Human Galactose-1-phosphate uridylyltransferase (GALT)

A Target Enabling Package (TEP)

Protein 1 Gene ID / UniProt ID / EC	GALT: 2592 / P07902 / EC 2.7.7.12
Protein 2 Gene ID / UniProt ID / EC	GALK1: 2584 / P51570 / EC 2.7.1.6
Target Nominators	Kent Lai (University of Utah)
SGC Authors	Gustavo Arruda Bezerra, Sabrina MacKinnon, Thomas McCorvie, Minghao Zhang, William Foster, Michael Fairhead, Anthony Tumber, Claire Strain-Damerell, Jolanta Kopec, Fiona Fitzpatrick, Wyatt W. Yue
Collaborating Authors	Manshu Tang[1], Frank Whitby[2], Kent Lai[1], Christopher P. Hill[2]
Target PI	Wyatt W. Yue (SGC Oxford)
Therapeutic Area(s)	Metabolic diseases
Disease Relevance	In the galactose metabolic pathway, GALK1 is the enzyme upstream of GALT, mutations of which lead to classic galactosemia (OMIM 230400).
Date Approved by TEP Evaluation Group	Dec 15, 2018
Document version	1
Document version date	May 2019
Citation	Gustavo Arruda Bezerra, Sabrina MacKinnon, Thomas McCorvie, Minhao Zhang, William Foster, Michael Fairhead, … Wyatt W. Yue. (2019). Human Galactose-1-phosphate Uridylyltransferase (GALT); A Target Enabling Package. Zenodo. http://doi.org/10.5281/zenodo.3234909
Affiliations	1. Department of Nutrition and Integrative Physiology, Pediatric Genetics, University of Utah; 2. Department of Biochemistry, University of Utah.

USEFUL LINKS

GALT / GALK1 | GALT / GALK1 | GALT / GALK1 | GALT / GALK1

(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

This project describes a drug discovery plan for Classic Galactosemia, a rare disorder of galactose metabolism where there is currently no disease-transforming therapy. This TEP provides the protein reagents, biophysical assays, and structural information to facilitate two aspects of galactosemia research: (i) to understand the molecular basis of the disease due to defective GALT enzyme, and (ii) to provide chemical starting points to explore metabolic intervention of the upstream enzyme GALK1 aimed at mitigating the GALT defect (substrate reduction).

SCIENTIFIC BACKGROUND

Leloir pathway & Galactosemia

The Leloir pathway is essential for the metabolism of dietary galactose (1), generating glucose units for glycolysis and biosynthesis of glycogen, glycoproteins and glycolipids. At the hub of the Leloir pathway is **galactose-1-phosphate uridylyltransferase** (GALT), which converts galactose-1-phosphate (Gal-1-P) and UDP-glucose (UDP-Glc) into glucose-1-phosphate (Glc-1-P) and UDP-galactose (**Fig 1**)(2). Inherited mutations of the *GALT* gene lead to the autosomal recessive disorder Classic Galactosemia (OMIM 230400)(3), affecting 1:16,000-60,000 live births. In the neonatal period, patients can manifest feeding difficulties, liver failure and *E. coli* sepsis. The current mainstay treatment is life-long dietary galactose restriction, which resolves the acute life-threatening symptoms, but fails to prevent long term complications that include cognitive, neurological and reproductive impairments (4).

To date 336 mutations have been reported in the GALT database (last accessed May 2018; http://www.arup.utah.edu/database/galt/GALT_display.php), of which ~60% represent missense changes (5). The most common galactosemia variant, p.Gln188Arg, accounts for ~60% of patients. A better understanding of the underlying molecular effects of the disease-causing mutations has been hindered by the lack of structural information for the human enzyme (hGALT).

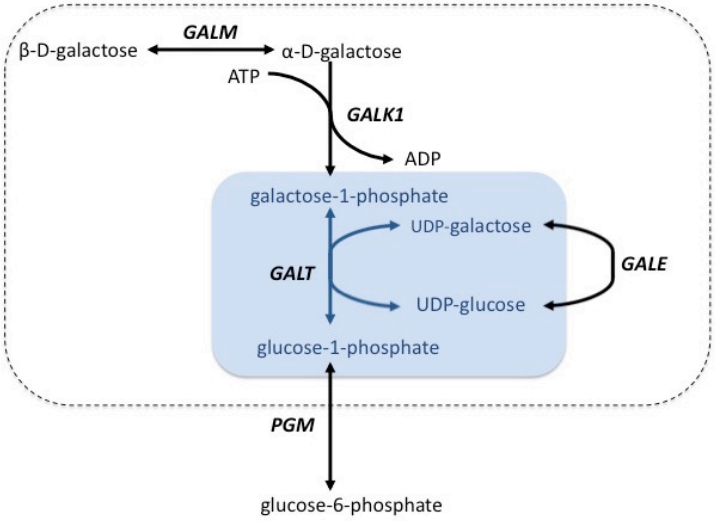


Figure 1 Leloir pathway of galactose metabolism

Rationale for GALK1 inhibitors

Elevated level of Gal-1-P, product of the upstream enzyme **galactokinase GALK1**, is shown to be a toxic pathogenic driver for classic galactosemia. Therefore, inhibition of human (hGALK1) to prevent Gal-1-P production could represent a potential therapeutic approach by substrate reduction (6). This is supported by the following experimental evidence:

- *galk1* knockout in drosophila rescued galactosemia phenotypes (7);
- *galk1* knockdown in GALT-deficient yeast abolishes sensitivity to galactose levels (8);
- patients with extremely rare cases of inherited GALK1 deficiency (OMIM 230200) present much milder phenotypes, and do not accumulate Gal-1-P (9).

RESULTS

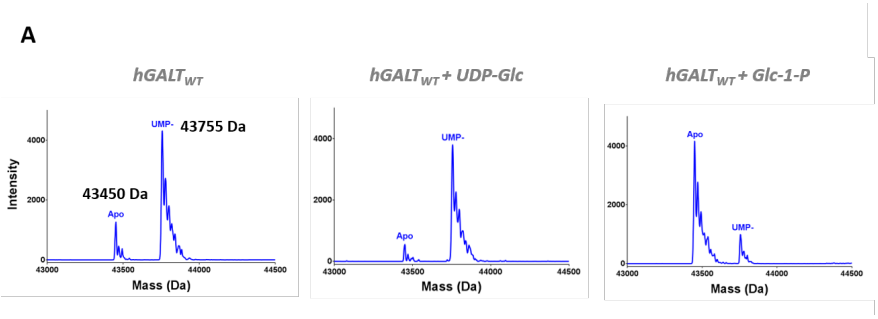
THE GALT TEP COMPONENTS

Proteins

hGALT catalyses its reaction via a covalent uridylylated intermediate. Recombinant hGALT was expressed in *E. coli*, existing in equilibrium between the *apo* (43450 Da) and uridylylated (43755 Da) states as

determined by intact mass spectrometry (MS). The ratios of the *apo* and uridylylated species are altered by incubation with either UDP-Glc or Glc-1-P, respectively (**Fig 2A**).

