

Galactokinase promiscuity: a question of flexibility?

Megan McAuley¹, Helena Kristiansson^{1,2}, Meilan Huang³, Angel L. Pey⁴ and David J. Timson^{1,5*}

¹ School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast, BT9 7BL. UK.

² Current address: Yale University School of Medicine, 310 Cedar Street, New Haven, CT 06510. USA.

³ School of Chemistry and Chemical Engineering, Queen's University, Belfast, David Kier Building, Stranmillis Road, Belfast, BT9 5AG. UK.

⁴ Department of Physical Chemistry, Faculty of Sciences, University of Granada, Av. Fuentenueva s/n, 18071, Spain.

⁵ School of Pharmacy and Biomolecular Sciences, University of Brighton, Huxley Building, Lewes Road, Brighton BN2 4GJ. UK.

* Corresponding author

School of Pharmacy and Biomolecular Sciences, University of Brighton, Huxley Building, Lewes Road, Brighton BN2 4GJ. UK.

d.timson@brighton.ac.uk

Abstract

Galactokinase catalyses the first committed step of the Leloir pathway, i.e. the ATP-dependent phosphorylation of α -D-galactose at C₁-OH. Reduced galactokinase activity results in the inherited metabolic disease type II galactosemia. However, inhibition of galactokinase is considered a viable approach to treating more severe forms of galactosemia (types I and III). Considerable progress has been made in the identification of high affinity, selective inhibitors. Although the structure of galactokinase from a variety of species is known, its catalytic mechanism remains uncertain. While the bulk of the evidence suggests that the reaction proceeds *via* an active site base mechanism, some experimental and theoretical studies contradict this. The enzyme has potential as a biocatalyst in the production of sugar 1-phosphates. This potential is limited by its high specificity. A variety of approaches have been taken to identify galactokinase variants which are more promiscuous. These have broadened galactokinase's specificity to include a wide range of D- and L-sugars. Initial studies suggest that some of these alterations result in increased flexibility at the active site. It is suggested that modulation of protein flexibility is at least as important as structural modifications in determining the success or failure of enzyme engineering.

Keywords: Enzyme engineering, galactosemia, GHMP kinase, specificity, galactose, enzyme mechanism

Introduction: Galactokinase

Galactokinase (galactose kinase, E.C. 2.7.1.6) is a member of the GHMP kinase family. This family of small molecule kinases phosphorylate a variety of small molecule substrates at the expense of ATP. The GHMP kinase family is named after the four originally identified members: galactokinase, homoserine kinase, mevalonate kinase and phosphomevalonate kinase [1,2]. Further members of the superfamily have been discovered subsequently and these include N-acetylgalactosamine kinase, galacturonic acid kinase, arabinose kinase and pantoate kinase [3-7]. Two non-enzymatic family members – the *Caenorhabditis elegans* sex-fate determining protein XOL-1 and *Saccharomyces cerevisiae* transcriptional regulator Gal3p – have also been recognised as family members [8-10]. Proteins in the family have three highly conserved motifs. Motif I is involved in binding of the small molecule substrate, motif II in binding of the nucleotide triphosphate and motif III in catalysis [1]. The sequences in between these conserved motifs vary in terms of residues and length but the overall fold of these proteins is similar [11].

Galactokinase catalyses the first committed step of the Leloir pathway, the magnesium dependent phosphorylation of α -D-galactose at position 1 of the sugar ring [12,13]. The function of the Leloir pathway is to convert galactose into glucose 1-phosphate (Figure 1). This compound is then isomerised into galactose 6-phosphate which can enter glycolysis. The pathway is especially important in young mammals since the main sugar in milk is lactose, a disaccharide of glucose and galactose.

