methanopyrus kandleri glutamyl-tRNA reductase. *J Biol Chem* **274:** 30679–30685

Moser J, Schubert WD, Beier V, Bringemeier I, Jahn D, Heinz DW (2001) V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *EMBO J* **20:** 6583–6590

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D Biol Crystallogr* **53:** 240–255

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol Crystallogr PT A Methods Enzymol* **276:** 307–326

Riddle RD, Yamamoto M, Engel JD (1989) Expression of delta-aminolevulinate synthase in avian cells: separate genes encode erythroid-specific and nonspecific isozymes. *Proc Natl Acad Sci USA* **86:** 792–796

Schmidt A, Sivaraman J, Li Y, Larocque R, Barbosa JA, Smith C, Matte A, Schrag JD, Cygler M (2001) Three-dimensional structure of 2-amino-3-ketobutyrate CoA ligase from *Escherichia coli* complexed with a PLP-substrate intermediate: inferred reaction mechanism. *Biochemistry* **40:** 5151–5160

Schneider G, Kack H, Lindqvist Y (2000) The manifold of vitamin B6 dependent enzymes. *Struct Fold Des* **8:** R1–R6

Shemin D, Kikuchi G (1958) Enzymatic synthesis of sigma-aminolevulinic acid. *Ann N Y Acad Sci* **75:** 122–128

Shoolingin-Jordan PM, Al-Daihan S, Alexeev D, Baxter RL, Bottomley SS, Kahari ID, Roy I, Sarwar M, Sawyer L, Wang SF (2003) 5-Aminolevulinic acid synthase: mechanism, mutations and medicine. *Biochim Biophys Acta* **1647:** 361–366

Shoolingin-Jordan PM, LeLean JE, Lloyd AJ (1997) Continuous coupled assay for 5-aminolevulinate synthase. *Methods Enzymol* **281:** 309–316

Tan D, Harrison T, Hunter GA, Ferreira GC (1998) Role of arginine 439 in substrate binding of 5-aminolevulinate synthase. *Biochemistry* **37:** 1478–1484

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673–4680

Vriend G (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph* **8:** 52–56, 29

Zhang J, Ferreira GC (2002) Transient state kinetic investigation of 5-aminolevulinate synthase reaction mechanism. *J Biol Chem* **277:** 44660–44669