We expressed and purified the active site variant p.His186Ala which exists only in the *apo* form, indicating that His186 is the site of uridylylation. We also produced the variant protein p.Gln188Arg, which represents the most prevalent disease mutation. Intact MS of as-purified hGALT-p.Gln188Arg revealed predominantly the *apo* form, which only became partially uridylylated (8% of total protein) upon overnight incubation with UDP-Glc (**Fig 2B**), suggesting it has considerably lower activity compared to wild-type.



B

Apo/UMP Ratio

	As Purified	+ UDP-Glc	+ Glc-1-P
hGALT	0.32/0.68	0.19/0.81	0.81/0.19
p.H186A	1.00/0.00	1.00/0.00	N.D.
p.Q188R	1.00/0.00	0.92/0.08	N.D.

Figure 2 (**A**) Intact denaturing MS of purified wild-type hGALT shows the presence of both *apo* and uridylylated states. (**B**) Table of corresponding ratios of *apo* vs uridylylated states for different hGALT proteins.

Using the phage display approach, three single-chain variable fragment (scFv) antibodies specific for human GALT have been generated (see *Materials and Methods*).

Structures

We have determined the 1.9 Å resolution structure of hGALT (PDB 5IN3)(**Fig 3A**), containing a covalent uridylylated intermediate (covalent uridine monophosphate-His186 linkage) and Glc-1-P molecule in the active site (**Fig 3B**), as well as a structural Zn^{2+}-binding site, per monomer. This structure therefore represents a ternary complex in the reaction path of hGALT, after hydrolysis of the UDP-Glc substrate. hGALT is an obligate dimer, where each active site is formed by both dimeric subunits (**Fig 3C**). We have also determined the 1.9 Å resolution structure of a hGALT variant harbouring substitutions at the protein surface to a crystallization-prone epitope motif (A21Y:A22T:T23P:R25L)(PDB 6QD), aimed at enhancing protein crystallizability.

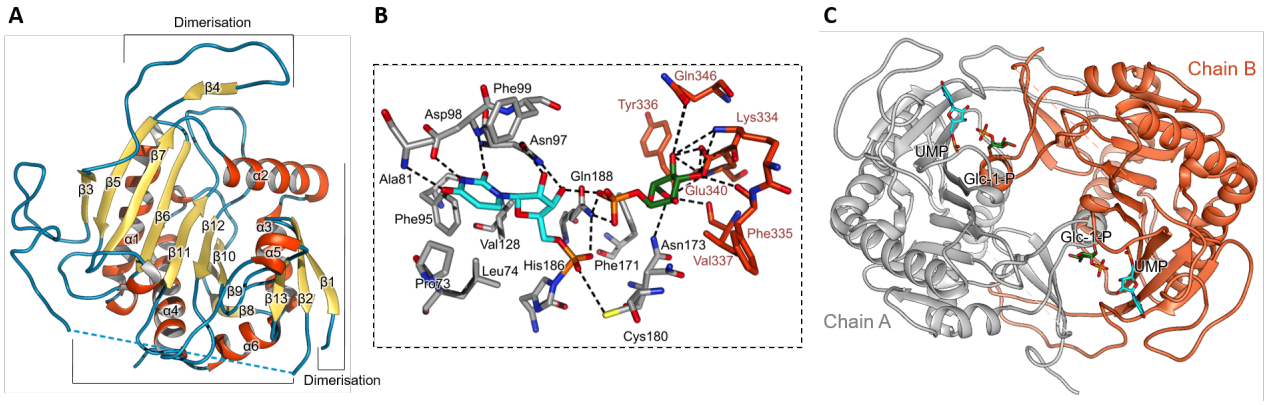


Figure 3 (**A**) Cartoon representation of hGALT structure showing secondary structure elements. (**B**) Active site residues involved in UMP (cyan carbon atoms) and Glc-1-P (green carbon atoms) binding. (**C**) View of the hGALT dimer along the 2-fold symmetry axis. UMP and Glc-1-P are represented as sticks.

To date, 336 mutations are annotated in the GALT database, with 178 of the missense type. On the basis of our hGALT structure, we categorized the missense variants into possible molecular defects of substrate binding, metal binding, dimerisation, and folding (**Fig 4A**). The most prevalent mutation p.Gln188Arg (accounting for ~60% of patients) affects the active site residue Gln188, which forms important hydrogen bonds with both the UMP and Glc-1-P moieties (**Fig 4B**). The essentiality of this residue is underlined in its strict invariance among GALT orthologues. The p.Gln188Arg variant protein lacks enzyme activity in our activity assay, and is destabilized and susceptible to degradation *in vitro* (data not shown).