Galactokinase deficiency: type II galactosemia

Dysfunction of any of the enzymes encoding enzymes of the Leloir pathway results in the autosomal recessive disorder galactosemia (Figure 2). This inherited metabolic disorder presents as a decreased ability to metabolise galactose [14,15]. There are three types of galactosemia each resulting from a mutation in a gene encoding a different enzyme of galactose metabolism. Type I galactosemia or classical galactosemia (OMIM #230400) is caused by mutations in the gene encoding galactose-1-phosphate uridylyltransferase (GALT; EC 2.7.7.12). Short term symptoms include vomiting, diarrhoea and sepsis. Longer term symptoms include speech impairments, cognitive disabilities, female reproductive problems and liver problems [14,16]. Approximately 200 disease-associated variants of GALT are known with these alterations spread throughout the sequence and structure [16,17].

Type II galactosemia (OMIM#230200) results from mutations in the *GALK1* gene which encodes galactokinase [18,19]. These mutations can be insertions, deletions and single base changes [20-25]. The main symptom of this type of galactosemia is early onset cataracts [20]. The condition is treated by surgery to remove the cataracts combined with the exclusion of galactose and lactose from the diet. Since galactose cannot be metabolised it builds up in the cells of patients. In the lens cells of the eye, galactose is efficiently converted to its corresponding alcohol galactitol (dulcitol) by the action of aldose reductase (EC 1.1.1.21). It has been shown in a mouse model that galactitol builds up, changing osmotic pressure and causing cell lysis [26]. No long-term complications have been definitively associated with this condition [27]. Thus, type II galactosemia is regarded as the most benign of the three types. *In vitro* studies on purified, recombinant galactokinase demonstrated a rough correlation between disease severity and effects on the enzyme [25,28]. In general, those variations associated with the most severe forms of the disease (cataracts developing within the first few months of life) resulted in proteins which were insoluble on expression in *Escherichia coli*. Consequently, it was not possible to isolate these proteins and study their enzymological properties. However, this suggests that these variants are misfolded and may well aggregate into insoluble, inactive complexes *in vivo* [28]. Those variants associated with

intermediate phenotypes (cataracts developing childhood) were soluble on expression in *E. coli* and had impaired kinetic parameters compared to the wild-type protein [28]. One variant, p.A198V, had enzyme kinetic properties which were essentially identical to the wild-type [28]. This variant is relatively common in East Asian populations and is associated with an increased risk of cataracts in mid-life [24]. Cells expressing this variant have lower amounts of galactokinase protein compared to those expressing the wild-type enzyme. This suggests that the protein is less stable under cellular conditions [24].

Type III galactosemia (OMIM#230250) results from mutations in the gene encoding UDP-galactose 4'-epimerase (GALE; EC 5.1.3.3) GALE [29]. This enzyme converts UDP-galactose to UDP-glucose (Figure 1). This is most likely the rarest form of galactosemia and the disease manifestations vary greatly in severity. The mildest forms are almost asymptomatic, resulting only in alterations in blood chemistry whereas the most severe cases are comparable to the worst cases of type I galactosemia. Complete removal of galactose from the diets of these patients is not advised. Since GALE is defective in this disease, UDP-galactose cannot be produced from glucose. This compound is required for the biosynthesis of glycolipids and glycoproteins. It is important therefore that galactose in the diet is present but carefully controlled [14].

The molecular pathology of galactosemia remains somewhat mysterious. There is considerable evidence that galactose 1-phosphate accumulation is toxic to cells, although the mechanism of this toxicity is unknown [30,31]. It has been proposed that inhibition of galactokinase may be beneficial in type I galactosemia (and severe forms of type III galactosemia) since this would reduce the accumulation of galactose 1-phosphate [27]. Potentially, this would “convert” type I and type III galactosemia into the much less severe type II (Figure 2). If this hypothesis is correct then these two life-threatening conditions which result in significant physical and cognitive disabilities could be significantly mitigated. It is likely that patients would have to continue on galactose (and lactose) restricted diets and may also require cataract removal at some point. However, they may be able to lead otherwise normal lives.

