Schulze, J.O., Schubert, W.-D., Moser, J., Jahn, D., Heinz, D.W.
Evolutionary Relationship between Initial Enzymes of Tetrapyrrole Biosynthesis
(2006) Journal of Molecular Biology, 358 (5), pp. 1212-1220.

Evolutionary Relationship between Initial Enzymes of Tetrapyrrole Biosynthesis

Jörg O. Schulze[1], Wolf-Dieter Schubert[1], Jürgen Moser[2], Dieter Jahn[2], Dirk W. Heinz[1,*]

[1]Division of Structural Biology, German Research Centre for Biotechnology (GBF), Mascheroder Weg 1, D-38124 Braunschweig, Germany

[2]Institute of Microbiology, Technical University Braunschweig, Spielmannstr. 7, D-38106 Braunschweig, Germany

*Corresponding author; E-mail: dih@gbf.de

Running title: Initial Enzymes of Tetrapyrrole Biosynthesis

Summary

Glutamate-1-semialdehyde 2,1-aminomutase (GSAM) is the second enzyme in the C_5 pathway of tetrapyrrole biosynthesis found in most bacteria, in archaea and in plants. It catalyzes the transamination of glutamate-1-semialdehyde to 5-aminolevulinic acid (ALA) in a pyridoxal 5'-phosphate-(PLP)-dependent manner. We present the crystal structure of GSAM from the thermophilic cyanobacterium *Thermosynechococcus elongatus* ($GSAM_{Tel}$) in its PLP-bound form at 2.8 Å resolution. $GSAM_{Tel}$ is a symmetric homodimer, whereas GSAM from *Synechococcus* ($GSAM_{Syn}$) was previously described as asymmetric. The symmetry of $GSAM_{Tel}$ thus challenges the previously proposed negative cooperativity between monomers of this enzyme. Furthermore, $GSAM_{Tel}$ reveals an extensive flexible region at the interface of the proposed complex of GSAM with glutamyl-tRNA reductase (GluTR), the preceding enzyme in tetrapyrrole biosynthesis. Compared to $GSAM_{Syn}$, the monomers of $GSAM_{Tel}$ are rotated away from each other along the dimerization interface by 10°. The associated flexibility of GSAM may be essential for complex formation with GluTR to occur.

Unexpectedly, we find that GSAM is structurally related to 5-aminolevulinate synthase (ALAS), the ALA producing enzyme in the Shemin pathway of α-proteobacteria and non-plant eukaryotes. This structural relationship also applies to the corresponding subfamilies of PLP-dependent enzymes. We thus propose that the CoA-subfamily (including ALAS) and the aminotransferase subfamily II (including GSAM) are evolutionarily closely related and that ALAS may thus have evolved from GSAM.

215 words (not more than 300 words)

Keywords: glutamate-1-semialdehyde 2,1-aminomutase, 5-aminolevulinate synthase, crystal structure, heme biosynthesis, aminotransferase

Introduction

The structural core of all tetrapyrroles, including essential cofactors such as heme, chlorophyll and vitamin B_{12}, is assembled from eight molecules of 5-aminolevulinic acid (ALA), the first common precursor of all tetrapyrroles. ALA was initially found to be synthesized from glycine and succinyl-coenzyme A by 5-aminolevulinate synthase (ALAS, EC 2.3.1.37) as part of the Shemin pathway.[1; 2] Much later the enzyme ALAS was found to be restricted to α-proteobacteria and non-plant eukaryotes. All other bacteria, archaea and plants produce ALA through the C_5 pathway using tRNA-bound glutamate as substrate.[3; 4] The thereby activated glutamate is first reduced to glutamate-1-semialdehyde (GSA) by the NADPH-dependent glutamyl-tRNA reductase (GluTR, EC 1.2.1.70)[5], and GSA is then rearranged to ALA by glutamate-1-semialdehyde 2,1-aminomutase (GSAM; EC 5.4.3.8)[6]. As animals and humans lack these enzymes, they are potential targets for herbicides and antibacterial agents.

GSAM, a homodimer, is a member of the α family of PLP-dependent enzymes.[7] Although it functions as a mutase, catalyzing the exchange of amino and oxo groups in GSA, GSAM is structurally[8] and mechanistically[6; 9; 10] an aminotransferase. One proposed reaction mechanism requires the cofactor to initially be present as pyridoxamine 5'-phosphate (PMP).[9; 10] Transfer of the amino group to atom C_5 of GSA produces the intermediate 4,5-diaminovalerate (DAVA) and PLP (Figure 1), while abstraction of the C_4 amino group finally yields ALA and regenerates PMP. An alternative mechanism starts with the PLP form of the enzyme and proceeds via a PMP-bound enzyme and 4,5-dioxovalerate (DOVA) as intermediate.[6] All analyzed GSAM enzymes were found capable of catalyzing both reactions, however, kinetic studies indicate a preference for the first described mechanism in *Synechococcus*.[9; 10]

