



Human 5'-Aminolevulinate Synthase, Erythroid-Specific (ALAS2)



A Target Enabling Package (TEP)

Gene ID / UniProt ID / EC	212 / P22557 / EC 2.3.1.37
Target Nominator	SGC Internal Nomination
SGC Authors	Gustavo A. Bezerra, Henry J. Bailey, William R. Foster, Elzbieta Rembeza
Collaborating Authors	Jason R. Marcero ¹ , Harry A. Dailey, Jr. ¹
Target PI	Wyatt W. Yue (SGC Oxford)
Therapeutic Area(s)	Metabolic disorders
Disease Relevance	Mutations of <i>ALAS2</i> lead to X-linked sideroblastic anaemia (MIM 300751) and X-linked protoporphyria (MIM 300752). <i>ALAS2</i> is also upstream of several haem biosynthetic steps associated with inherited porphyria disorders.
Date Approved by TEP Evaluation Group	June 10, 2019
Document version	Version 2
Document version date	October 2020
Citation	Gustavo A. Bezerra, Henry J. Bailey, Jason R. Marcero, William R. Foster, Elzbieta Rembeza, Harry A. Dailey, Jr., & Wyatt W. Yue. (2019). Human 5'-Aminolevulinate Synthase, Erythroid-Specific (<i>ALAS2</i>); A Target Enabling Package (Version 0) [Data set]. Zenodo. 10.5281/zenodo.3245202
Affiliations	1. Department of Biochemistry and Molecular Biology, University of Georgia, Athens GA

USEFUL LINKS



(Please note that the inclusion of links to external sites should not be taken as an endorsement of that site by the SGC in any way)

SUMMARY OF PROJECT

Erythroid specific 5'-Aminolevulinate synthase (*ALAS2*) catalyses the first and rate-determining step in haem biosynthesis during erythroid development. This TEP provides the structural biology tools to facilitate two aspects of *ALAS2* research: (i) understanding how mutations in the *ALAS2* C-terminus, a region unique to higher eukaryotes, lead to over-activity of the enzyme associated with a gain-of-function disorder; and (ii) providing chemical starting points to explore metabolic intervention of *ALAS2*, as substrate reduction therapy for the group of porphyria disorders that are associated with haem biosynthetic steps downstream of *ALAS2*.

SCIENTIFIC BACKGROUND

The tetrapyrrole cofactor haem is essential for various cellular processes across lifeforms. In metazoans, haem biosynthesis occurs via a conserved eight-reaction pathway that requires iron, glycine and succinyl-CoA (**Fig. 1**). The first and rate-limiting step is catalysed by 5'-aminolevulinic acid synthase (**ALAS**; EC 2.3.1.37) in the mitochondria (1,2), which condenses succinyl-CoA and glycine in a pyridoxal phosphate-dependent manner, to form aminolevulinic acid (ALA) and by-products CoA and CO₂. Vertebrates encode the ubiquitously-expressed ALAS1 isozyme providing a basal level of haem for cytochromes and other hemoproteins, as well as the ALAS2 isozyme predominantly expressed in erythroid progenitor cells, synthesizing 85-90% of total body haem specifically for haemoglobin production.

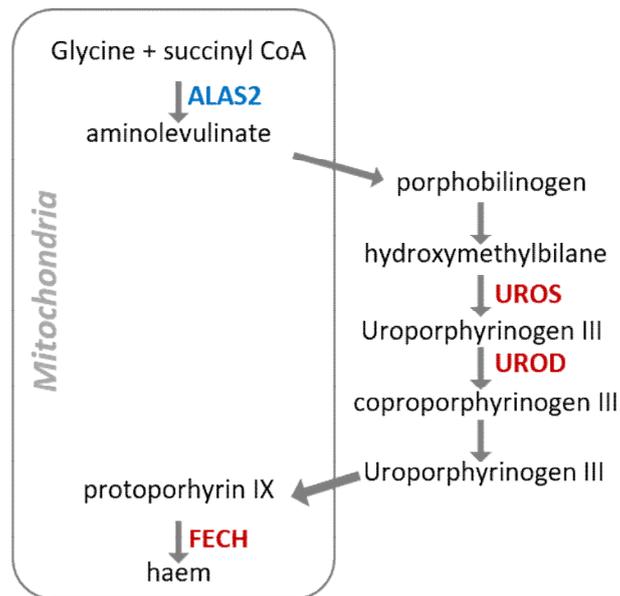


Fig. 1 Pathway of haem biosynthesis. Enzymes associated with erythropoietic porphyria in red, and the proposed therapeutic target ALAS2 in blue.

While the ALAS catalytic mechanism is well conserved across all phyla (3), eukaryotic enzymes have evolved a 35-60 amino acid extension C-terminal to the catalytic core (4). *Indel* and frameshift mutations in exon 11 of the *ALAS2* gene, resulting in deletion, replacement, or elongation of this C-terminal extension, are the molecular causes of X-linked protoporphyria (XLP, MIM 300752), an inherited blood disorder characterised by high levels of the toxic haem intermediate (e.g. protoporphyrinogen PPIX), resulting in painful phototoxicity and an increased risk for liver dysfunction (5,6). At the protein level, these genetic lesions result in enhanced enzyme activity, lending to XLP being referred to as a gain-of-function (GOF) disorder (7-9). The molecular basis of the GOF phenotype with respect to the ALAS2 C-terminal extension remains unclear, although a self-inhibitory role has been proposed for this region (10).

Defective haem biosynthesis is associated with a group of inherited blood disorders, generally named porphyrias. Three of them are erythroid-specific and associated with the loss of function of a haem biosynthetic enzyme downstream of the ALAS2 step (**Fig. 1, red**). These porphyrias result in the accumulation of toxic haem intermediates, and point to the inhibition of ALAS2 (**Fig. 1, blue**) as a potential therapeutic target via the substrate reduction approach that has precedence in other inherited metabolic disorders (11). This is lent support by recent siRNA knockdown of the *ALAS1* gene showing rescue for the hepatic forms of porphyria (12,13), pointing to similar therapeutic benefit for ALAS2 inhibition towards the erythropoietic forms of porphyria.

RESULTS – THE TEP

Proteins Purified

The human ALAS2 (hsALAS2) polypeptide is 587aa long (**Fig. 2**) and consists of a mitochondrial targeting sequence at the N-terminus (residues 1-49), a low-homology region of unknown function (residues 50-142), a conserved catalytic core (residues 143-545), and a C-terminal extension of 50 aa unique to metazoans (residues 546-587).



Fig. 2 Domain diagram of human ALAS2 showing constructs generated for wild-type and disease variant proteins.

We employed baculovirus-infected insect *Sf9* cells to express several constructs of N-terminally truncated hsALAS2 bearing an intact C-terminal extension, namely residues 79-587 (hsALAS2 $_{\Delta N78}$) and residues 143-587 (hsALAS2 $_{\Delta N142}$). Both hsALAS2 $_{\Delta N78}$ and hsALAS2 $_{\Delta N142}$ proteins are solubly expressed, loaded with PLP cofactor, and display no signs of proteolytic degradation which were reported in *E. coli* expression (8).

We also generated 4 recombinant variants (hsALAS2 $_{Q548X}$, hsALAS2 $_{delAT}$, hsALAS2 $_{delAGTG}$, and hsALAS2 $_{delG}$) reconstructing the *indel* mutations of the C-terminal extension that lead to the XLP disorder (**Fig. 2**).