The identification of high affinity, selective inhibitors of galactokinase is challenging. The only naturally occurring inhibitors known are the substrate galactose and the product galactose 1-phosphate [28,32]. Neither would be suitable therapeutically and both have relatively low affinities. The similarity of the GHMP kinase family members also presents a problem. While there is some variation in the small molecule substrate site, the ATP binding site is well-conserved. A particular problem is presented by *N*-acetylgalactosamine kinase (GALK2; EC 2.7.1.157). This enzyme has a highly similar active site and also has low levels of activity against galactose [33,34]. The similarity to other GHMP kinases, especially GALK2, results in a risk of off-target effects. However, any drug used to treat galactosemia needs to be well-tolerated over a long period of time. It is likely that patients would need to take the drug regularly throughout childhood and, quite possibly, for the duration of their lives.

Through high-throughput screening a number of promising inhibitors of galactokinase have been identified [35,36]. An initial screen identified 34 compounds, with a variety of chemical scaffolds which had micromolar IC₅₀ values [37]. Three of these compounds showed selectively for galactokinase over mevalonate kinase and were also capable of reducing galactose 1-phosphate levels in cells [38]. A second screen identified another 200 inhibitors. From these a 4-oxo-3,4-dihydro-2*H*-1,3-thiazine-5-carbonitrile scaffold was used as the basis for further optimisation resulting in compounds with IC₅₀ values in the low micromolar range [39]. Subsequent work resulted in the discovery of a spiro-benzoxazole compound which inhibited galactokinase, but had no activity against GALK2 [40]. Work on the identification of galactokinase inhibitors has also been assisted by molecular dynamics simulations of protein-inhibitor complexes [41].

Structure and mechanism of galactokinase

The crystal structure of human galactokinase has been solved. The enzyme has two domains: an N-terminal domain with a long α -helix composed of 15 residues alongside a 6-stranded mixed β -sheet and a C-terminal domain with six α -helices and two layers of β -sheet each containing four strands [42]. Structures from other species have also been solved, including *Saccharomyces cerevisiae*, *Pyrococcus furiosus* and *Lactococcus lactis* [43-45]. These proteins, from widely different parts of the tree of life, have broadly similar structures to the human enzyme.

Human galactokinase is specific to α -D-galactose, 2-deoxy- α -D-galactose and α -D-galactosamine [46]. It is believed that only the first of these three compounds is a physiologically relevant substrate. Interestingly, galactokinases from different species have different kinetic mechanisms. Human, pig, rat, yeast and *Bifidobacterium bifidum* galactokinases all operate *via* an ordered ternary complex mechanism in which ATP binds first [28,47-50]. The enzymes from fenugreek seeds and *Vicia fabia* (broad bean) also have ordered mechanisms; however, galactose binds first [51,52]. It is interesting to speculate that this might be the case for all plant galactokinases. In the case of *Escherichia coli* galactokinase, the enzyme operates *via* a random ternary complex mechanism in which the two substrates can bind in any order [53]. It is not currently known what structural and dynamic features of the enzyme result in these different mechanisms. Only very limited changes have been detected when x-ray crystal structures of the enzyme with and without substrates are compared [42].

GHMP kinases appear to act by at least two different catalytic mechanisms [2,20]. The majority are thought to act through a catalytic base in the active site (e.g. rat mevalonate kinase [54]). In this mechanism, an aspartate or glutamate residue in the active site removes a proton from the substrate (generally from a hydroxyl group). This renders the substrate negatively charged, highly nucleophilic and thus able to attack the γ -phosphate of ATP. This attack breaks the bond linking the β - and γ -phosphates of ATP, transferring the γ -phosphate to the substrate. The active site base can be regenerated through loss of the proton to water, thus restoring the enzyme to its original state. However, homoserine kinase lacks this active site base and has an asparagine residue at the structurally equivalent site [55]. It is hypothesised that, in this enzyme, the active site promotes stabilisation of the transition state, enabling catalysis [55]. It is also possible that one of the negatively charged oxygens on the γ -phosphate of the ATP acts as a base, wholly or partially abstracting a proton from homoserine [20]. The existence of this alternative mechanism suggests that it is possible that transition state stabilisation also plays a role in other GHMP kinase family members, including galactokinase.