Fig. 1

GSA, a highly reactive α-amino aldehyde, has a half-life of ~3 min under physiological conditions.[11] Based on the crystal structures of GluTR from *Methanopyrus kandleri*[12] and GSAM from *Synechococcus* ($GSAM_{Syn}$)[8] we have previously proposed that both dimeric enzymes form a complex allowing GSA to channel directly from GluTR to GSAM ensuring

efficient synthesis of ALA and minimizing the loss of this unstable intermediate.[12] We and others have confirmed the existence of this complex for the enzymes from *Escherichia coli*[13] and *Chlamydomonas reinhardtii*, respectively[14]

Here we present the crystal structure of GSAM from *Thermosynechococcus elongatus* ($GSAM_{Tel}$) in its PLP-bound form at 2.8 Å resolution. In contrast to the asymmetric homodimer from *Synechococcus* ($GSAM_{Syn}$), $GSAM_{Tel}$ is perfectly symmetric. An extensive flexible region covering the active site is located at the proposed binding interface with GluTR. Based on the unexpected structural similarity of $GSAM_{Tel}$ to 5-aminolevulinate synthase (ALAS), we propose an evolutionary relationship between the ALA-generating enzymes of the independent C_5 and Shemin pathways.

Results

Overall structure

GSAM from the thermophilic cyanobacterium *Thermosynechococcus elongatus* ($GSAM_{Tel}$) was cloned, produced, purified and crystallized as described in Materials and Methods. The structure was solved by molecular replacement using the structure of GSAM from *Synechococcus* ($GSAM_{Syn}$)[8] (PDB-code 2GSA) and refined to a resolution of 2.8 Å. Data collection and refinement statistics are summarized in Table 1. In agreement with previous studies[15], dynamic light scattering and analytical gel filtration confirm $GSAM_{Tel}$ to be a structural dimer. Crystal packing analyses revealed only one $GSAM_{Tel}$ monomer (415 residues) per asymmetric unit. Two crystallographically related monomers thus constitute a symmetric homodimer for $GSAM_{Tel}$ - compared to the asymmetric $GSAM_{Syn}$ dimer[8].

Fig. 2(a) Each GSAM monomer consists of three domains (Figure 2a): a small N-terminal domain (residues 1-48) encompassing a three-stranded, antiparallel β-sheet; a catalytic domain (residues 49-304) dominated by a seven-stranded β-sheet (β-strand order 3245671 where only strand 7 is antiparallel) flanked on either side by α-helices; and a C-terminal domain (residues 305-411) including a three-stranded antiparallel β-sheet and four α-helices. The symmetrically positioned active sites, identified by the bound cofactor PLP, are located at the dimer interface.

A root-mean-square deviation of 0.7 Å for 360 common C_α atoms underscores the similarity (77% sequence identity) of $GSAM_{Tel}$ and $GSAM_{Syn}$[8]. The topology of GSAM is, furthermore, clearly related to that of other members of the AT II subfamily of PLP-dependent enzymes such as 4-aminobutyrate aminotransferase[16] and dialkylglycine decarboxylase[17].

Cofactor binding

UV-visible absorption spectroscopy of $GSAM_{Tel}$ indicates that the cofactor is lost during purification. Although both PMP and PLP had been added to the protein in sufficient amounts

to allow active-site saturation prior to crystallization, only PLP is observed within the active site. PLP is covalently linked to the catalytic residue Lys251 constituting the "internal aldimine" (Figure 2(b)). Correspondingly, the pyridinium ring of PLP adopts an orientation characteristic for lysine-bound PLP. Bound PMP, in contrast, is usually tilted by 20-30° moving the amino group away from the catalytic lysine as observed in the PMP-containing subunit of $GSAM_{Syn}$ and aspartate aminotransferase[18]. The phosphate group of PLP is recognized via hydrogen bonds to Gly101-N, Thr102-N, Thr102-O_γ and to a water molecule. The pyridinium ring of PLP is sandwiched between Val225 and Tyr128 in such a way that the phenoxy ring of Tyr128 is almost perpendicular to the pyridinium ring (Figure 2(b)), a distinctive feature of GSAM and closely related aminotransferases[16; 19]. A salt bridge to atom $O_{\delta 2}$ of Asp223, an invariant residue in all aminotransferases[20; 21], ensures that PLP-N_1 remains protonated strengthening the electron withdrawing capacity of the cofactor[22; 23]. Residues from the second monomer that normally participate in PLP-binding[24] are disordered in $GSAM_{Tel}$ (see below).