In Vitro Assays

We adopted the discontinuous colorimetric activity assay as per Shoolingin-Jordan *et al* (14) to demonstrate that recombinant hsALAS2 $_{\Delta N78}$ is enzymatically active (**Fig. 3a,b**) with a k_{cat} of $3.4 \pm 0.2 \text{ min}^{-1}$ and K_m values determined for the substrates glycine ($7.6 \pm 0.5 \text{ mM}$) and succinyl-CoA ($30 \pm 3 \text{ }\mu\text{M}$) close to published results (8). Using this assay, we showed that the XLP-associated variants lead to significant increased hsALAS2 activity, attributable to either higher k_{cat} (hsALAS2 $_{Q548X}$, $17 \pm 1.1 \text{ min}^{-1}$; hsALAS2 $_{delAT}$, $15.4 \pm 0.8 \text{ min}^{-1}$) or lower K_m for succinyl-CoA (hsALAS2 $_{delG}$, $19 \text{ }\mu\text{M}$).

Surface plasmon resonance (SPR) was used to determine *in vitro* binding of substrates and inhibitors to immobilised His-tagged hsALAS2 $_{\Delta N78}$, and has confirmed its binding to succinyl-CoA, and succinyl-CoA+glycine concomitantly (**Fig. 3c**). Using cross-linking by disuccinimidyl glutarate we also detected a direct interaction between with hsALAS2 $_{\Delta N78}$ and a previously reported binding partner, namely the beta subunit (SUCLA2) of the succinyl-CoA synthetase complex SUCLG1-SUCLA2 (15,16).

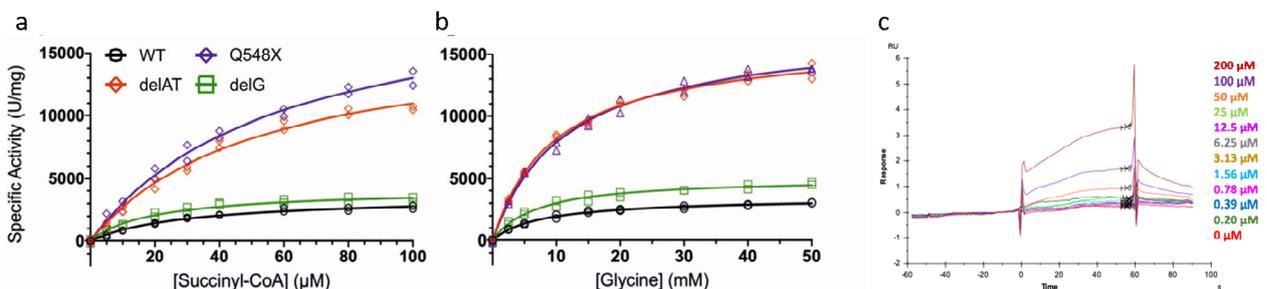


Fig. 3 Michaelis-menton kinetics of hsALAS2 $_{\Delta N78}$. Wild-type and disease-associated variants of hsALAS2 $_{\Delta N78}$ were titrated with either (a) succinyl-CoA and (b) glycine. (c) Sensorgram for the binding of succinyl-CoA to immobilised His-tagged hsALAS2 $_{\Delta N78}$.

Structural Data

We first crystallised hsALAS2 $_{\Delta N78}$ and determined its crystal structure to 2.7 Å resolution, showing that residues 79-142 N-terminal to the catalytic core were not visible in the electron density. Subsequently, the shorter hsALAS2 $_{\Delta N142}$ protein was crystallised, and its higher resolution structure (2.3 Å) was determined and described hereafter.

The hsALAS2 structure reveals a tightly interlocked homodimer (**Fig. 4a**), consistent with dimer formation by size exclusion chromatography and small angle x-ray scattering. Each active site, bound with PLP in a highly conserved pocket, is formed by residues from both dimeric subunits. The extensive dimer interface (burying 4850 Å² accessible surface per monomer) is predominantly derived from the catalytic core. The hsALAS2 catalytic core is highly homologous to structurally characterised orthologues from *R. capsulatus* (rcALAS, sequence identity 50%, RMSD 1.1 Å)(17) and *S. cerevisiae* (scALAS, 45%, 1.5 Å)(18), all belonging to the superfamily of PLP-dependent enzyme fold type I.

The most significant structural difference of hsALAS2, compared to the structural homologues, is found in its C-terminal extension that follows the catalytic core (**Fig. 4b**). This 40 residue region threads along the exterior of the catalytic core, and uses a two-turn helix ($\alpha 15$) to form a lid atop the PLP-bound active site, through several charged interactions with the catalytic core. These include (i) salt bridges of Glu571 with both Lys299 and Arg293 side-chains, (ii) a hydrogen bond between Gly576 main-chain and Arg293, and (iii) a salt bridge between side-chains of Glu569 and Arg511 (**Fig. 4c**). It is of note that Arg511 is part of the active site loop (residues 505-514), shown previously in the rcALAS structure to exhibit conformational mobility during catalysis. Helix $\alpha 15$ also contributes three aromatic residues (Trp570, Tyr574 and Phe575) that stack among themselves through π -stacking and form inter-molecular hydrophobic interactions with the opposite dimeric subunit (Phe267 and Thr268 from helix $\alpha 4$).

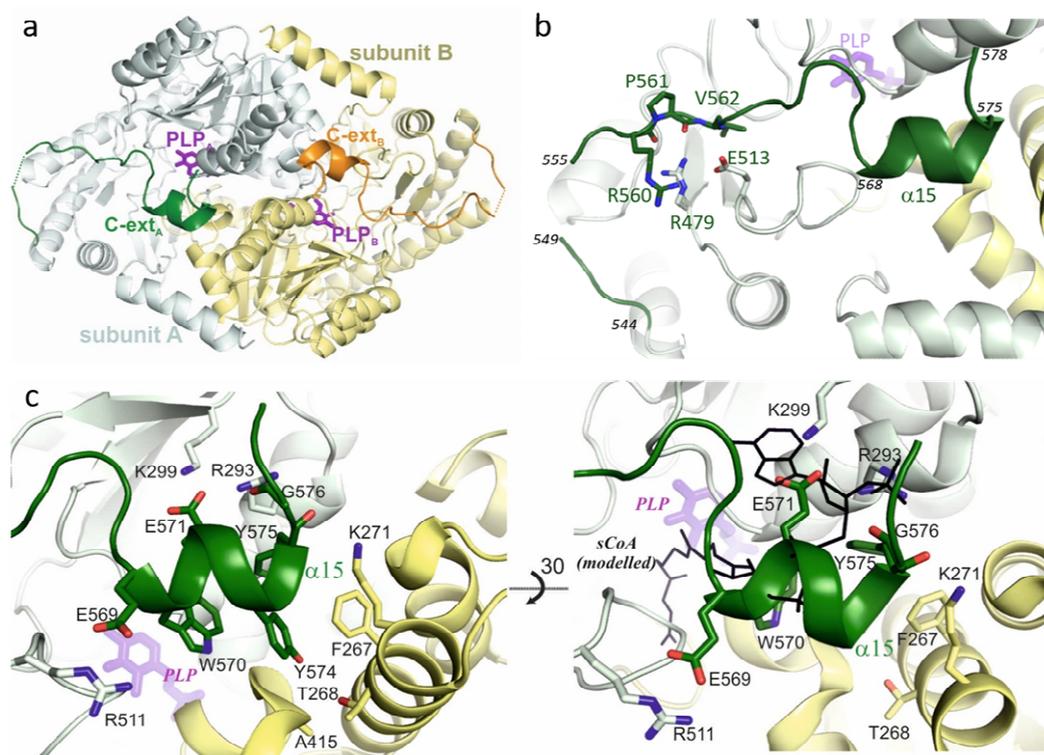


Fig. 4 Structure of hsALAS2 (a) The homodimer arrangement with PLP ligand and C-terminal extension from each subunit is highlighted; (b) Topology of an hsALAS2 C-terminal extension (dark green) packing against the catalytic core of the same subunit (light green) and of dimeric subunit (yellow); (c) Interactions of helix $\alpha 15$ from the C-terminal extension with both subunits of the homodimer. Right panel shows an overlay of succinyl-CoA taken from the rcALAS homologue structure to highlight the steric clashes with helix $\alpha 15$ of hsALAS2.