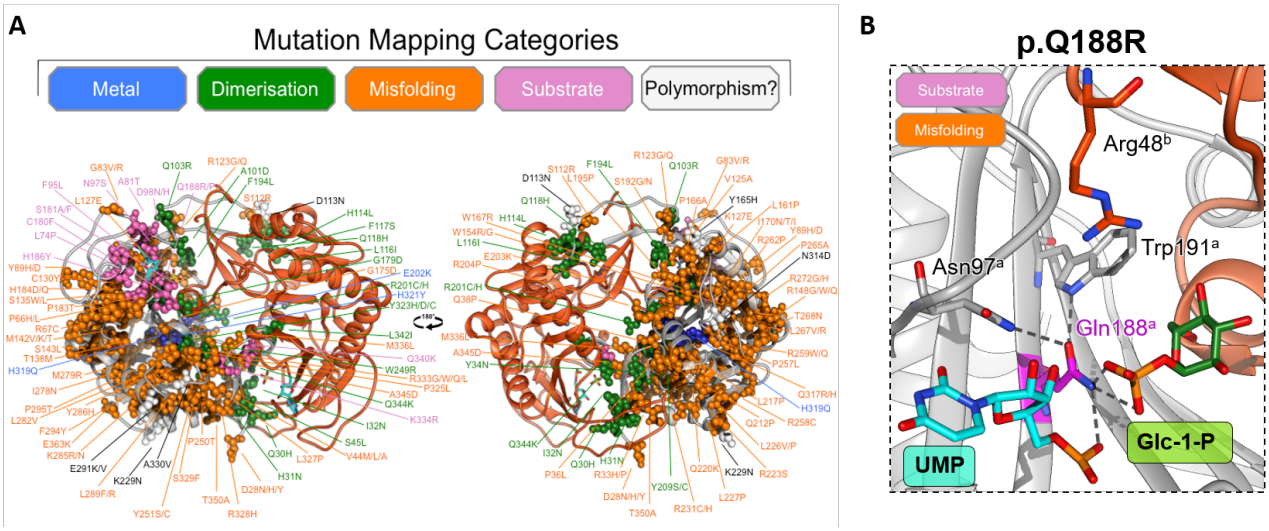


Figure 4 (**A**) Structural mapping and categorization of hGALT missense mutations that cause classic galactosemia. (**B**) Atomic environment around Gln188, site of the most prevalent disease causing mutation p.Gln188Arg. This mutation likely alters interactions with the substrate and the surrounding residues.

Assays

hGALT catalyses its reaction in two steps (**Fig 5A**). In step 1, hydrolysis of UDP-Glc results in a phospho-histidine bond between uridine monophosphate (UMP) and His186, with the release of Glc-1-P. In step 2, this covalent uridylylated intermediate is attacked by an incoming Gal-1-P, forming the product UDP-galactose. We have developed an enzyme activity assay for recombinant hGALT, to measure the formation of uridylylated-hGALT upon addition of UDP-Glc (step 1) by RapidFire mass spectrometry. Using this assay, we have confirmed that our recombinant wild-type protein is active, determining a K_M of 0.802 µM for UDP-Glc (**Fig 5B**). We also showed that recombinant hGALT-p.Gln188Arg, representing the prevalent disease-causing mutation, has negligible activity under the experimental conditions of our assay.

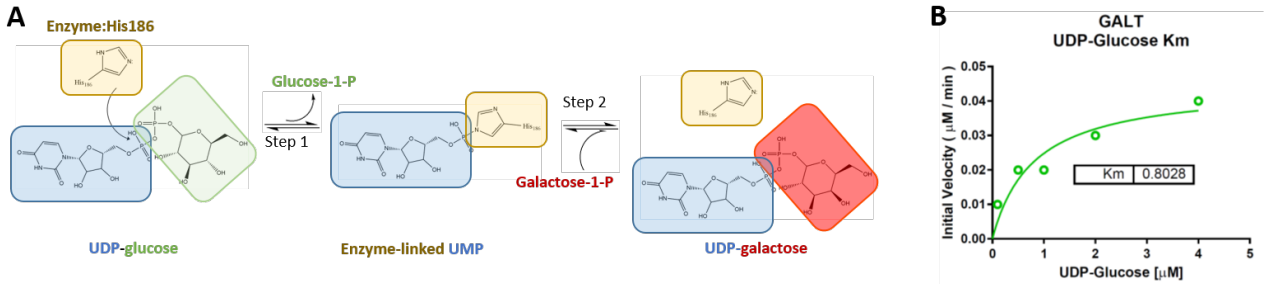


Figure 5 (**A**) hGALT carries out its reversible catalysis with a covalent uridylylated (UMP-) intermediate. Our RapidFire mass spectrometry assay follows the first step of hGALT reaction. Here, the UMP group of UDP-Glc substrate is reversibly attached to the active site His186 resulting in the release of the Glc-1-P. (**B**) Kinetic characterization of recombinant hGALT.

THE GALK1 TEP COMPONENTS

Structural biology of GALK1

Prompted by knockdown studies (7,8) indicating that hGALK1 inhibition could present a potential therapy, we initiated work on structure-guided drug discovery. The structure of hGALK1 was previously reported in complex with galactose and the ATP analogue AMP-PNP (PDB 1WUU)(10). We have additionally determined the structure of hGALK1 in complex with galactose and ADP at 2.5 Å resolution (**Fig 6**). The hGALK1 overall fold has a V-shaped architecture which can be divided into N and C-terminal domains containing α-helical and β-sheet elements. This results in a central cleft where the galactose and ATP binding site are situated (**Fig 6, inset**).

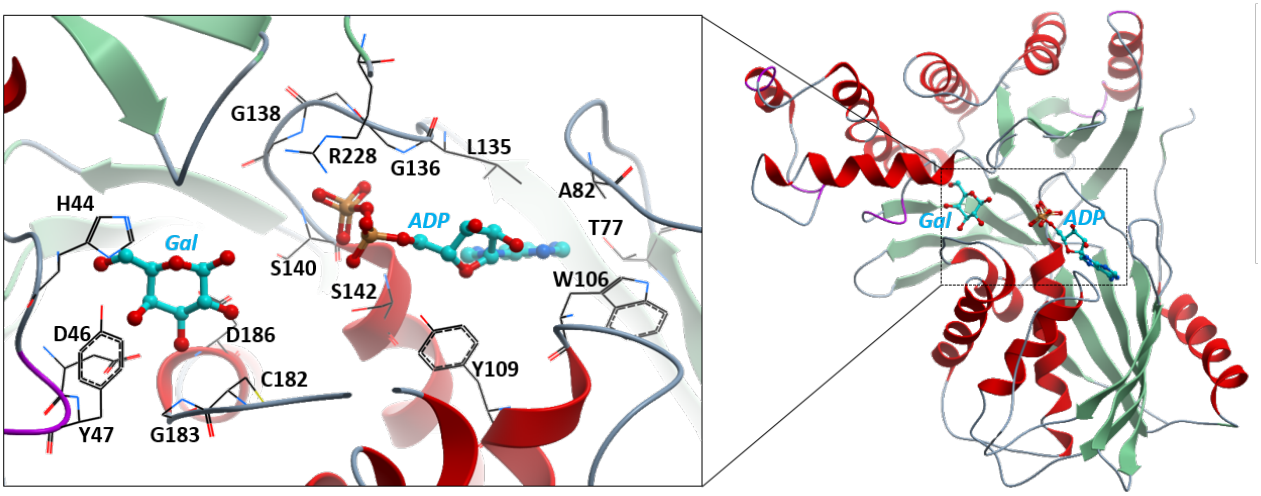


Figure 6 Structure of hGALK1 in complex with galactose and ADP, with inset showing the active site with bound ligands.