Fig. 2b

Loops covering the active sites are flexible in the $GSAM_{Tel}$ dimer

Discontinuous electron density for three mutually adjacent loops (residues 1-18, 132-159 and 278–282) indicates that they are disordered in the $GSAM_{Tel}$ crystal. Excluding these loops from the refined structure results in a deep cleft at the dimer interface exposing both active sites (Figure 2(c)). One of the loops (132-159) was similarly found to be disordered in subunit B of $GSAM_{Syn}$ (153-181) while it is ordered in subunit A. In addition, the disordered 18 N-terminal residues of $GSAM_{Tel}$ correspond to residues 30-42 of $GSAM_{Syn}$, observed to show different backbone conformations in subunits A and B[8]. Val31 and Arg32 of $GSAM_{Syn}$ delimit the substrate-binding pocket. Arg32, an invariant residue in GSAM from all species, presumably binds the carboxylate group of the substrate GSA.[8]

Fig. 2c

Quaternary structures of GSAM$_{Tel}$ and GSAM$_{Syn}$ differ

The r.m.s. deviation GSAM$_{Tel}$- and GSAM$_{Syn}$-monomers following their superposition is 0.7 Å (see above). The corresponding value for the dimers is 2.1 Å. This noticeably larger value for the dimer is caused by an "opening up" by about 10° of the GSAM$_{Tel}$ dimer relative to GSAM$_{Syn}$ (Figure 2(d)). While the separation and mutual orientation of α-helices α2 and α2* (* for the second monomer), located near the two-fold dimer axis, are quite similar in both proteins, adjoining regions are progressively further apart in moving away from the fulcrum positioned near α2 and α2*. The largest difference of 6 Å between GSAM$_{Tel}$ than GSAM$_{Syn}$ is observed for residues located in α-helices α5 and α5*.

Fig. 2(d)

Discussion

GSAM is an unusually flexible aminotransferase

GSAM$_{Tel}$ is a symmetric, PLP-binding dimer with equally disordered regions in both subunits. In the asymmetric GSAM$_{Syn}$ the PLP-binding monomer B has a similarly disordered lid[25], while the lid is closed and well-ordered in PMP-binding monomer A.[8] This observation of PLP in combination with a disordered or "open" lid is unexpected as the catalyzed reaction for GSAM$_{Syn}$ is believed to be initiated by PMP[9; 10] (Figure 1), while the closed, PMP-bound state would interfere both with substrate uptake and product release. In the PLP-bound form, corresponding to a reaction intermediate, GSAM$_{Syn}$ adopts an "open" conformation allowing for the potential loss of the product-intermediate 4,5-diaminovalerate arresting the enzyme in its PLP form. Monomers were suggested to show negative cooperativity[8], ensuring that one monomer remains in the closed and PMP-bound conformation while the other is open and binds PLP. The symmetric GSAM$_{Tel}$, however, challenges the proposed negative cooperativity as both monomers can clearly adopt the same state simultaneously. This is supported by the observation that GSAM$_{Tel}$ exclusively binds PLP, despite PMP and PLP

having been added. PLP may have a higher binding affinity than PMP because of Schiff base formation, whereas PMP does not bind covalently. This finding supports an alternative mechanism suggested for GSAM, where catalysis starts with the enzyme in its PLP form abstracting the amino-group of GSA and yielding 4,5-dioxovalerate as product intermediate (Figure 1)[6]. Only crystals of intermediate or inhibitor complexes, or further spectroscopic data would allow the true mechanism to be established.

The asymmetry of $GSAM_{Syn}$, the symmetry of the closely related $GSAM_{Tel}$, and the open conformation of the latter, indicate that uncomplexed GSAM is unusually flexible compared to other PLP-dependent enzymes.

The flexible region is located in the putative GluTR binding interface

In a proposed model of a GluTR/GSAM complex, we placed $GSAM_{Syn}$[8] into the V-shaped GluTR-dimer from *Methanopyrus kandleri*[12] *inter alia* to allow for substrate channeling between the two enzymes of the C_5 pathway[13; 14]. Replacing $GSAM_{Syn}$ by $GSAM_{Tel}$ in this model, results in the flexible lids of $GSAM_{Tel}$ being positioned at the interface with GluTR (Figure 3) opposite the "back door" of GluTR[12]. GluTR would therefore potentially stabilize the flexible loops of GSAM and establish a substrate channel that enables efficient transfer of the reactive intermediate GSA between the active sites of both enzymes. Complex formation of GSAM with GluTR may also prevent loss of the proposed reaction intermediate 4,5-diaminovalerate, as is observed to an significant extent in *in vitro* analyses of isolated GSAM[26].