When superimposing hsALAS2 structure with that of rcALAS bound with the substrate succinyl-CoA (PDB 2BWO), we observed to our surprise that the expected hsALAS2 succinyl-CoA site is sterically blocked by helix α 15 in the C-terminal extension (Fig. 4c, right). Our hsALAS2 structure therefore represents a conformation of the C-terminal extension that precludes binding of succinyl-CoA to the active site and implies that during catalysis this C-terminal extension will undergo conformational rearrangement to mediate substrate binding and product release. Rearrangement of the C-terminal extension will likely act in concert with the active site loop, since the two regions are connected by the Glu569:Arg511 interaction.

Chemical Matter

We aim to identify chemical matter that serve as a starting point for inhibitor development, to treat inherited defects of erythroid haem biosynthesis situated downstream of the ALAS2 reaction step. To this end, we have performed a crystallography-based fragment screening (XChem) campaign. 425 hsALAS2 Δ N142 crystals were grown, yielding 295 data sets collected in the resolution range of 1.44-1.93 Å, and resulting in 25 bound fragments identified. Discarding 6 fragments mediated by artefactual crystal contacts, we have in total 19 fragments clustered into 3 different regions (Fig. 5) of the protein.

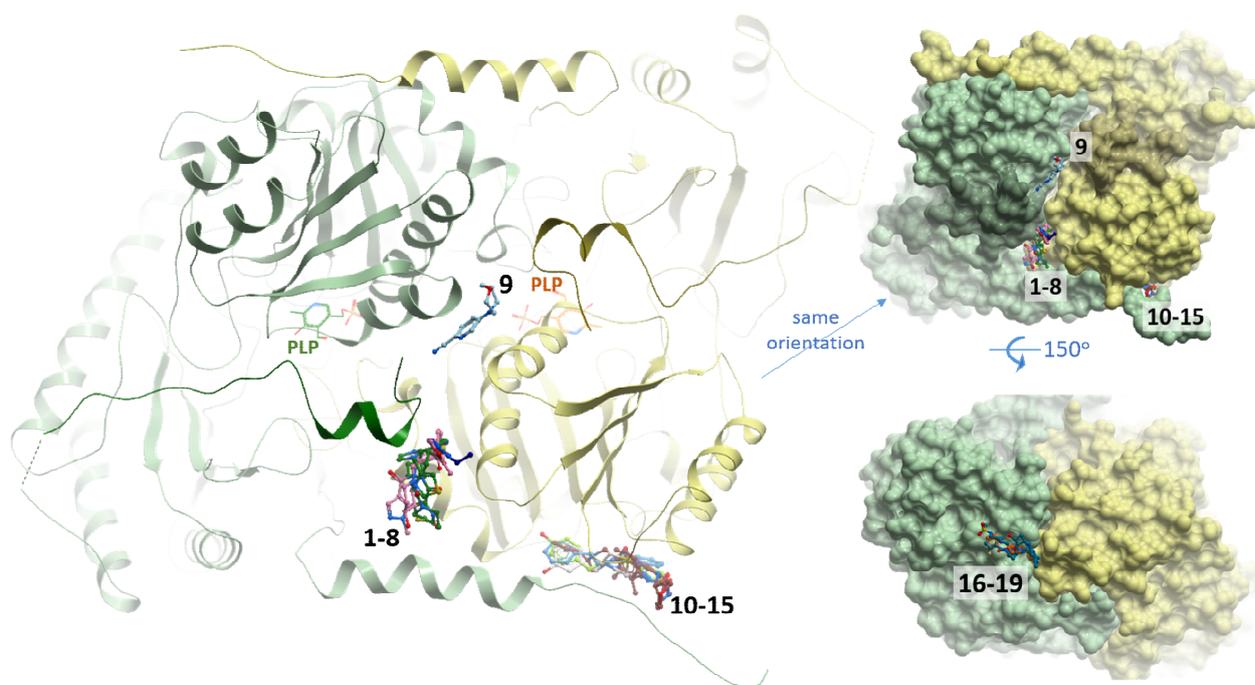


Fig. 5 ALAS2-bound fragments are clustered into distinct regions of the protein. The two subunits of an ALAS2 homodimer are coloured green and yellow. Inset shows two rotated views of ALAS2 in surface representation.

Cluster 1: C-terminal extension and between subunits A and B

8 fragments are accommodated around the hsALAS2 C-terminal extension, in a pocket between helix α 15 of one subunit and its interacting regions (helix α 4, as described above) in the opposite subunit (Fig. 6). This pocket is mainly hydrophobic in nature, contributed by Tyr574 from helix α 15 and residues Phe267, Thr268, Lys271, Ile272, Tyr413 from the opposite chain. 5 of the fragments (1-5) extend into a subsite (called 'Ia') packing against Glu155 and Lys152 from the first helix (α 1) of the catalytic core. Helix α 1 is known to stabilize the active site mobile loop and C-terminal helix α 15, by contributing Asp159 to a network of charged interactions Asp159:Arg511:Glu569. The aromatic side-chain of Tyr574 (helix α 15) adopts an alternate conformation when bound to fragments 2-4. 3 other fragments (6-8) occupy another subsite ('Ib') in direct proximity to the last turn of helix α 15 (e.g. Phe575 and Gly576). We rationalize that designing molecules to simultaneously occupy both subsites (Ia and Ib) could improve affinity and potency.

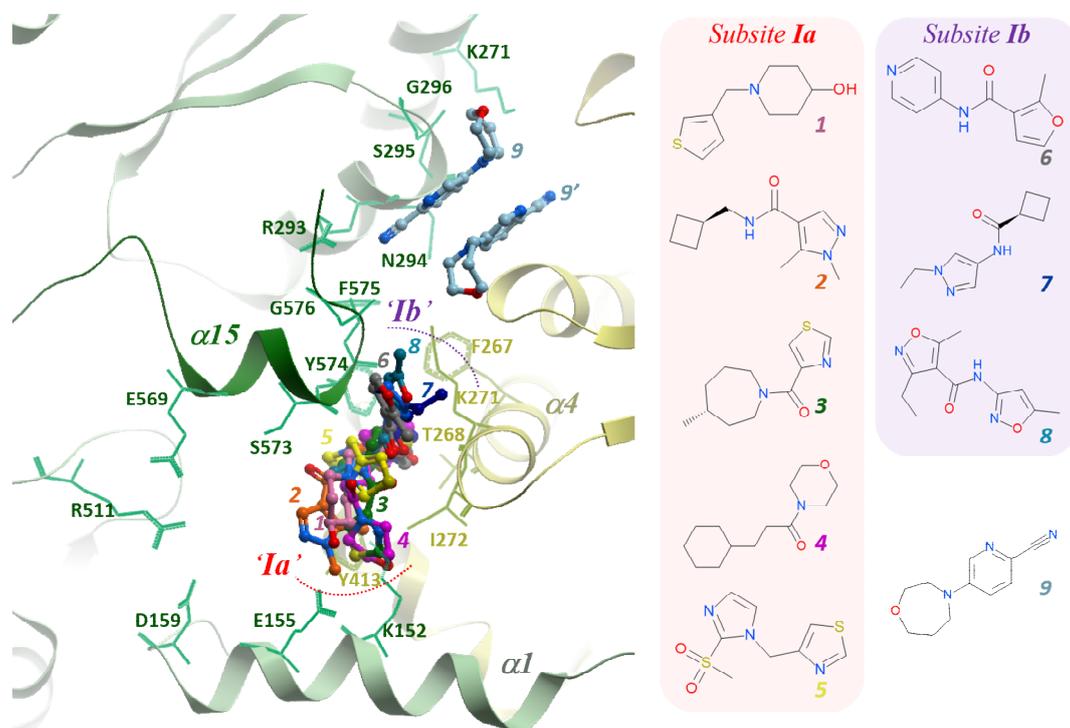


Fig. 6 Cluster 1 of fragments around the C-terminal extension. Close-up view of the C-terminal extension region of one subunit (green) and its neighbouring subunit (yellow) is shown. For fragment 9, the second molecule (9') bound to the yellow dimeric subunit is also shown. The labels of the 2D chemical structures of fragments 1-9 are individually coloured-coded according to the stick representation in the left image.