Biophysical characterisation of hGALK1 tool compounds

We have set up an hGALK1 activity assay to monitor ATP depletion by luminescence detection (Kinase-Glo, Promega), determining its K_M for galactose as 100 µM (**Fig 7A**). We further applied this assay to characterise a selection of tool inhibitors reported in a previous high-throughput screening campaign (11). We have determined the IC_{50} for the tool compounds as 70 µM (cpd1) and 17 µM (cpd3)(**Fig 7B, C**), in the same order of magnitude as published values in the low-to-mid µM range (11). We attempted to determine the IC_{50} for cpd7, but due to its low affinity (**Fig 7D**) it was not possible to obtain a reliable fit.

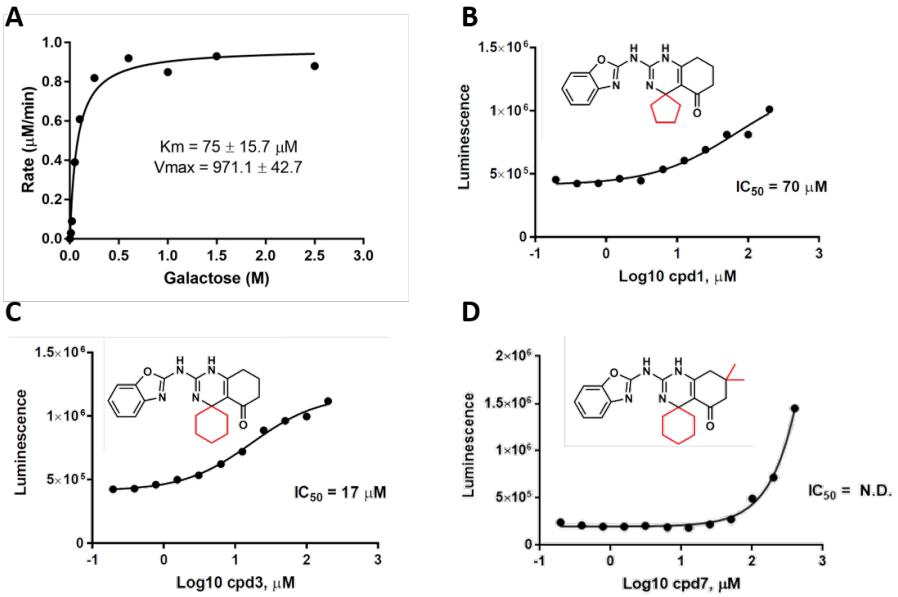


Figure 7 Kinetic characterisation of GALK *in vitro* activity. (**A**) Km determination for Galactose. (**B**) IC_{50} for cpd1 (Interbioscreen: STOCK4S-65953). (**C**) IC_{50} for cpd3 (Interbioscreen: STOCK4S-53098). (**D**) IC_{50} for cpd7 (Interbioscreen STOCK5S-50080).

We have also applied the surface plasmon resonance (SPR) method to validate binding of tool compounds to immobilized hGALK1 protein in solution (**Fig 8**). We observed no detectable binding for cpd2 and cpd7 binding was too weak, at accessible concentrations, for quantitative analysis. Cpd1 and cpd3 showed good binding to hGALK1, along with a small amount of non-specific binding to the sensor, with calculated dissociation constants (K_D) of 2.6 and 4.8 µM respectively.

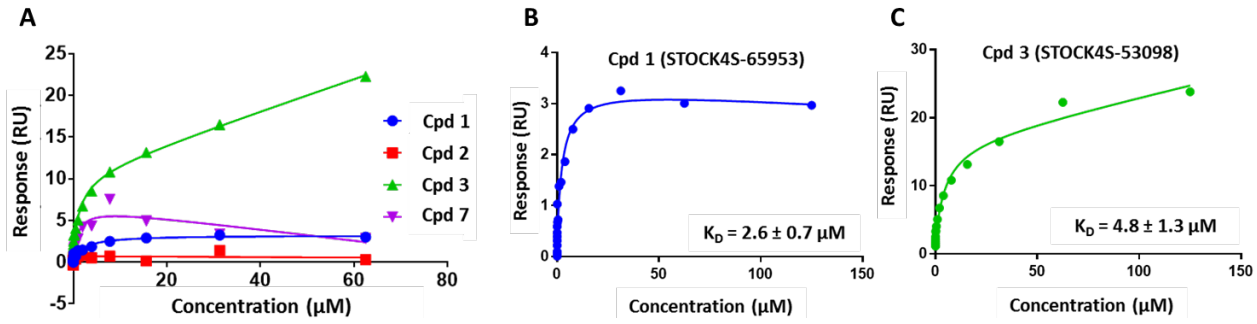


Figure 8 Surface plasmon resonance analysis. (**A**) Binding response curves for co-crystallised cpd1,2,3,7. (**B**) Binding response curve for cpd 1. (**C**) Binding response curve for cpd 3.

Crystallography-based fragment screening

We previously found that crystals of hGALK1 in absence of galactose or ATP analogues were not of the required reproducibility and diffraction consistency for large-scale fragment screening. Intensive efforts on remedial strategies at the construct/protein/crystal level nonetheless did not yield significant improvement. In light of our results from co-crystallization of tool compounds that provided diffraction data, we pre-incubated hGALK1 with the ATP-analogue cpd3 and substrate galactose to successfully generate crystals of required quality and quantity for one round of fragment screening campaign. There is also rationale that pre-loading the protein with these ligands would enable us to find fragment hits away from the canonical ATP/galactose pockets within the active site, or in other allosteric sites of the protein altogether.

In this campaign, diffraction datasets from 180 hGALK1 crystals, each soaked with an individual fragment, have been collected and processed. This campaign yielded 86 fragment binding events from across 50 crystals. After model building and inspection, we identified that 27 fragments are clustered into three well resolved binding sites of hGALK1. An overview of bound fragments can be found in **Fig 9**.