Fig. 3

The open conformation of $GSAM_{Tel}$ allows it to be accommodated within the V-shaped GluTR with less steric hindrance than $GSAM_{Syn}$ implying that flexibility of GSAM is a necessary feature for complex formation. Instead of acting as an active-site lid, the flexible regions of GSAM would facilitate complex formation with GluTR creating the connecting substrate channel in the process.

Structural homology of GSAM and ALAS

We recently reported the crystal structure of 5-aminolevulinate synthase (ALAS)[27] from *Rhodobacter capsulatus* (Figure 4(a)). ALAS, like GSAM, produces ALA. It utilizes the substrates glycine and succinyl-coenzyme A rather than glutamate-1-semialdehyde (GSA). This route constitutes the Shemin pathway of ALA-synthesis found in non-plant eukaryotes and α-proteobacteria. Mechanistically, ALAS belongs to the CoA subfamily of the α family of PLP-dependent enzymes[7] and is hence structurally closely related to 2-amino-3-ketobutyrate CoA ligase (KBL)[28] and 8-amino-7-oxononanoate synthase (AONS)[29], two members of this subfamily (Table 2). Despite a sequence identity of only 18%, GSAM is the next closest structural neighbor of ALAS (Figure 4(b)). This is unexpected, as GSAM belongs to the subfamily AT II of the α family, proposed to be only distantly related to the CoA subfamily[7]. The overall structural similarity of GSAM and ALAS includes the substrate-binding pockets where both enzymes share a conserved arginine residue required for substrate binding (Arg21 in ALAS, Arg32 in $GSAM_{Syn}$, Arg10 in $GSAM_{Tel}$). Differences are restricted to the active sites and substrate channels, and include a significantly shorter loop near the active site compared to the long, flexible lid in GSAM (residues 132-159)[8; 25]. As a result ALAS possesses a distinct active-site channel and a hydrophobic binding pocket for the 3'-phosphate ADP moiety of succinyl-CoA. The N-terminal stretches (residues 1-23 in ALAS and 1-18 in GSAM) also deviate significantly. In ALAS an α-helix wraps around the second subunit stabilizing the interface, while in GSAM α-helix α1 folds back onto the same subunit[8].

Fig. 4a

Table 2

Fig. 4b

CoA- and AT II-subfamilies of PLP-dependent enzymes are evolutionary neighbors

PLP-dependent enzymes of four structurally unrelated families have been identified[30]. Elaborate evolutionary pedigrees for each family have been proposed based on family profile analysis and protein structure.[7] Within the α family, CoA-utilizing enzymes constitute the CoA subfamily and were found to be most closely related to the γ subfamily and more

Fig. 4c

distantly to subfamilies AT I and IV (Figure 4(c)).[7] The AT II subfamily, of which GSAM is a member, was found to be more distantly related to the CoA subfamily.

The high structural similarity of ALAS (CoA subfamily) and GSAM (AT II subfamily) implies a close evolutionary relationship that does not fit the above classification. Strikingly, both enzymes synthesize the identical product as part of the two alternate pathways initiating tetrapyrrole biosynthesis. Structural comparison analyses (DALI Z-scores, Table 2) indicate that enzymes of the CoA subfamily are, in fact, more similar to AT II enzymes than to members of subfamilies γ, AT I or AT IV (Figure 4(d)-(f)). Structural analysis of AONS (CoA subfamily) identified dialkylglycine decarboxylase, an AT II-enzyme, as the closest structural and possibly evolutionary neighbor.[29] Later, 7,8-diaminopelargonic acid (DAPA) synthase (AT II) was found to be structurally similar to AONS (CoA), the preceding enzyme in biotin biosynthesis[21], again indicating a close evolutionary relationship.[31] All members of both subfamilies share a unique 3-stranded β-meander in the N-terminal domain[21] not present in enzymes of the other subfamilies.

Fig. 4d-f

In summary, we suggest that within the classification of PLP-dependent enzymes by Mehta & Christen[7], the CoA subfamily needs to be re-assigned to the branch of the evolutionary tree accommodating subfamily AT II (Figure 4(c)).

Fig. 4c

ALAS evolved from GSAM

ALAS and GSAM, enzymes of two related, yet independent pathways, synthesize the same product. As part of the Shemin pathway, ALAS is highly conserved between α-proteobacteria as well as unicellular and multicellular eukaryotes.[32] GSAM as part of the C_5 pathway is found in all plants, archaea and eubacteria, except α-proteobacteria. As a presumed remnant of the RNA world[33], the glutamyl-tRNA substrate of GluTR, the first enzyme of the C_5 pathway, indicates this pathway to be particularly ancient. Succinyl-coenzyme A, appearing late in the evolution of the Krebs cycle[34] (itself younger than the biosynthesis of tetrapyrroles[35]) would be much younger.