The catalytic mechanism of human galactokinase is currently uncertain. However, the structure of the active site resembles that of mevalonate kinase with an aspartate (Asp-186 in the human enzyme) appropriately positioned to act as a catalytic base. As in many other GHMP kinase family members, this putative base is located close to a positively charged residue, Arg-37. Although this active site architecture suggests a similar mechanism to mevalonate kinase, there is some evidence to the contrary. In the *S. cerevisiae* enzyme there was no major effects of pH on the steady-state kinetic parameters and no detectable deuterium kinetic isotope effect [49]. No deuterium kinetic isotope effect was detected with the human enzyme either [28]. These results imply that proton transfer is unlikely to be an important part of the catalytic mechanism. However, both results can be confounded by other effects and, consequently, should be interpreted with caution. The aspartate residue has been shown to be essential to catalysis in both rat and human galactokinase [38,48,56]. However, these results also require caution. While alteration of Asp-186 in the human enzyme to either alanine or asparagine reduced activity to below detectable levels, both changes also destabilised the enzyme towards denaturation by the chaotropic compound urea [56]. Arg-37 was also shown to be essential, but again alteration of this to alanine reduced the stability of the enzyme [56]. Thus these losses of activity could result from either removal of key chemical groups in the

active site or a failure of the protein to fold correctly. Interestingly, alteration of Arg-37 to lysine resulted in reduced, but not zero, activity suggesting that a positive charge at this position is required [56]. A detailed study on the effects of pH on the kinetic parameters of the *L. lactis* galactokinase showed effects consistent with an active site base mechanism, in which the pK_a of the active site aspartate is approximately 6.9 (compared to its free solution value of 3.9) [57]. The role of the adjacent arginine residue, Arg-36, was proposed to be to reduce the pK_a of the C₁-OH to a value which enables efficient proton transfer (Figure 3) [57]. However, QM/MM studies on the mechanism of human galactokinase have identified three residues which are proposed to stabilise MgATP in the active site, Glu-174, Arg-228 and Arg-105. These studies also suggested that the enzyme operates via a direct phosphorylation mechanism [58]. It seems unlikely that the bacterial and human enzymes act by radically different catalytic mechanisms, especially considering the high degree of similarities of their active sites [42,44] (Figure 3). It is possible that elements of all proposed mechanisms occur, with the enzyme manipulating the pK_a of active site residues and the substrates, stabilising the transition state and also acting on the MgATP substrate. To date, the role of the sugar's conformation and dynamics in the mechanism has been underexplored and we suggest that this may warrant further study.

Galactokinase: potential role in biotechnology

Sugar 1-phosphates are important precursors for the synthesis of some drugs and drug-like molecules, including aminoglycoside antibiotics. Although these occur naturally and can be purified from these sources, semi- or fully synthetic variants may also be useful. Chemical synthesis of glycosylated compounds is technically challenging. In part, this arises from the similar reactivity of the hydroxyl groups present in sugars. In the case of the pyranose ring form of galactose, there are six hydroxyl groups and directing chemical modifications to just one of these is difficult with conventional chemical agents. However, galactokinase is highly site-specific in the reactions it catalyses – only α -D-galactose 1-phosphate is produced. However, it is also specific to galactose and closely related sugars. In the case of the human enzyme only a small number of very closely related structures can act as substrates. Although some bacterial galactokinases have wider specificity, the range of possible substrates is still somewhat limited [59].