We also observed a ninth fragment (**9**) at the opposite side of C-terminal helix $\alpha 15$ with respect to the above 8 fragments (**Fig. 6**). This site is situated along the 2-fold axis of the hSALAS2 homodimer. One face of the fragment contacts a hydrophobic interaction surface (Asn294, Ser295, Gly296 and Lys271) from its own subunit; its nitrile group captures hydrogen bonds with the main chain carbonyl of Phe575 (in C-terminal helix $\alpha 15$) and guanidinium side-chain from Arg293. The opposite face of the fragment forms van der Waals contact with the same fragment bound to the dimeric subunit (**9'**).

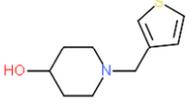
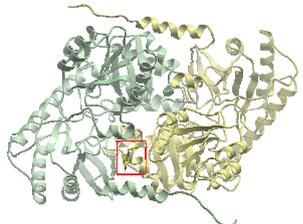
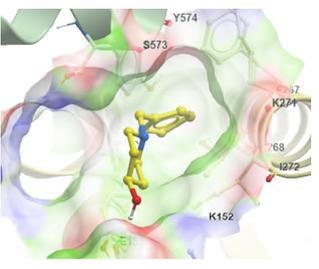
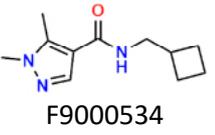
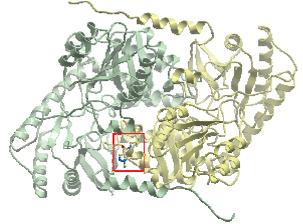
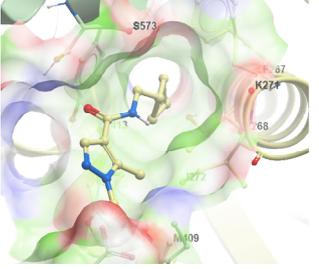
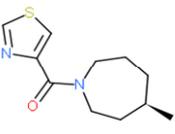
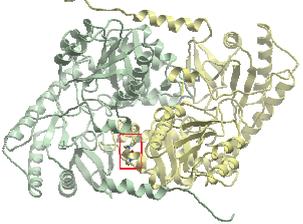
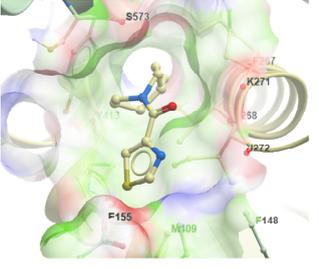
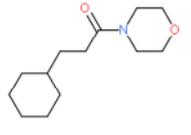
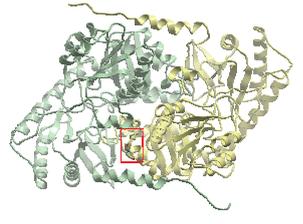
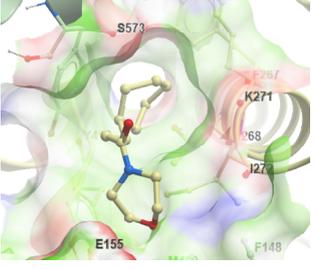
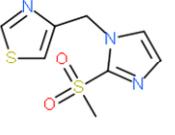
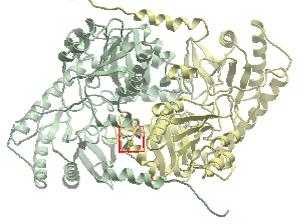
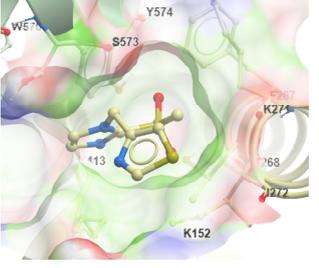
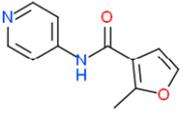
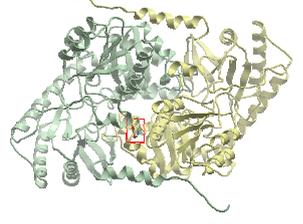
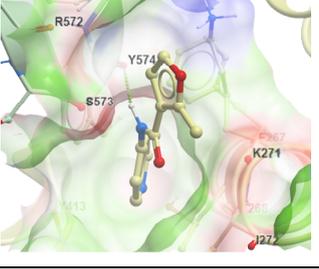
Altogether, we rationalise that fragments from cluster 1 have potential for further optimization into inhibitors aimed at increasing the residence time of the C-terminal extension on top of the active site, therefore blocking substrate access and product release to inhibit hSALAS2 catalytic activity.

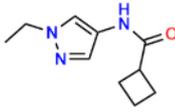
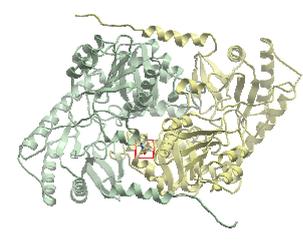
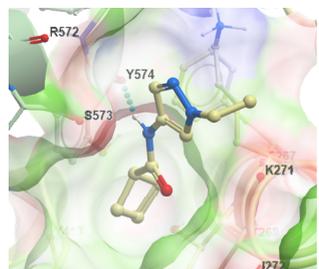
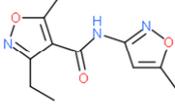
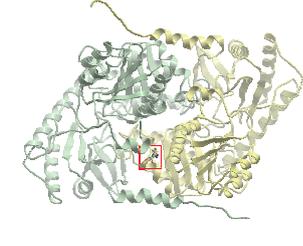
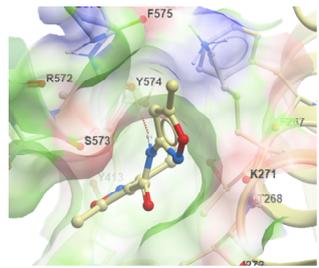
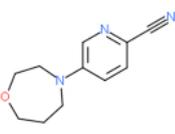
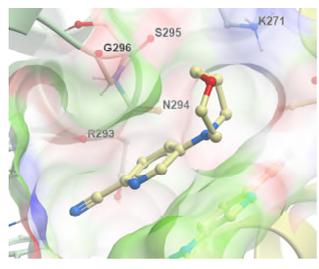
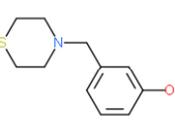
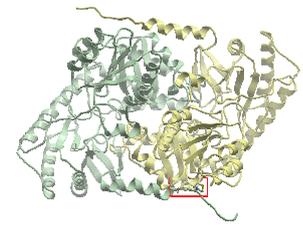
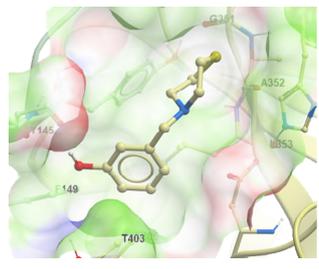
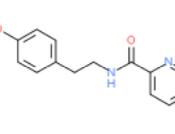
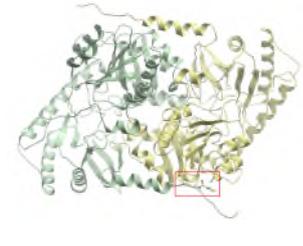
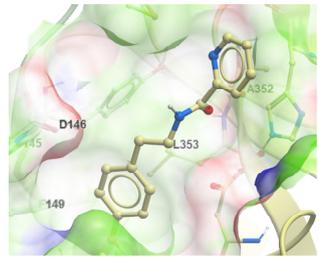
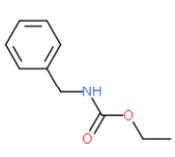
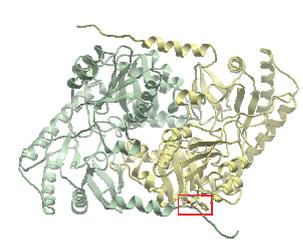
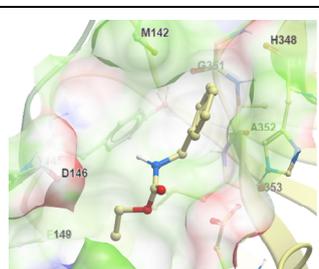
Cluster 2: Another interface between dimer subunits