The three identified binding sites are described as follows, using 5 fragment-bound structures that we have currently deposited:

(1) 2 fragments (**b**, **c**) were identified at the entrance to the ATP binding pocket. Both fragments cause small movements in Arg105 and interact with Tyr109. Fragment **c** forms additional interactions with Asn108 while fragment **b** is close to the backbone of Ala178. These fragments represent an opportunity for merging with ATP-analogues that bind within the active site.

(2) 3 fragments (typified by fragment **a**) were identified in a shallow surface groove, close to where the N and C termini of the protein meet. This groove is primarily lined with neutral residues – Leu13, Val32, Ser33, Met60, Val63, Leu390 – that form hydrogen bonds with the bound fragments. Mutations on residues near this pocket – such as Val32Met – have been shown to abolish activity in patient-derived materials and cause insolubility of recombinantly expressed protein (10), indicating that the structural integrity of this pocket is required for hGALK1 activity.

(3) The remaining 22 fragments cluster in a binding hotspot, on the opposite side of the protein to the galactose binding site. This hotspot is formed by a hydrophobic tunnel, running behind a surface loop (Pro212-Val217) and lined on one side by three beta strands (Gly350-Ala358, Tyr339-Met343 and Leu218-Thr223). Fragments bind either to the upper opening of the tunnel (15 fragments, represented by fragment **d,** interacting with surface-exposed charged residues, Lys217 and Gln382), or the lower opening of the tunnel (7 fragments, represented by fragment **e,** displacing Tyr300). **Fig 9 (green inset)** illustrates the potential for merging fragments, such as **d** and **e,** from the two regions of the tunnel to rapidly generate the next generation of ligands.

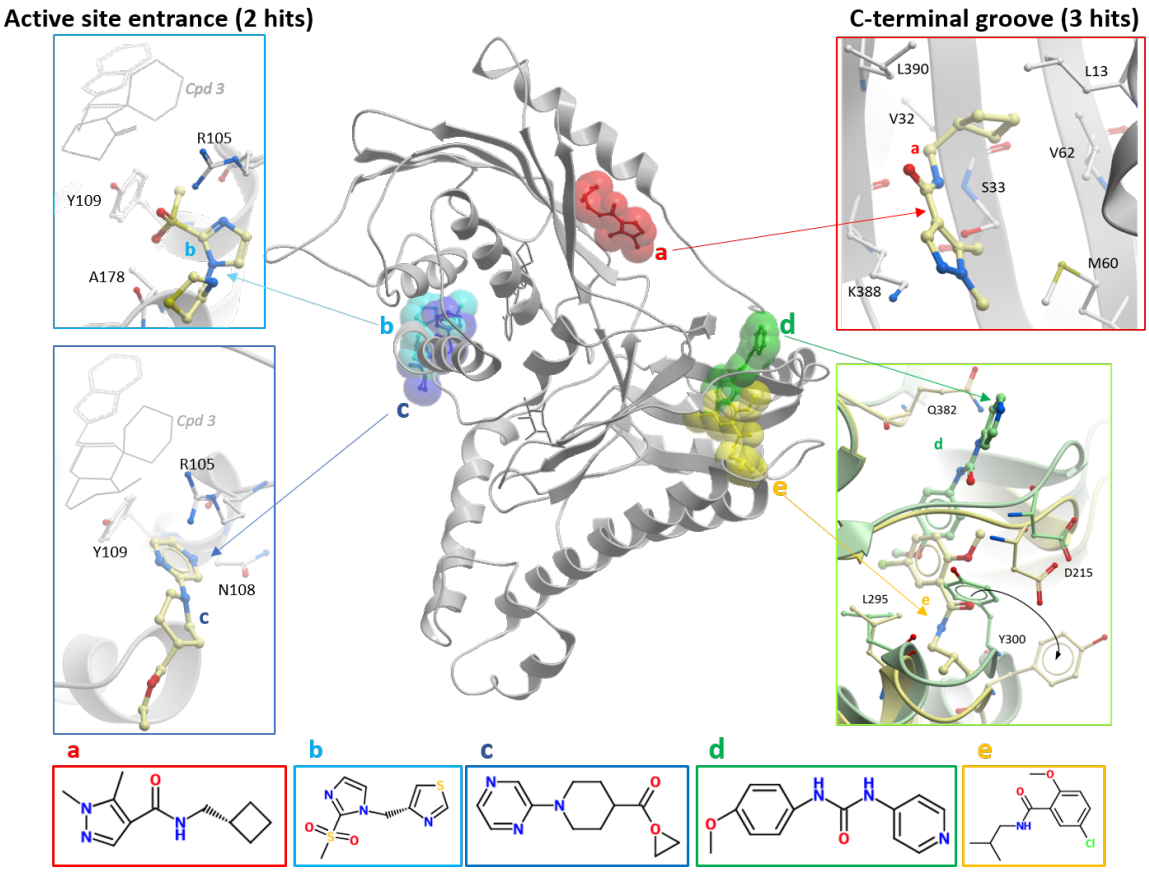
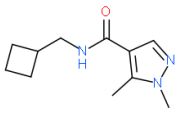
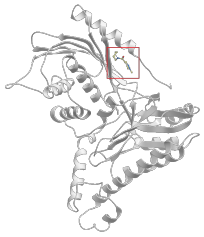
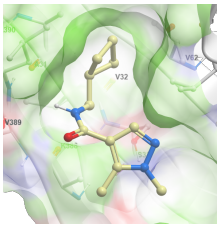
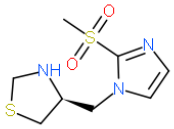
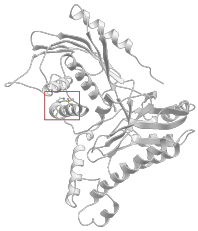
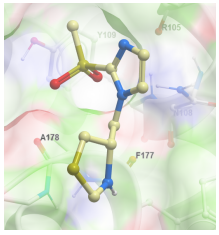
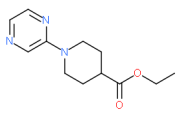
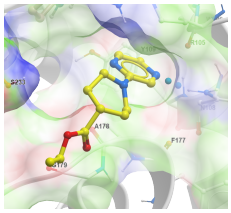


Figure 9 GALK1-bound fragments are clustered into three distinct sites. The 5 fragments described in this document are shown in spheres & sticks (their interactions with hGALK1 residues are shown in ***inset***).