One important methodology which relies on sugar 1-phosphates is *in vitro* glycorandomisation (IVG) [60,61]). This process was developed for the production of unique variants of vancomycin (and other antibiotics), a drug used to treat methicillin resistant *Staphylococcus aureus* (MRSA). The aim was to produce novel variants of the drug with enhanced activity. However, IVG has potential as a process for the production of novel glycosylated products [62]. Promiscuous nucleotidyltransferases (E_p) and glycosyltransferases have been engineered but a promiscuous sugar 1-kinase is also required [60]. Along with glycogen phosphorylases and fucokinases, galactokinases are a group of enzymes which catalyses phosphorylation at position 1 of the sugar ring. Consequently, there have been efforts made to broaden the specificity of galactokinase while retaining a level of activity consistent with its use in industrial processes.

Human galactokinase will tolerate small changes to the sugar substrate at carbon-2 but not at carbons 4 and 6. Even at position 2, an *N*-acetyl group appears to be too large to be accommodated and so *N*-acetylgalactosamine is not a substrate [46]. Asp-46, Tyr-236 and Glu-43 are partly responsible for anchoring the sugar in the active site; Gly-183 and His-44 form hydrogen bonds with the sugar [42]. Alteration of Glu-43 and His-44 to glycine and isoleucine respectively in order to match the equivalent sequence in arabinose kinase resulted in insoluble protein and alterations at Asp-46 resulted in a soluble enzyme with no detectable activity [46]. Taken together with the data on disease-associated variants, these results suggest that human galactokinase is not very tolerant to even single amino acid changes.

Although *E. coli* galactokinase has wider substrate specificity than the human enzyme, it is still limited having little or no activity towards sugars with alterations at carbons 3, 4 or 6 [63]. Directed evolution experiments carried out on *E. coli* galactokinase identified a tyrosine residue (Tyr-371), which is distant from the active site and alteration of this residue to histidine conferred increased substrate promiscuity. The increased range of substrates included both D- and L-sugars (e.g. D-galacturonic acid, D-talose, L-altrose, and L-glucose) [64]. Alteration of another tyrosine residue (Tyr-233; equivalent to Tyr-236 in the human enzyme) also resulted in greater promiscuity [65]. *Lactococcus lactis* galactokinase has low activity against D-glucose. Comparison of the active sites of *E. coli* and *L. lactis* galactokinases suggested that this was due to the presence of a methionine residue (Met-173) in the *E. coli* enzyme's active site. Alteration of this to leucine (the equivalent residue in *L. lactis* galactokinase) resulted in the enzyme becoming active towards D-glucose. Combining this p.M173L variant with the previously described p.Y371H variant resulted in an enzyme capable of catalysing the phosphorylation of at least 28 different monosaccharides [66]. Bacteria expressing this variant enzyme were able to produce a variety of sugar 1-phosphates when fed with the appropriate precursor [66].

Interestingly alteration of the equivalent tyrosine in the human enzyme (Tyr-379) to histidine did not increase substrate range but alteration to a tryptophan, lysine, arginine, serine or cysteine did albeit with some reduction in overall activity [67]. The ability of a residue distant from the active site to influence specificity suggests that the region around Tyr-379 is able to influence the active site structure or dynamics. Preliminary molecular modelling studies suggested that this may occur by a modulation of the flexibility of the active site. Although this requires verification experimentally and from more thorough molecular dynamics studies, it is consistent with the experimental data: a more flexible active site would be able to accommodate a wider range of molecules, but the increased mobility of active site residues may reduce the catalytic efficiency and activity [67]. Interestingly, human GALK2 (which has a phenylalanine residue at the equivalent position to Tyr-379) naturally has relatively broad specificity being able to catalyse the phosphorylation of *N*-acetylglucosamine and *N*-acetylmannosamine in addition to its natural substrate, *N*-acetylgalactosamine [68,69]. The active site of GALK2 is predicted to be slightly more flexible than the one in galactokinase [68].

Conclusions: is flexibility the key?