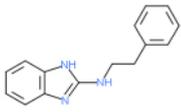
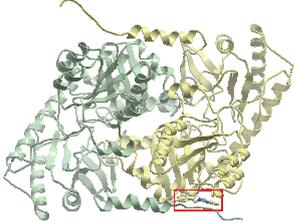
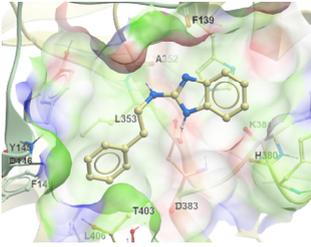
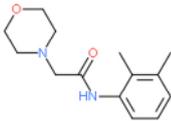
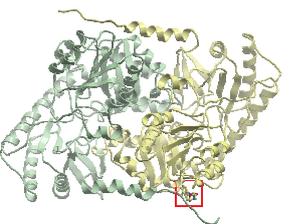
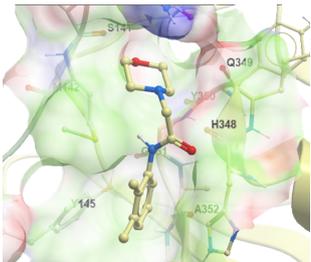
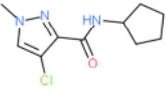
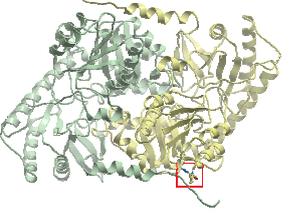
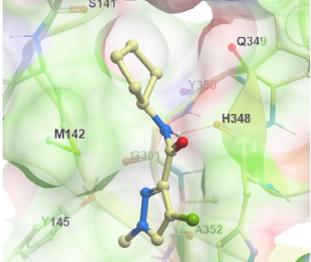
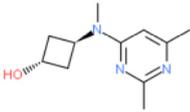
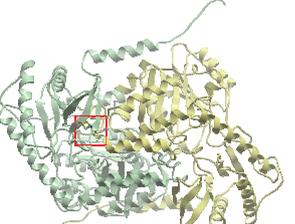
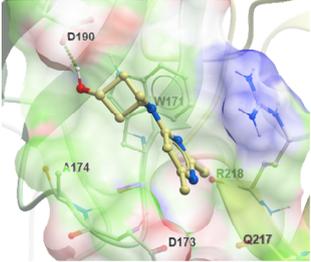
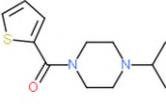
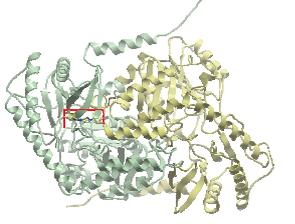
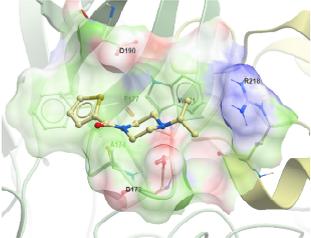
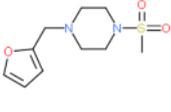
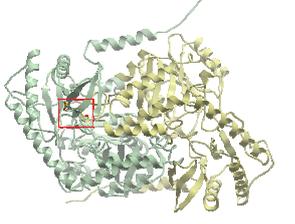
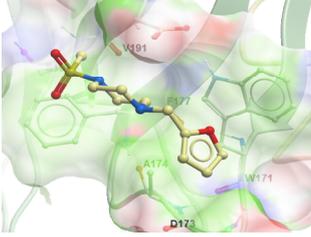
We observed a cluster of six fragments (**10-15**) situated at another interfacial region of the hSALAS2 homodimer that is distant from the enzyme's active site and C-terminal extension (**Fig. 5 & inset top**). These fragments are accommodated between the stretch before helix $\alpha 1$ of one subunit and the central $\alpha\beta$ core of the other subunit. This extensive surface, structurally conserved among ALAS orthologues, is important for maintaining the dimer assembly through several hydrophobic residues (e.g. Asp145, Asp383, Ile384, Ser402, Thr403 and Leu 406). We envisage that fragments from cluster 2 might be further explored to evaluate their potential in preventing or disrupting hSALAS2 dimerization, an essential catalytic feature of the enzyme (15).

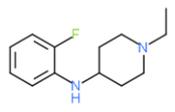
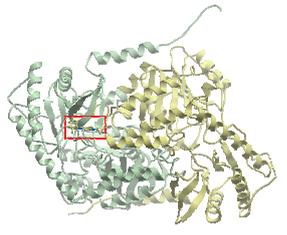
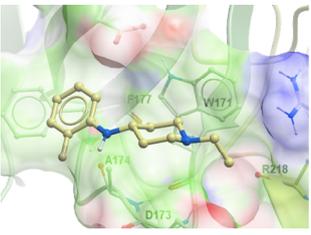
Cluster 3: Surface patch

Four fragments (**16-19**) are bound at the 'back of the protein' (i.e. opposite face from the active site) (**Fig. 5, inset bottom**). This surface patch is formed by Arg218 from one subunit and is lined up residues Trp171-Phe177 from the opposite subunit. Since this is distant from the active site, the functional relevance of cluster 3 remains to be determined in future.

PDBID	Ligand	Binding Location	Binding Pocket	Resolution (Å)
5QQY <i>fragment 1</i>	 F9000350			1.49
5QR1 <i>fragment 2</i>	 F9000534			1.44
5QRA <i>fragment 3</i>	 F9000413			1.72
5QRC <i>fragment 4</i>	 FM010110			1.82
5QRD <i>fragment 5</i>	 F9000430			1.76
5QQW <i>fragment 6</i>	 XS179878			1.56

<p>5QXX fragment 7</p>	 <p>F9000532</p>			<p>1.50</p>
<p>5QRE fragment 8</p>	 <p>Z117233350</p>			<p>1.67</p>
<p>5QQU fragment 9</p>	 <p>F9000370</p>			<p>1.55</p>
<p>5QQQ fragment 10</p>	 <p>F9000344</p>			<p>1.93</p>
<p>5QR0 fragment 11</p>	 <p>XS096095</p>			<p>1.65</p>
<p>5QR9 fragment 12</p>	 <p>FM002200</p>			<p>1.62</p>

<p>5QR7 fragment 13</p>	 <p>XS035550</p>			<p>1.74</p>
<p>5QQT fragment 14</p>	 <p>F9000368</p>			<p>1.67</p>
<p>5QQS fragment 15</p>	 <p>XS115742</p>			<p>1.85</p>
<p>5QQZ fragment 16</p>	 <p>F9000355</p>	 <p>Flipped 150°</p>		<p>1.55</p>
<p>5QR4 fragment 17</p>	 <p>F9000380</p>	 <p>Flipped 150°</p>		<p>1.57</p>
<p>5QR5 fragment 18</p>	 <p>F9000383</p>	 <p>Flipped 150°</p>		<p>1.49</p>

<p>5QRB fragment 19</p>	 <p>F9000416</p>	 <p>Flipped 150°</p>		<p>1.72</p>
---------------------------------	---	---	--	-------------

IMPORTANT: Please note that the existence of small molecules within this TEP indicates only that chemical matter might bind to the protein in potentially functionally relevant locations. The small molecule ligands are intended to be used as the basis for future chemistry optimisation to increase potency and selectivity and yield a chemical probe or lead series. As such, the molecules within this TEP should not be used as tools for functional studies of the protein, unless otherwise stated, as they are not sufficiently potent or well-characterised to be used in cellular studies.