PDBID	Ligand	Binding Location	Binding Pocket	Resolution (Å)
6Q8Z *(fragment **a**)*	 F9000534			2.4
6QJE *(fragment **b**)*	 F9000952			2.4
6Q3W *(fragment **c**)*	 DA000165			2.0

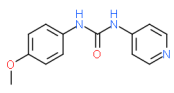
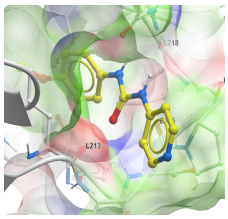
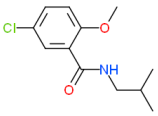
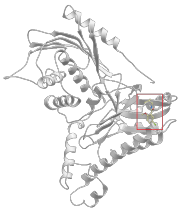
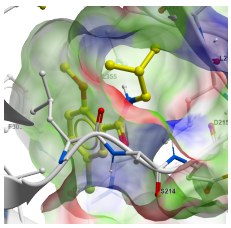
6Q90 *(fragment **d**)*	 XS092188			2.4
6Q91 *(fragment **e**)*	 FM010048			2.4

Table 1: Summary of fragment-bound GALK1 structures

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

This project first sets out to determine the novel crystal structure of hGALT, with a view to understanding the molecular basis of missense mutations causing Classic Galactosemia. The structural data provided the starting point for biophysical and biochemical characterization of several hGALT disease variant proteins, including the elucidation of the molecular mechanism for the p.Gln188Arg defect. The structural and mutagenesis studies of hGALT have now been published in Human Molecular Genetics, featuring in the issue's front cover (14).

Our publication has generated immediate impact – we have since been approached by researchers, clinicians and patient groups in the field of galactosemia to join efforts in developing substrate reduction therapy by targeting the upstream enzyme hGALK1, which becomes the second aim of this TEP project.

The co-crystal structures with the spiro-benzoxazole chemotypes in this TEP represent the first structural data available for hGALK1 chemical matters, providing a view of the binding modes for these early inhibitors and the opportunities in the active site for further chemistry design. We have established collaborations with the research teams in US (Kent Lai, University of Utah; Matthew Hall, NIH-NCATS) that originally identified this series of inhibitors. We are currently working with them towards structural characterization of their newer generations of compounds based on this scaffold.

In the past year we have received generous funding from the Galactosemia Foundation to initiate a fragment screen campaign (XChem) for hGALK1. Having now identified a way forward with reproducible crystals for XChem and completed one screening campaign, we plan to characterise the existing fragment hits in solution and to screen more drug-like molecule libraries using the same approach. For the novel binding sites identified (fragments **a**, **d** and **e**), we will use a combination of our existing assays (activity assay, SPR) to validate their interaction with hGALK1 in solution, molecular dynamics and hotspot computational methods to propose a mechanism of action, and additional assays (Tm shift assay, circular dichroism) to probe function of these sites. For the fragments (**b**, **c**) at the active site entrance, there is a path to pursue merging with existing ATP-analogue inhibitors (e.g. the existing spiro-benzoxazole series in (11)), to improve their affinity and potency.

Engagement with the patient organisation has also helped us reach out to galactosemia researchers to facilitate the next steps in hGALK1 drug discovery after our fragment screening campaign. We are also working with the group of galactosemia clinicians in Europe ('Galnet') who have the relevant patient-derived cell lines, metabolomic assays, and organism models that are key assets for downstream optimisation of our drug discovery project.

In short, this will be a TEP in joint efforts with patient and clinician groups.

FUNDING INFORMATION

The work performed at the SGC has been funded by a grant from the Wellcome [106169/ZZ14/Z], as well as the Galactosemia Foundation research grant (2017-2018).

ADDITIONAL INFORMATION

Structure Files

PDB ID	Structure Details
5IN3	Uridylylated hGALT bound with Glc-1-P
6GQD	hGALT with epitope mutations A21Y:A22T:T23P:R25L
6GR2	hGALK1 bound with galactose and ADP
6Q8Z	hGALK1 bound with galactose and **Fragment a** (N-(Cyclobutylmethyl)-1,5-dimethyl-1H-pyrazole-4-carboxamide, x717)
6QJE	hGALK1 bound with galactose and **Fragment b** (4-{[2-(Methylsulfonyl)-1H-imidazol-1-yl]methyl}-1,3-thiazole, x843)
6Q3W	hGALK1 bound with galactose and **Fragment c** (Ethyl 1-(2-pyrazinyl)-4-piperidinecarboxylate, x682)
6Q90	hGALK1 bound with galactose and **Fragment d** (1-(4-Methoxyphenyl)-3-(4-pyridinyl)urea, x747)
6Q91	hGALK1 bound with galactose and **Fragment e** (5-Chloro-N-isobutyl-2-methoxybenzamide, x841)

Materials and Methods

GALT – protein expression and purification

DNA fragment encoding full-length hGALT was amplified from a cDNA clone (IMAGE: 3922902) that harbours the p.Asn314Asp polymorphism (11), and subcloned into the pNIC28-Bsa4 vector that incorporates an N-terminal TEV-cleavable His6-tag. hGALT variants were created by either QuikChange site-directed mutagenesis or primer extension PCR method. hGALT (wt or variants) was cultured in 6 L of Terrific Broth at 37 °C, and induced with 0.1 mM IPTG overnight at 18 °C. Cell pellets were harvested, homogenized in lysis buffer (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP) and centrifuged to remove insoluble material. The supernatant was purified by immobilized metal affinity (Talon resin; GE Healthcare) and size-exclusion chromatography in Superdex 200 Hi-Load 16/60 column (GE Healthcare), pre-equilibrated with buffer 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP. Purified protein was treated with His-tagged TEV protease overnight at 4 °C, and further purified by reverse Nickel affinity and size exclusion chromatography. The final purified protein was concentrated to 20 mg/ml, flash cooled in liquid nitrogen and stored at −80 °C. To produce apo and uridylylated protein, 0.5 mM tag-removed hGALT was incubated overnight at 4 °C with 41 mM Glc-1-P or UDP-Glc before size-exclusion chromatography (Superdex 200 10/300 GL, GE Healthcare). Samples were analysed by denaturing mass spectrometry.