It is now clear that the dynamic behaviour of proteins is critically important in determining their function [70]. Traditionally, enzyme engineering has focussed on the alteration of structural aspects of the protein (e.g. the modification of active site residues to change the shape and/or chemical reactivity of that space). However, the substrate specificity of enzymes can also be influenced by the dynamics and it appears that this may be the case with galactokinase. Increasing the flexibility of the active site seems to increase the range of substrates processed. Further experimental studies are required to determine the nature of this increased flexibility. Exploitation of this enzyme as a biocatalyst will require more than just increasing its promiscuity. Maintaining a reasonably high activity level will also be required, preferably combined with an enzyme which is stable for long periods of time under the conditions of industrial synthesis. Paradoxically, it may be necessary to simultaneously increase the flexibility of the active site, while increasing the global, overall stability of the protein. In pursuit of the latter goal, consensus methods may be a viable method to achieve increased stability without loss of activity [71-73]. Overall, it is clear that enzyme engineering methods which rely on altering structure may ultimately be limited: more successful outcomes may be achieved where local and global flexibility are also considered.

Funding

MM and HK thank the Department of Employment and Learning (Northern Ireland, UK) and the Environmental Protection Agency (Republic of Ireland) [2007-PhD-ET-9] respectively for PhD studentships. MM and DJT thank the Biochemical Society for financial assistance which enabled their attendance of the focussed meeting on Carbohydrate Active Enzymes in Medicine and Biotechnology (St Andrews, August 2015). Work in DJT's laboratory on Leloir pathway enzymes was partly funded by the Royal Society (UK; [2004/R1]). ALP acknowledges support from MINECO [CSD2009-00088] and [BIO2012-34937], Junta de Andalucia [CTS-11-07187] and FEDER Funds. ALP is a recipient of a Ramon y Cajal contract from MINECO [RYC2009-04147].

References

- 1 Bork, P., Sander, C. and Valencia, A. (1993) Convergent evolution of similar enzymatic function on different protein folds: The hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* **2**, 31-40
- 2 Timson, D. J. (2007) GHMP kinases - structures, mechanisms and potential for therapeutically relevant inhibition. *Curr. Enz. Inhib.* **3**, 77-94
- 3 Pastuszak, I., O'Donnell, J. and Elbein, A. D. (1996) Identification of the GalNAc kinase amino acid sequence. *J. Biol. Chem.* **271**, 23653-23656
- 4 Pastuszak, I., Drake, R. and Elbein, A. D. (1996) Kidney N-acetylgalactosamine (GalNAc)-1-phosphate kinase, a new pathway of GalNAc activation. *J. Biol. Chem.* **271**, 20776-20782
- 5 Yang, T., Bar-Peled, L., Gebhart, L., Lee, S. G. and Bar-Peled, M. (2009) Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants, and comparison with galactose-1-phosphate kinase. *J. Biol. Chem.* **284**, 21526-21535.
- 6 Tomita, H., Yokooji, Y., Ishibashi, T., Imanaka, T. and Atomi, H. (2012) Biochemical characterization of pantoate kinase, a novel enzyme necessary for coenzyme A biosynthesis in the archaea. *J. Bacteriol.* **194**, 5434-5443.
- 7 Sherson, S., Gy, I., Medd, J., Schmidt, R., Dean, C., Kreis, M., Lecharny, A. and Cobbett, C. (1999) The arabinose kinase, *ARA1*, gene of *Arabidopsis* is a novel member of the galactose kinase gene family. *Plant Mol. Biol.* **39**, 1003-1012
- 8 Bajwa, W., Torchia, T. E. and Hopper, J. E. (1988) Yeast regulatory gene *GAL3*: Carbon regulation; UASGal elements in common with *GAL1*, *GAL2*, *GAL7*, *GAL10*, *GAL80*, and *MEL1*; encoded protein strikingly similar to yeast and *escherichia coli* galactokinases. *Mol. Cell. Biol.* **8**, 3439-3447
- 9 Lavy, T., Kumar, P. R., He, H. and Joshua-Tor, L. (2012) The Gal3p transducer of the *GAL* regulon interacts with the Gal80p repressor in its ligand-induced closed conformation. *Genes Dev.* **26**, 294-303.
- 10 Luz, J. G., Hassig, C. A., Pickle, C., Godzik, A., Meyer, B. J. and Wilson, I. A. (2003) XOL-1, primary determinant of sexual fate in *C. elegans*, is a GHMP kinase family member and a structural prototype for a class of developmental regulators. *Genes Dev.* **17**, 977-990.
- 11 Andreassi, J. L., 2nd and Leyh, T. S. (2004) Molecular functions of conserved aspects of the GHMP kinase family. *Biochemistry.* **43**, 14594-14601.
- 12 Trucco, R. E., Caputto, R., Leloir, L. F. and Mittleman, N. (1948) Galactokinase. *Arch. Biochem.* **18**, 137-146
- 13 Frey, P. A. (1996) The Leloir pathway: A mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J.* **10**, 461-470
- 14 Fridovich-Keil, J. L. and Walter, J. H. (2008) Galactosemia. In *The Online Metabolic and Molecular Bases of Inherited Diseases* (Valle, D., Beaudet, A. L., Vogelstein, B., Kinzler, K. W., Antonarakis, S. E. and Ballabio, A., eds.), McGraw-Hill, New York
- 15 Timson, D. J. (2015) The molecular basis of galactosemia - past, present and future. *Gene*. **In press**.