The C_5 pathway (including GSAM) must hence be significantly older than the Shemin pathway (ALAS). Assuming that the C_5 pathway evolved into the latter would require the conversion of GSAM into ALAS but would eliminate the enzyme GluTR of the C_5 pathway. It would furthermore decouple heme biosynthesis from protein biosynthesis, making heme biosynthesis dependent on the metabolically more central Krebs cycle instead. Only after the evolution of the Shemin pathway, will an α-proteobacterium have given rise to the eukaryotic mitochondrion, as propounded by the endosymbiotic theory[36]. Together with other genes of energy metabolism or cofactor biosynthesis[37], the gene for ALAS would have been transferred to the nuclear genome, requiring the unfolded enzyme in turn to be translocated across the outer and inner mitochondrial membranes before adopting its productive state in the mitochondrial lumen.

Materials and Methods

Cloning, production and purification of GSAM

The gene for GSAM (*hemL*) of *Thermosynechococcus elongatus* strain BP-1 was amplified from genomic DNA and cloned into the *Nco*I/*Bam*HI sites of the *E. coli* expression vector pETM30 to produce a C-terminal GST-fusion protein. The plasmid was transformed into *E. coli* BL21 CodonPlus cells (Stratagene), cultivated in LB medium containing 30 µg/ml kanamycin at 37°C to an OD_{600} of 0.6. The temperature was lowered to 20°C, protein production induced by 100 µM isopropyl-ß-D-thiogalactopyranosid and cells cultivated for 20 h. Cells were harvested by centrifugation, disrupted by French press and cell debris removed by centrifugation at 4°C. $GSAM_{Tel}$-GST was isolated from the soluble fraction by glutathion sepharose affinity chromatography and enzymatically cleaved with TEV-protease. $GSAM_{Tel}$ was purified by anion exchange (MonoQ HR 10/10, Amersham Biosciences) and gel filtration chromatography (Superdex 200 26/60, Amersham Biosciences). The purified protein in 20 mM Hepes pH 7.9, 20 mM NaCl and 10 mM DTT was concentrated to 10 mg/ml using a Vivaspin 20 centrifugal concentrator (Vivascience). The yield was ~25 mg of pure protein per liter of bacterial culture. Protein integrity and purity were verified using SDS-PAGE, N-terminal protein sequencing, dynamic light scattering and mass-spectrometry.

Crystallization

GSAM was crystallized by hanging drop vapor diffusion at 4°C. 3 µl of protein (containing 1 mM of both PMP and PLP) was added to 3 µl of reservoir solution (100 mM Pipes pH 6.5, 200 mM Mg sulfate and 17% PEG 3350). Crystals grew to a size of 400 x 150 x 150 μm^3 within 3 weeks. Prior to X-ray data collection 20 % (v/v) ethylene glycol was added for cryo-protection. Crystals belong to space group $P4_12_12$ with cell constants a = b = 109.5 Å and c = 95.4 Å. A V_M[38] of 3.26 $Å^3$/Da indicated the presence of one GSAM monomer per asymmetric

unit, corresponding to a solvent content of 62%. Data reduction in a space group of lower symmetry, e.g. $P4_1$ did not result in better data statistics or improved electron density.

Data collection, structure determination and analysis

X-ray diffraction data were collected using a RU-H3R rotating anode X-ray generator and an R-axis IV++ detector (MSC-Rigaku). Data were processed and scaled using the XDS program package[39]. $GSAM_{Syn}$ (PDB-code 2GSA, monomer A)[8] was used as a model for molecular replacement in EPMR[40]. Rigid body and simulated annealing refinement was performed using CNS[41], while REFMAC5 including TLS-refinement protocols[42] was used for subsequent refinement. Coot was used for manual model building and structural analysis[43], WHAT IF[44] and Procheck[45] for structure validation, and PyMOL[46] for molecular depictions. LSQKAB of the CCP4 program suite[47] was used for structural superpositions and ClustalW[48] for sequence alignments, and DALI[49] for structure-based sequence alignments.

Accession number

The coordinates of the structure have been deposited in the Protein Data Bank (Entry code: XXX) and will be released upon publication.

Acknowledgements

We thank Dr. Joop van den Heuvel for help in cloning and protein purification techniques. We are indebted to Dr. Günter Stier (EMBL, Heidelberg) for providing us with the pETM30 vector. This work was funded by the Deutsche Forschungsgemeinschaft (DFG) to D.W.H. (He 1852/5-3) and D.J. (Ja 470/7-3).