CONCLUSION

ALAS2 is of notable importance to the field of inborn errors of metabolism as it represents one of few rare examples where inherited mutations lead to enzyme super-activity resulting in a gain-of-function XLP disorder. This TEP provides the first relevant structure to explain how frameshift/indel mutations in the ALAS2 C-terminus, which would otherwise function as an 'auto-inhibitory' element for the active site, will result in a variant protein (without the C-terminal extension) that is relieved from this inhibition, thereby becoming more enzymatically active than wild-type, as supported by our activity and binding assay data. We are currently studying the dynamics of the ALAS2 C-terminal extension, through collaboration with a Molecular Dynamics (MD) simulation team (Gopal Bulusu, TCS Innovation Labs). It also remains to be seen if the N-terminal extension to the catalytic core, a region not resolved in our crystal structures, plays a role in ALAS2 enzyme regulation.

This TEP also provides the starting point for small molecule drug discovery. There are seven porphyria disorders each associated with the loss of function of a haem biosynthetic enzyme, resulting in accumulation of toxic haem intermediates. Three of them are classified as erythroid-specific, pointing to the ALAS2 isozyme as a potential target for therapeutic inhibition via substrate reduction. Small molecule inhibitors that take advantage of the enzyme's auto-inhibitory mechanism mediated by the C-terminal extension, are potential avenues for drug development. To this end, follow-up experiments are currently underway via our binding and enzyme assays, to evaluate if the current set of fragments around the C-terminal extension (or the next generation of optimized compounds) would interfere with the mobility of the C-terminal extension that could downregulate ALAS2 and haem biosynthesis. Downstream collaborations with the relevant porphyria clinicians (Bob Desnick, Mount Sinai Hospital) are also in place to carry this work further.

As of June 2020, data from this TEP project have been published in a *Nature Communications* article (<https://www.nature.com/articles/s41467-020-16586-x>). This manuscript also reports MD simulation data from our collaborator Bulusu's team, showing that the ALAS2 C-terminal extension is mobile and switches between various conformations during catalysis. It was further shown that fragment **1** inhibited ALAS2 activity by 15% at 1 mM fragment concentration (28% at 5 mM), and during MD simulations it stabilised a conformation of the C-terminal extension as shown in the crystal structure.

FUNDING INFORMATION

The work performed at the SGC has been funded by a grant from Wellcome [106169/ZZ14/Z].

ADDITIONAL INFORMATION

Structure Files

PDB ID	Structure Details	Resolution (Å)
6HRH	Structure of hsALAS2 bound with PLP	2.30
5QQY	Structure of hsALAS2 bound with PLP and fragment 1 (x1131): (1-(thiophen-3-ylmethyl)piperidin-4-ol)	1.49
5QR1	Structure of hsALAS2 bound with PLP and fragment 2 (x1154): <i>N</i> -(cyclobutylmethyl)-1,5-dimethyl-1 <i>H</i> -pyrazole-4-carboxamide	1.44
5QRA	Structure of hsALAS2 bound with PLP and fragment 3 (x1277): (4-methylazepan-1-yl)(thiazol-4-yl)methanone	1.73
5QRC	Structure of hsALAS2 bound with PLP and fragment 4 (x1311): 3-cyclohexyl-1-morpholinopropan-1-one	1.83
5QRD	Structure of hsALAS2 bound with PLP and fragment 5 (x1317): 4-((2-(methylsulfonyl)-1 <i>H</i> -imidazol-1-yl)methyl)thiazole	1.76
5QQW	Structure of hsALAS2 bound with PLP and fragment 6 (x1084): 2-methyl- <i>N</i> -(pyridin-4-yl)furan-3-carboxamide	1.56
5QQX	Structure of hsALAS2 bound with PLP and fragment 7 (x1097): <i>N</i> -(1-ethyl-1 <i>H</i> -pyrazol-4-yl)cyclobutanecarboxamide	1.51
5QRE	Structure of hsALAS2 bound with PLP and fragment 8 (x1359): 3-ethyl-5-methyl- <i>N</i> -(5-methylisoxazol-3-yl)isoxazole-4-carboxamide	1.68
5QQU	Structure of hsALAS2 bound with PLP and fragment 9 (x1082): 5-(1,4-oxazepan-4-yl)picolinonitrile	1.55
5QQQ	Structure of hsALAS2 bound with PLP and fragment 10 (x1036): 3-thiomorpholinomethyl)phenol	1.93
5QR0	Structure of hsALAS2 bound with PLP and fragment 11 (x1147): <i>N</i> -benzyl-1-methyl-1 <i>H</i> -pyrazol-3-amine	1.66
5QR9	Structure of hsALAS2 bound with PLP and fragment 12 (x1235): ethyl benzylcarbamate	1.63
5QR7	Structure of hsALAS2 bound with PLP and fragment 13 (x1208): <i>N</i> -phenethyl-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-amine	1.75
5QQT	Structure of hsALAS2 bound with PLP and fragment 14 (x1080): 4-chloro- <i>N</i> -cyclopentyl-1-methyl-1 <i>H</i> -pyrazole-3-carboxamide	1.67
5QQS	Structure of hsALAS2 bound with PLP and fragment 15 (x1070): <i>N</i> -(2,3-dimethylphenyl)-2-morpholinoacetamide	1.85
5QQZ	Structure of hsALAS2 bound with PLP and fragment 16 (x1141): 3-((2,6-dimethylpyrimidin-4-yl)(methyl)amino)cyclobutan-1-ol	1.55
5QR4	Structure of hsALAS2 bound with PLP and fragment 17 (x1185): (4-isopropylpiperazin-1-yl)(thiophen-2-yl)methanone	1.57
5QR5	Structure of hsALAS2 bound with PLP and fragment 18 (x1196): 2-((4-(furan-2-ylmethyl)-1-(methylsulfonyl)piperazin-2-yl)oxy)-2-oxoacetic acid	1.49
5QRB	Structure of hsALAS2 bound with PLP and fragment 19 (x1281): 1-ethyl- <i>N</i> -(2-fluorophenyl)piperidin-4-amine	1.72

Non-SGC resources

Commercially available CRISPR/Cas9 knockout plasmids	Commercially available antibodies
SCBT: Cat # sc-403752	Thermofisher: TA810000 (monoclonal)
Genscript: Cat # 212 These sgRNA sequences were validated in Sanjana N.E., Shalem O., Zhang F. Improved vectors and genome-wide libraries for CRISPR screening . Nat Methods. 2014, 11(8):783-4.	SCBT: 166739 (monoclonal)