GALT – crystallization and structure determination

For the native hGALT protein, crystals were grown by vapour diffusion at 20 °C, from sitting drops mixing 200 nl of protein (20 mg/ml; pre-incubated with 5 m M UDP-Glc) and 100 nl of reservoir solution containing 0.2 M ammonium sulphate, 30% (w/v) PEG 8000. For the hGALT crystal epitope variant (A21Y:A22T:T23P:R25L), crystals were obtained at 4 °C in reservoir solution containing 0.2M ammonium sulphate and 30% PEG4000. Crystals were cryo-protected with reservoir solution supplemented with 25% (v/v) ethylene glycol and flash-cooled in liquid nitrogen. The structures were solved by molecular replacement with PHASER (15), using the structure of *E. coli* GALT (1HXP) as template. Modelling and refinement were carried out using Refmac (16) and Coot (17).

GALT – Rapid Fire Mass Spectrometry assay

In a typical experiment, 50 µL out of 540 µL assay buffer containing 0.1 µM hGALT (20 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, 0.5 mM TCEP) is injected prior to the addition of UDP-Glc. Next, 10 µL of UDP-Glc stock is added to reach the desired final concentration (0.1, 0.5, 1.0, 2.0 and 4.0 µM) of each experiment. The reactions were transferred to a RapidFire RF360 high-throughput sampling robot connected to an Agilent 6530 Accurate-Mass Quadrupole time-of-flight (Q-TOF) mass spectrometer operated in the positive ion mode (Agilent, Wakefield, MA, USA). The peak-area data of each injection were integrated and the percentage activity was calculated. All values were normalised using the percentage of uridylylation obtained from the

first time point of each experiment. The initial slope for enzyme progress curves was taken for each UDP-Glc concentration, and data were fitted to the Michaelis–Menten equation in GraphPad Prism 7.

GALT – Single-chain fragment variable antibody

Three scFv antibody fragments for hGALT have been generated via phage-display approach. The scFvs constructs are encoded in the pHAT6 vector with a LacZ promoter, OmpA leader and C-terminal Triple FLAG + His6 tags. Around 1 mg/L is expected from expression in 1.5 L scale from LB or TB media. The scFvs are isolated from cell lysate by affinity chromatography using HiTrap protein A, and dialyzed to 1xPBS overnight for storage.

	Kd from SPR Single Cycle Kinetics	**IP-MS Normalized Spectral Abundance Factor**
J-GALTA-2	1.88 nM	21
J-GALTA-3	2.28 nM	30
J-GALTA-4	5.85 nM	33

J-GALTA-2

EVQLLESGGGLVQPGGSLRLSCAASGFTFYSSYMYWVRQAPGKGLEWVSYISGGGSSTGYADSVKGRFTISRDNSKNTLYLQ
MNSLRAEDTAVYYCARESGSVGFMDYWGQGTLVTVSSGGGGSGGGGSGGGGSDIQMTQSPSSLSASVGDRVTITCRASQ
SISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYYYPSTFGQGTKLEIKRTDYK
DHDGDYKDHDIDYKDDDDKAAAHHHHHH*

J-GALTA-3

EVQLLESGGGLVQPGGSLRLSCAASGFTFSSSYMGWVRQAPGKGLEWVSSIGSYGSGTGYADSVKGRFTISRDNSKNTLYLQ
MNSLRAEDTAVYYCARSYTNYSLVSIGIFDYWGQGTLVTVSSGGGGSGGGGSGGGGSDIQMTQSPSSLSASVGDRVTITCR
ASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYYYPLTFGQGTKLEIKRT
DYKDHDGDYKDHDIDYKDDDDKAAAHHHHHH*

J-GALTA-4

EVQLLESGGGLVQPGGSLRLSCAASGFTFSYYGMYWVRQAPGKGLEWVSSISSSSSSTYYADSVKGRFTISRDNSKNTLYLQ
MNSLRAEDTAVYYCARYAGGYGIDYWGQGTLVTVSSGGGGSGGGGSGGGGSDIQMTQSPSSLSASVGDRVTITCRASQSI
SSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYSTYLLSTFGQGTKLEIKRTDYKD
HDGDYKDHDIDYKDDDDKAAAHHHHHH*

The three scFvs have been validated to bind endogenous hGALT using IP-MS, and display dissociation constants (Kd) in the range of 1.9 – 5.8 nM for recombinant hGALT as measured by SPR.

GALK1 – protein expression and purification

The GALK1 construct encodes full-length protein harbouring the surface entropy mutations K252A:E253A, with an engineered N-terminal His_6-tag subcloned into the pET21d vector. hGALK1 was cultured in Terrific Broth with 0.1 mM IPTG induction. Cell pellets were harvested, homogenized in lysis buffer (50 mM sodium phosphate pH 7.4, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP) and centrifuged to remove insoluble material. The supernatant was purified by Nickel affinity (Thermo Fisher Scientific) followed by ion exchange (Resource Q, GE Healthcare) and size exclusion (Superdex 200 Hi-Load 16/60, GE Healthcare) chromatography into crystallisation buffer (50 mM sodium phosphate pH7.4, 500 mM NaCl, 5% glycerol, 30mM Galactose and 0.5 mM TCEP). Protein was concentrated to 20 mg/mL and flash-cooled for storage at −80 °C.

Kinase-Glo activity assay for GALK1

GALK activity *in vitro* was determined using the Luminescent Kinase Assay (Kinase-Glo Promega), which detects ATP depletion, according to the manufacturer's protocol. Specifically, ten µL/well of reaction containing 5 nM GALK, 100 µM Galactose and 35 µM ATP in assay buffer (20 mM HEPES pH 7.5, 5 mM $MgCl_2$, 60 mM NaCl, 1 mM DTT, 0.01% BSA final concentration) as well as 100 µM of respective compounds was dispensed into 384-well assay plates (Greiner®). Following a 1 hour room temperature incubation, 10 µL of Kinase-Glo Plus detection reagent was added to provide an ATP-dependent luminescent readout (final assay volume: 20 µL/well). Luminescence was detected using a Pherastar Luminescence plate reader (BMG Labtech) after a10 minute incubation, using a 1 second exposure time and 2x binning. For IC50 determination, compounds were

serially diluted from 400 µM to 0.2 µM. Experiments were carried out as determined in the manufacturer protocol. Data were plotted in Graphpad Prism.