References

1. Shemin, D. & Kikuchi, G. (1958). Enzymatic synthesis of sigma-aminolevulinic acid. *Ann N Y Acad Sci* 75, 122-8.
2. Gibson, K. D., Laver, W. G. & 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.
3. Beale, S. I. & Castelfranco, P. A. (1973). 14 C incorporation from exogenous compounds into -aminolevulinic acid by greening cucumber cotyledons. *Biochem Biophys Res Commun* 52, 143-9.
4. Jahn, D., Verkamp, E. & Soll, D. (1992). Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends in Biochemical Sciences* 17, 215-218.
5. Moser, J., Lorenz, S., Hubschwerlen, C., Rompf, A. & Jahn, D. (1999). Methanopyrus kandleri glutamyl-tRNA reductase. *J. Biol. Chem.* 274, 30679-30685.
6. Ilag, L. L. & Jahn, D. (1992). Activity and spectroscopic properties of the Escherichia coli glutamate 1-semialdehyde aminotransferase and the putative active site mutant K265R. *Biochemistry* 31, 7143-51.
7. Mehta, P. K. & Christen, P. (2000). The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. In *Adv Enzymol Relat Areas Mol Biol.*, Vol. 74, pp. 129-184.
8. Hennig, M., Grimm, B., Contestabile, R., John, R. A. & Jansonius, J. N. (1997). Crystal structure of glutamate-1-semialdehyde aminomutase: An alpha 2-dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. *PNAS* 94, 4866-4871.
9. Smith, M. A., Kannangara, C. G., Grimm, B. & von Wettstein, D. (1991). Characterization of glutamate-1-semialdehyde aminotransferase of *Synechococcus*. Steady-state kinetic analysis. *Eur J Biochem* 202, 749-57.
10. Pugh, C., Harwood, J. & John, R. (1992). Mechanism of glutamate semialdehyde aminotransferase. Roles of diamino- and dioxo-intermediates in the synthesis of aminolevulinate. *J. Biol. Chem.* 267, 1584-1588.
11. Hoober, J. K., Kahn, A., Ash, D. E., Gough, S. & Kannangara, C. G. (1988). Biosynthesis of delta-aminolevulinate in greening barley leaves. IX. Structure of the substrate, mode of gabaculine inhibition, and the catalytic mechanism of glutamate 1-semialdehyde aminotransferase. *Carlsberg Res Commun* 53, 11-25.
12. Moser, J., Schubert, W. D., Beier, V., Bringemeier, I., Jahn, D. & Heinz, D. W. (2001). V-shaped structure of glutamyl-tRNA reductase, the first enzyme of tRNA-dependent tetrapyrrole biosynthesis. *Embo J* 20, 6583-90.
13. Luer, C., Schauer, S., Mobius, K., Schulze, J., Schubert, W. D., Heinz, D. W., Jahn, D. & Moser, J. (2005). Complex formation between glutamyl-tRNA reductase and glutamate-1-semialdehyde 2,1-aminomutase in Escherichia coli during the initial reactions of porphyrin biosynthesis. *J Biol Chem* 280, 18568-72.
14. Nogaj, L. A. & Beale, S. I. (2005). Physical and kinetic interactions between glutamyl-tRNA reductase and glutamate-1-semialdehyde aminotransferase of Chlamydomonas reinhardtii. *J Biol Chem* 280, 24301-7.
15. Ilag, L. L., Jahn, D., Eggertsson, G. & Soll, D. (1991). The Escherichia coli hemL gene encodes glutamate 1-semialdehyde aminotransferase. *J Bacteriol* 173, 3408-13.
16. Storici, P., Capitani, G., De Biase, D., Moser, M., John, R. A., Jansonius, J. N. & Schirmer, T. (1999). Crystal structure of GABA-aminotransferase, a target for antiepileptic drug therapy. *Biochemistry* 38, 8628-34.
17. Toney, M. D., Hohenester, E., Cowan, S. W. & Jansonius, J. N. (1993). Dialkylglycine decarboxylase structure: bifunctional active site and alkali metal sites. *Science* 261, 756-9.