Materials and Methods

Protein expression and purification of *hsALAS2*_{ΔN78}

Vector: pFB-LIC-Bse

Entry clone accession: BC030230

Cell line: DH10Bac

Tags and additions: N-terminal, TEV protease cleavable hexahistidine tag

Construct protein sequence (underlined sequence contains vector encoded His-tag and TEV protease cleavage site*):

*hsALAS2*_{ΔN78} (SGC ID: *hALAS2-c016*; aa 79-587)

MGHHHHHSSGVDLGTENLYFQ*SMFSYDQFFRDKIMEKKQDHTYRVFKTVNRWADAYPFAQHFSEASVASKDVSVWCS
NDYLGMSRHPQVLQATQETLQRHG V GAGGTRNISGTSKFHVELEQELAEHQKDSALLFSSCFVANDSTLFTLAKILPGCEIY
SDAGNHASMIQGIRNSGAAKFVFRHNDPDHLKLLKLEKSNPKIPKIVAFETVHSMGDAICPLEELCDVSHQYGALTFVDEVHA
VGLYGSRGAGIGERD GIMHKIDIISGTLGKAFGCVGGYIASTRDLVDMVRSYAAGFIFTTSLPPMVLSGALESVRLKGGEEGQA
LRAHQNRNVKHMRLQMLMDRGLPVPICPSHIPIRVGNAALNSKLCDLLSKHGIYVQAINYPTVPRGEELLRLAPSPHHSPQM
MEDFVEKLLLAWTAVGLPLQDVSVAACNFCRRPVHFEMLSEWERSYFGNMGPQYVTTYA

hALAS2-c017 (aa 79-545)

MGHHHHHSSGVDLGTENLYFQ*SMFSYDQFFRDKIMEKKQDHTYRVFKTVNRWADAYPFAQHFSEASVASKDVSVWCS
NDYLGMSRHPQVLQATQETLQRHG V GAGGTRNISGTSKFHVELEQELAEHQKDSALLFSSCFVANDSTLFTLAKILPGCEIY
SDAGNHASMIQGIRNSGAAKFVFRHNDPDHLKLLKLEKSNPKIPKIVAFETVHSMGDAICPLEELCDVSHQYGALTFVDEVHA
VGLYGSRGAGIGERD GIMHKIDIISGTLGKAFGCVGGYIASTRDLVDMVRSYAAGFIFTTSLPPMVLSGALESVRLKGGEEGQA
LRAHQNRNVKHMRLQMLMDRGLPVPICPSHIPIRVGNAALNSKLCDLLSKHGIYVQAINYPTVPRGEELLRLAPSPHHSPQM
MEDFVEKLLLAWTAVGL

Bacmid DNA was prepared from DH10Bac cells and used to transfect Sf9 insect cells for the preparation of initial baculovirus. *hsALAS2* protein was expressed from infected Sf9 cells cultivated in InsectXpress medium (Lonza) for 72 hours at 27°C.

Harvested cells were resuspended in lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% Glycerol, 20 mM Imidazole pH 7.5, 0.5 mM TCEP, 1 μL per 1 mL protease inhibitor cocktail EDTA-free).

The cell pellet was dissolved in approximately 200 mL lysis buffer and broken by sonication, performed at 35% amplitude for 5 minutes with cycles of 5 seconds on and 10 seconds off. The cell debris was pelleted at 35000 x g, 1h and the supernatant used for purification with Nickel resin.

Buffers used are detailed hereafter;

Binding Buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% Glycerol, 20 mM Imidazole pH 7.5, 0.5 mM TCEP

Wash Buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% Glycerol, 40 mM Imidazole pH 7.5, 0.5 mM TCEP

Elution Buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% Glycerol, 250 mM Imidazole pH 7.5, 0.5 mM TCEP

The clarified cell extract was added to 5 ml of Ni-NTA resin pre-equilibrated with lysis buffer and passed through a glass column. The column was then washed with Binding Buffer (2 x 50 mL) and Wash Buffer (2 x 50 mL). The protein was eluted with Elution Buffer in 5 x 5 mL fractions. The eluted fractions from column 1 were pooled and concentrated to 10 mL with a 30 kDa MWCO spin concentrator and injected into an S200 16/60 column (pre-equilibrated in GF Buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 5% Glycerol)) at 1.0 mL/min. 1.5 mL-fractions were collected. The eluted protein was pooled and concentrated to 10 mg/mL using a 30 kDa mwco concentrator.

Activity assay

Reaction mixtures consisted of 50 mM potassium phosphate buffer, pH 7.4, 50 mM PLP, 1 mM DTT, and 10 mM MgCl₂, various concentrations of glycine and succinyl-CoA (Sigma), and 1-4 μg/mL fresh purified ALAS2 enzyme (175 μL total). After incubation at 37°C for 15 minutes (previously checked for linear ALA formation with each enzyme concentration), reactions were terminated with 100 μL trichloroacetic acid and centrifuged at 13,000xg for 5 min to remove protein. Supernatants (240 μL) were added to freshly prepared mixtures of

240 μL of 1 M sodium acetate, pH 4.7, and 20 μL acetylacetone (500 μL total), and boiled for 10 min to derivatize the ALA product. Samples were cooled and three 150 μL aliquots per reaction (three technical replicates) were further derivatized with 150 μL modified Ehrlich's reagent and monitored at 554 nm every 60 s in a CLARIOstar microplate reader (BMG Labtech). Absorbance values collected after 5 minutes were converted to molar quantities of ALA using an extinction coefficient of $60.4 \text{ mM}^{-1} \text{ cm}^{-1}$. For kinetic studies, apparent V_{max} , K_{m} , and k_{cat} values were determined by titrating 2.5-50 mM glycine in the presence of 100 μM succinyl-CoA and 5-100 μM succinyl-CoA in the presence of 50 mM glycine. Michaelis-Menten nonlinear regression analysis was subsequently carried out on data for two separate protein preparations (two biological replicates) with Prism software (GraphPad 8.0).

Crystallisation

To crystallise hsALAS2 $_{\Delta\text{N}142}$, sitting drops containing 75 nL protein (17 mg/mL) and 75 nL well solution containing 25% (w/v) PEG 3350, 0.1 M bis-tris pH 6.7 and 0.3 M MgCl_2 were equilibrated at 20 °C by vapour diffusion. Crystals were cryo-protected using 25% (v/v) ethylene glycol and flash-cooled in liquid nitrogen. Diffraction data were collected at the DLS beamlines i04 and i24.

Structure Determination

hsALAS2 $_{\Delta\text{N}142}$ crystallised in the monoclinic space group C2 with two molecules (chains A & B) in the asymmetric unit. The data was processed using the Xia2 autoprocessing dial pipeline. Structure was solved by molecular replacement using the program PHASER (19) and the rcALAS structure (PDB code 2BWN)(17) as search model. The final model was produced by iterative cycles of restrained refinement and model building using COOT (20), REFMAC5 (21) and Phenix.refine (22). The final model consists of Phe143-Met578 of chains A and B. In both chains the last 9 residues aa 579-587 and a loop between Ser182 and Ser188 are not visible in the electron density map indicating that this region is largely unstructured.

Crystallography-based fragment screening

To grow crystals for the fragment screening campaign, 10 mg/mL of ALAS2 was pre-incubated with 5 mM hydroxylamine for 30 minutes on ice to convert PLP into the homogenous, non-covalent form. Crystals were grown by vapour diffusion in 400 nL sitting drops in the presence of seeds, at 20°C equilibrated against well solutions of 0.1 M bis-tris pH 7.0, 0.3 M magnesium chloride and 23% PEG3350.