GALK1 – crystallization and structure determination

For co-crystallisation with ADP, 20mg/ml of GALK1 was pre-incubated with 10mM of ADP for 15 minutes on ice. Crystals were grown by vapour diffusion at 20 °C in 150 nL sitting drops equilibrated against well solution of 0.1M cacodylate pH 6.3, 17.5% PEG8K and 0.2M ammonium sulphate (optimised from Hampton Crystal Screen condition B3). Crystals were cryo-protected by addition of well solution supplemented with 25% ethylene glycol and flash-cooled before diffraction data collection at beamline I03 of the Diamond Light Source. The hGALK1-galactose-ADP structure was solved by molecular replacement using the published hGALK1 structure as template (PDB code: 1WUU)(10). PHENIX (18) and COOT were employed for refinement via iterative cycles.

GALK1 – crystallography-based fragment screening

To produce crystals for XChem, 24 mg/mL hGALK1 was pre-incubated with 0.75 mM of an ATP-analogue (Interbioscreen STOCK4S-53098; referred to in text as cpd 3 (11)) dissolved in NMP (final NMP concentration of 5%). Crystals were grown by vapour diffusion in 300 nL sitting drops at 20°C, equilibrated against well solutions of 0.1 M MOPS/sodium HEPES pH 7.0-7.5, 40-50 % Morpheus® Precipitant Mix 4 (50% mix = 12.5% MPD, 12.5% PEG1000, 12.5% PEG3350), 0.1 M Morpheus® Carboxylic acids mix (0.02M each of:sodium formate, ammonium acetate, sodium citrate tribasic dehydrate, sodium potassium tartrate tetrahydrate and sodium oxamate).

For soaking, 50 nL of each fragment compound (final concentration of 125 mM) was added to a crystallization 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 XChemXplorer pipeline (19) was used for structure solution with parallel molecular replacement using DIMPLE (20), followed by map averaging and statistical modelling to identify weak electron densities generated from low occupancy fragments using Pandda software (21). Coordinates and structure factors for exemplary data sets with bound fragments are deposited in the RCSB Protein Data Bank with the remainder to follow.

Non-SGC Resources

GALT	GALK1
Commercially available CRISPR/Cas9 knockout plasmids	
SCBT: Cat # sc-420482	SCBT: Cat # sc-420578
Genscript: Cat # 2592 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.	Addgene #76703, 76704, 76705 Nat Biotechnol. 2016, (2):184-191
Commercially available antibodies	
SCBT: sc-365577 (monoclonal)	SCBT: sc-393404 (monoclonal)

References

1. Holden, H. M., Rayment, I., and Thoden, J. B. (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J Biol Chem* **278**, 43885-43888
2. McCorvie, T. J., and Timson, D. J. (2011) The structural and molecular biology of type I galactosemia: Enzymology of galactose 1-phosphate uridylyltransferase. *IUBMB Life* **63**, 694-700
3. McCorvie, T. J., and Timson, D. J. (2011) Structural and molecular biology of type I galactosemia: disease-associated mutations. *IUBMB Life* **63**, 949-954
4. Waisbren, S. E., Potter, N. L., Gordon, C. M., Green, R. C., Greenstein, P., Gubbels, C. S., Rubio-Gozalbo, E., Schomer, D., Welt, C., Anastasoaie, V., D'Anna, K., Gentile, J., Guo, C. Y., Hecht, L.,

Jackson, R., Jansma, B. M., Li, Y., Lip, V., Miller, D. T., Murray, M., Power, L., Quinn, N., Rohr, F., Shen, Y., Skinder-Meredith, A., Timmers, I., Tunick, R., Wessel, A., Wu, B. L., Levy, H., Elsas, L., and Berry, G. T. (2012) The adult galactosemic phenotype. *J Inherit Metab Dis* **35**, 279-286

5. Calderon, F. R., Phansalkar, A. R., Crockett, D. K., Miller, M., and Mao, R. (2007) Mutation database for the galactose-1-phosphate uridyltransferase (GALT) gene. *Hum Mutat* **28**, 939-943
6. Tang, M., Odejinmi, S. I., Vankayalapati, H., Wierenga, K. J., and Lai, K. (2012) Innovative therapy for Classic Galactosemia - tale of two HTS. *Mol Genet Metab* **105**, 44-55
7. Jumbo-Lucioni, P., Parkinson, W., and Broadie, K. (2014) Overelaborated synaptic architecture and reduced synaptomatrix glycosylation in a Drosophila classic galactosemia disease model. *Dis Model Mech* **7**, 1365-1378
8. Ross, K. L., Davis, C. N., and Fridovich-Keil, J. L. (2004) Differential roles of the Leloir pathway enzymes and metabolites in defining galactose sensitivity in yeast. *Mol Genet Metab* **83**, 103-116
9. Bosch, A. M., Bakker, H. D., van Gennip, A. H., van Kempen, J. V., Wanders, R. J., and Wijburg, F. A. (2002) Clinical features of galactokinase deficiency: a review of the literature. *J Inherit Metab Dis* **25**, 629-634
10. Thoden, J. B., Timson, D. J., Reece, R. J., and Holden, H. M. (2005) Molecular structure of human galactokinase: implications for type II galactosemia. *J Biol Chem* **280**, 9662-9670
11. Liu, L., Tang, M., Walsh, M. J., Brimacombe, K. R., Pragani, R., Tanega, C., Rohde, J. M., Baker, H. L., Fernandez, E., Blackman, B., Bougie, J. M., Leister, W. H., Auld, D. S., Shen, M., Lai, K., and Boxer, M. B. (2015) Structure activity relationships of human galactokinase inhibitors. *Bioorg Med Chem Lett* **25**, 721-727
12. McAuley, M., Huang, M., and Timson, D. J. (2017) Insight into the mechanism of galactokinase: Role of a critical glutamate residue and helix/coil transitions. *Biochim Biophys Acta Proteins Proteom* **1865**, 321-328
13. Huang, M., Li, X., Zou, J. W., and Timson, D. J. (2013) Role of Arg228 in the phosphorylation of galactokinase: the mechanism of GHMP kinases by quantum mechanics/molecular mechanics studies. *Biochemistry* **52**, 4858-4868
14. McCorvie, T. J., Kopec, J., Pey, A. L., Fitzpatrick, F., Patel, D., Chalk, R., Shrestha, L., and Yue, W. W. (2016) Molecular basis of classic galactosemia from the structure of human galactose 1-phosphate uridylyltransferase. *Hum Mol Genet* **25**, 2234-2244
15. 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
16. 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
17. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132
18. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L. W., Jain, S., Kapral, G. J., Grosse Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94-106
19. 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
20. 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., Potterton, 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
21. 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.