18. Jansonius, J. N. & Vincent, M. G. (1987). Structural basis for catalysis by aspartate aminotransferase. In *Biological Macromolecules and Assemblies* (Jurnak, F. A. & McPherson, A., eds.), Vol. 3 Active Sites of Enzymes, pp. 187-285. Wiley & Sons, New York.
19. Toney, M. D., Hohenester, E., Keller, J. W. & Jansonius, J. N. (1995). Structural and Mechanistic Analysis of Two Refined Crystal Structures of the Pyridoxal Phosphate-dependent Enzyme Dialkylglycine Decarboxylase. *Journal of Molecular Biology* 245, 151-179.
20. Mehta, P. K., Hale, T. I. & Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur J Biochem* 214, 549-61.
21. Käck, H., Sandmark, J., Gibson, K., Schneider, G. & Lindqvist, Y. (1999). Crystal structure of diaminopelargonic acid synthase: evolutionary relationships between pyridoxal-5'-phosphate-dependent enzymes. *Journal of Molecular Biology* 291, 857-876.
22. Yano, T., Kuramitsu, S., Tanase, S., Morino, Y. & Kagamiyama, H. (1992). Role of Asp222 in the catalytic mechanism of Escherichia coli aspartate aminotransferase: the amino acid residue which enhances the function of the enzyme-bound coenzyme pyridoxal 5'-phosphate. *Biochemistry* 31, 5878-87.
23. Onuffer, J. J. & Kirsch, J. F. (1994). Characterization of the apparent negative co-operativity induced in Escherichia coli aspartate aminotransferase by the replacement of Asp222 with alanine. Evidence for an extremely slow conformational change. *Protein Eng* 7, 413-24.
24. Eliot, A. C. & Kirsch, J. F. (2004). Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annual Review of Biochemistry* 73, 383-415.
25. Contestabile, R., Angelaccio, S., Maytum, R., Bossa, F. & John, R. A. (2000). The Contribution of a Conformationally Mobile, Active Site Loop to the Reaction Catalyzed by Glutamate Semialdehyde Aminomutase. *J. Biol. Chem.* 275, 3879-3886.
26. Tyacke, R. J., Harwood, J. L. & John, R. A. (1993). Properties of the pyridoxaldimine form of glutamate semialdehyde aminotransferase (glutamate-1-semialdehyde 2,1-aminomutase) and analysis of its role as an intermediate in the formation of aminolaevulinate. *Biochem J* 293, 697-701.
27. Astner, I., Schulze, J. O., van den Heuvel, J., Jahn, D., Schubert, W. D. & Heinz, D. W. (2005). Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *Embo J* 24, 3166-77.
28. Schmidt, A., Sivaraman, J., Li, Y., Larocque, R., Barbosa, J. A., Smith, C., Matte, A., Schrag, J. D. & 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, 5151-60.
29. Alexeev, D., Alexeeva, M., Baxter, R. L., Campopiano, D. J., Webster, S. P. & Sawyer, L. (1998). The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *J Mol Biol* 284, 401-19.
30. Jansonius, J. N. (1998). Structure, evolution and action of vitamin B6-dependent enzymes. *Current Opinion in Structural Biology* 8, 759-769.
31. Otsuka, A., Buoncristiani, M., Howard, P., Flamm, J., Johnson, C., Yamamoto, R., Uchida, K., Cook, C., Ruppert, J. & Matsuzaki, J. (1988). The Escherichia coli biotin biosynthetic enzyme sequences predicted from the nucleotide sequence of the bio operon. *J. Biol. Chem.* 263, 19577-19585.
32. Duncan, R., Faggart, M. A. & Cornell, N. W. (1997). Phylogenetic Analysis of the 5-Aminolevulinate Synthase Gene. *Biol Bull* 193, 247-248.
33. Gilbert, W. (1986). The RNA world. *Nature* 319, 618.

34. Cunchillos, C. & Lecointre, G. (2003). Evolution of Amino Acid Metabolism Inferred through Cladistic Analysis. *J. Biol. Chem.* 278, 47960-47970.
35. Benner, S. A., Ellington, A. D. & Tauer, A. (1989). Modern metabolism as a palimpsest of the RNA world. *Proc Natl Acad Sci U S A* 86, 7054-8.
36. Gray, M. W., Burger, G. & Lang, B. F. (1999). Mitochondrial Evolution. *Science* 283, 1476-1481.
37. Rivera, M. C., Jain, R., Moore, J. E. & Lake, J. A. (1998). Genomic evidence for two functionally distinct gene classes. *PNAS* 95, 6239-6244.
38. Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-7.
39. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* 26, 795-800.
40. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* 55, 484-91.
41. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5), 905-21.
42. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D - Biological Crystallography* 53, 240-255.
43. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-32.
44. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 8, 52-6, 29.
45. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
46. DeLano, W. L. (2002). *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA, USA.
47. Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-3.
48. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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, 4673-80.
49. Holm, L. & Sander, C. (1996). Mapping the protein universe. *Science* 273, 595-603.
50. Nakai, T., Okada, K., Akutsu, S., Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S. & Hirotsu, K. (1999). Structure of Thermus thermophilus HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry* 38, 2413-24.

Figure Legends

Figure 1: Scheme of the reaction catalyzed by GSAM. Glutamate-1-semialdehyde (GSA) is converted to 5-aminolevulinic acid (ALA) via the intermediate 4,5-diaminovalerate (DAVA). During the first half of the reaction pyridoxamine 5'-phosphate (PMP) is converted to pyridoxal 5'-phosphate (PLP), in the second half PMP is regenerated. An alternative mechanism (dotted arrows) proceeds via the intermediate 4,5-dioxovalerate (DOVA; gray). Here, PLP is converted to PMP and regenerated upon ALA formation.