For soaking, 50 nL of each fragment compound from the XChem fragment library (final concentration of 125 mM) was added to a crystallisation drop using an ECHO acoustic liquid handler dispenser at the Diamond light source XChem facility. Crystals were soaked for two hours with fragments from the Diamond-SGC Poised Library before being harvested using XChem SHIFTER technology, cryo-cooled in liquid nitrogen and data sets collected at the beamline I04-1 in "automated unattended" mode. The XChem Explorer pipeline (23) was used for structure solution with parallel molecular replacement using DIMPLE (24), followed by map averaging and statistical modelling to identify weak electron densities generated from low occupancy fragments using PANDDA software (25). Coordinates and structure factors for exemplary data sets with bound fragments are deposited in the RCSB Protein Data Bank.

References

1. Gibson, K. D., Laver, W. G., and Neuberger, A. (1958) Initial stages in the biosynthesis of porphyrins. 2. The formation of delta-aminolaevulinic acid from glycine and succinyl-coenzyme A by particles from chicken erythrocytes. *Biochem J* **70**, 71-81
2. Akhtar, M., Abboud, M. M., Barnard, G., Jordan, P., and Zaman, Z. (1976) Mechanism and stereochemistry of enzymic reactions involved in porphyrin biosynthesis. *Philos Trans R Soc Lond B Biol Sci* **273**, 117-136
3. Hunter, G. A., Zhang, J., and Ferreira, G. C. (2007) Transient kinetic studies support refinements to the chemical and kinetic mechanisms of aminolevulinic synthase. *J Biol Chem* **282**, 23025-23035
4. Munakata, H., Yamagami, T., Nagai, T., Yamamoto, M., and Hayashi, N. (1993) Purification and structure of rat erythroid-specific delta-aminolevulinic synthase. *J Biochem* **114**, 103-111

5. Whatley, S. D., Ducamp, S., Gouya, L., Grandchamp, B., Beaumont, C., Badminton, M. N., Elder, G. H., Holme, S. A., Anstey, A. V., Parker, M., Corrigall, A. V., Meissner, P. N., Hift, R. J., Marsden, J. T., Ma, Y., Mieli-Vergani, G., Deybach, J. C., and Puy, H. (2008) [C-terminal deletions in the ALAS2 gene lead to gain of function and cause X-linked dominant protoporphyria without anemia or iron overload.](#) *Am J Hum Genet* **83**, 408-414
6. Ducamp, S., Schneider-Yin, X., de Rooij, F., Clayton, J., Fratz, E. J., Rudd, A., Ostapowicz, G., Varigos, G., Lefebvre, T., Deybach, J. C., Gouya, L., Wilson, P., Ferreira, G. C., Minder, E. I., and Puy, H. (2013) [Molecular and functional analysis of the C-terminal region of human erythroid-specific 5-aminolevulinic synthase associated with X-linked dominant protoporphyria \(XLDPP\).](#) *Hum Mol Genet* **22**, 1280-1288
7. Balwani, M., Doheny, D., Bishop, D. F., Nazarenko, I., Yasuda, M., Dailey, H. A., Anderson, K. E., Bissell, D. M., Bloomer, J., Bonkovsky, H. L., Phillips, J. D., Liu, L., Desnick, R. J., and Porphyrias Consortium of the National Institutes of Health Rare Diseases Clinical Research, N. (2013) [Loss-of-function ferrochelatase and gain-of-function erythroid-specific 5-aminolevulinic synthase mutations causing erythropoietic protoporphyria and x-linked protoporphyria in North American patients reveal novel mutations and a high prevalence of X-linked protoporphyria.](#) *Mol Med* **19**, 26-35
8. Bishop, D. F., Tchaikovskii, V., Nazarenko, I., and Desnick, R. J. (2013) [Molecular expression and characterization of erythroid-specific 5-aminolevulinic synthase gain-of-function mutations causing X-linked protoporphyria.](#) *Mol Med* **19**, 18-25
9. Tchaikovskii, V., Desnick, R. J., and Bishop, D. F. (2019) [Molecular expression, characterization and mechanism of ALAS2 gain-of-function mutants.](#) *Mol Med* **25**, 4
10. Kadirvel, S., Furuyama, K., Harigae, H., Kaneko, K., Tamai, Y., Ishida, Y., and Shibahara, S. (2012) [The carboxyl-terminal region of erythroid-specific 5-aminolevulinic synthase acts as an intrinsic modifier for its catalytic activity and protein stability.](#) *Exp Hematol* **40**, 477-486 e471
11. Yue, W. W., Mackinnon, S., and Bezerra, G. A. (2019) [Substrate reduction therapy for inborn errors of metabolism.](#) *Emerging Topics in Life Sciences*, ETL20180058
12. Sardh, E., Harper, P., Balwani, M., Stein, P., Rees, D., Bissell, D. M., Desnick, R., Parker, C., Phillips, J., Bonkovsky, H. L., Vassiliou, D., Penz, C., Chan-Daniels, A., He, Q., Querbes, W., Fitzgerald, K., Kim, J. B., Garg, P., Vaishnav, A., Simon, A. R., and Anderson, K. E. (2019) [Phase 1 Trial of an RNA Interference Therapy for Acute Intermittent Porphyria.](#) *N Engl J Med* **380**, 549-558
13. Yasuda, M., Gan, L., Chen, B., Kadirvel, S., Yu, C., Phillips, J. D., New, M. I., Liebow, A., Fitzgerald, K., Querbes, W., and Desnick, R. J. (2014) [RNAi-mediated silencing of hepatic Alas1 effectively prevents and treats the induced acute attacks in acute intermittent porphyria mice.](#) *Proc Natl Acad Sci U S A* **111**, 7777-7782
14. Shoolingin-Jordan, P. M., LeLean, J. E., and Lloyd, A. J. (1997) [Continuous coupled assay for 5-aminolevulinic synthase.](#) *Methods Enzymol* **281**, 309-316
15. Bishop, D. F., Tchaikovskii, V., Hoffbrand, A. V., Fraser, M. E., and Margolis, S. (2012) [X-linked sideroblastic anemia due to carboxyl-terminal ALAS2 mutations that cause loss of binding to the beta-subunit of succinyl-CoA synthetase \(SUCLA2\).](#) *J Biol Chem* **287**, 28943-28955
16. Furuyama, K., and 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
17. Astner, I., Schulze, J. O., van den Heuvel, J., Jahn, D., Schubert, W. D., and Heinz, D. W. (2005) [Crystal structure of 5-aminolevulinic synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans.](#) *EMBO J* **24**, 3166-3177
18. Brown, B. L., Kardon, J. R., Sauer, R. T., and Baker, T. A. (2018) [Structure of the Mitochondrial Aminolevulinic Acid Synthase, a Key Heme Biosynthetic Enzyme.](#) *Structure* **26**, 580-589 e584
19. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) [Likelihood-enhanced fast translation functions.](#) *Acta Crystallogr D Biol Crystallogr* **61**, 458-464
20. Emsley, P., and Cowtan, K. (2004) [Coot: model-building tools for molecular graphics.](#) *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132
21. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) [Refinement of macromolecular structures by the maximum-likelihood method.](#) *Acta Crystallogr D Biol Crystallogr* **53**, 240-255

22. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) [Towards automated crystallographic structure refinement with phenix.refine](#). *Acta Crystallogr D Biol Crystallogr* **68**, 352-367
23. Krojer, T., Talon, R., Pearce, N., Collins, P., Douangamath, A., Brandao-Neto, J., Dias, A., Marsden, B., and von Delft, F. (2017) [The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination](#). *Acta Crystallogr D Struct Biol* **73**, 267-278
24. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Pottterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) [Overview of the CCP4 suite and current developments](#). *Acta Crystallogr D Biol Crystallogr* **67**, 235-242
25. Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M., and von Delft, F. (2017) [A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density](#). *Nat Commun* **8**, 15123

We respectfully request that this document is cited using the DOI value as given above if the content is used in your work.