Figure 2: Structure of GSAM from *Thermosynechococcus elongatus* ($GSAM_{Tel}$). One subunit of the homodimer is colored in yellow, the color of the second subunit changes from light blue for the N-terminus to dark blue for the C-terminus. (a) Cartoon representation with the cofactor PLP (green) depicted in ball-and-stick representation, carbon in green, oxygen in red, nitrogen in blue, phosphorus in orange. (b) Electron density of the "internal aldimine" consisting of PLP (green) covalently bound to Lys251. Only residues in contact with PLP are shown. (c) Surface representation with unresolved residues modeled in and shown in transparent colors. The orange residues belong to the yellow subunit, the cyan residues to the blue monomer. The molecule is rotated by 90° along the dimerization interface of Figure 2(a). (d) Structures of $GSAM_{Tel}$ (blue and yellow) and $GSAM_{Syn}$[8] (green and gray) superimposed on the subunit in the back. The lines subtending an angle of 10° indicate a rotational displacement of the subunit of $GSAM_{Tel}$ in the front in comparison to $GSAM_{Syn}$. The figures were produced using PyMOL[46].

Figure 3: Model of the proposed complex between $GSAM_{Tel}$ and glutamyl-tRNA reductase (GluTR) from *Methanopyrus kandleri*[12]. GluTR is shown in green as cartoon representation

and $GSAM_{Tel}$ as in Figure 2(b). The red arrow indicates the direction of the flexible movement in $GSAM_{Tel}$. The co-crystallized GluTR inhibitor glutamycin is shown as ball-and-stick representation with carbon atoms in green. The substrate channeling between the active sites of both enzymes is indicated by a green arrow.

Figure 4: Structural comparison of PLP-dependent enzymes of the α family in cartoon representation with PLP in ball-and-stick representation (a) Structure of the 5-aminolevulinic acid synthase (ALAS) homodimer from *Rhodobacter capsulatus* with one subunit colored in gray and the other subunit in shades of orange. (b) Structure of $GSAM_{Tel}$, rotated by 90° along the dimerization interface of Figure 2(a). (c) Evolutionary pedigree of α-family PLP-dependent enzymes according to Mehta and Christen[7]. The previous position of the CoA subfamily is depicted in gray; the suggested new location is shown in red. (d)-(f) Structures of aspartate aminotransferase[50], serine hydroxymethyltransferase and kynureninase, the enzymes of subfamilies AT I, AT IV and γ resulting in the highest Z-scores when compared to ALAS. One subunit is rendered in green, the other subunit in yellow.

Tables

Table 1: Data collection and structure determination statistics

Data collection	
Unit cell dimensions, a, b, c (Å)	108.9, 108.9, 93.4
Space group	$P4_12_12$
Wavelength (Å)	1.54
Number of unique reflections	26282
Resolution range (Å)	50-2.8 (2.9-2.8)
Completeness of data (%)	99.5 (99.7)
Redundancy	4.2 (4.2)
R_{merge} (%)	15.0 (47.1)
I/σ	10.0 (3.0)

Refinement statistics	
Maximum resolution (Å)	2.85 (2.92-2.85)
No. of atoms: protein, cofactor, water	2716, 15, 13
Monomers per asymmetric unit	1
R-factor (%)	17.7 (28.8)
R_{free} (%)	21.7 (22.8)
Average B-factor ($Å^2$)	32.2
R.m.s.d. bond length (Å)	0.02
R.m.s.d. bond angles (°)	1.63
Ramachandran plot[#]	88.3/10.4/1.3/0.0

Values in parentheses refer to shell of highest resolution.

[#] Procheck[45]: most favored/additionally allowed/generously allowed/disallowed regions

Table 2: Enzymes structurally most similar to ALAS

enzyme	PDB-ID	DALI Z-score	r.m.s.d.[#] (Å)	number of equivalent amino acids	sequence identity (%)	aminotransferase subfamily
2-amino-3-ketobutyrate CoA ligase	1FC4	49.7	1.6	377	32	CoA
8-amino-7-oxoanoate synthase	1BS0	48.4	2.2	380	28	CoA
glutamate-1-semialdehyde aminotransferase	**2GSA**	**30.5**	**2.9**	**343**	**18**	**AT II**
2,2-dialkylglycine decarboxylase	2DKB	29.0	3.7	345	14	AT II
adenosylmethionine 8-amino-oxononanoate aminotransferase	1DTY	27.7	3.1	338	15	AT II
serine hydroxymethyltransferase	1BJ4	25.9	3.6	327	15	AT IV
aspartate aminotransferase	1BJW	25.8	4.7	326	15	AT I
kynureninase	1QZ9	23.5	3.5	308	15	γ

root-mean-square deviations for common C_α atoms of one monomer as calculated by DALI[49]