- 16 McCorvie, T. J. and Timson, D. J. (2011) Structural and molecular biology of type I galactosemia: Disease-associated mutations. *IUBMB Life*. **63**, 949-954.
- 17 d'Acierno, A., Facchiano, A. and Marabotti, A. (2014) GALT protein database: Querying structural and functional features of GALT enzyme. *Hum. Mutat.* **35**, 1060-1067.
- 18 Gitzelmann, R. (1967) Hereditary galactokinase deficiency, a newly recognized cause of juvenile cataracts. *Pediatr. Res.* **1**, 14-23
- 19 Timson, D. J., Reece, R. J., Thoden, J. B. and Holden, H. M. (2009) Galactokinase deficiency. In *Encyclopedia of Molecular Mechanisms of Disease* (Lang, F., ed.), pp. 679-680, Springer, Berlin and Heidelberg
- 20 Holden, H. M., Thoden, J. B., Timson, D. J. and Reece, R. J. (2004) Galactokinase: Structure, function and role in type II galactosemia. *Cell Mol. Life Sci.* **61**, 2471-2484.
- 21 Kalaydjieva, L., Perez-Lezaun, A., Angelicheva, D., Onengut, S., Dye, D., Bosshard, N. U., Jordanova, A., Savov, A., Yanakiev, P., Kremensky, I., Radeva, B., Hallmayer, J., Markov, A., Nedkova, V., Tournev, I., Aneva, L. and Gitzelmann, R. (1999) A founder mutation in the GK1 gene is responsible for galactokinase deficiency in roma (gypsies). *Am. J. Hum. Genet.* **65**, 1299-1307.
- 22 Stambolian, D., Ai, Y., Sidjanin, D., Nesburn, K., Sathe, G., Rosenberg, M. and Bergsma, D. J. (1995) Cloning of the galactokinase cDNA and identification of mutations in two families with cataracts. *Nat. Genet.* **10**, 307-312
- 23 Kolosha, V., Anoaia, E., de Cespedes, C., Gitzelmann, R., Shih, L., Casco, T., Saborio, M., Trejos, R., Buist, N., Tedesco, T., Skach, W., Mitelmann, O., Ledee, D., Huang, K. and Stambolian, D. (2000) Novel mutations in 13 probands with galactokinase deficiency. *Hum. Mutat.* **15**, 447-453.
- 24 Okano, Y., Asada, M., Fujimoto, A., Ohtake, A., Murayama, K., Hsiao, K. J., Choeh, K., Yang, Y., Cao, Q., Reichardt, J. K., Niihira, S., Imamura, T. and Yamano, T. (2001) A genetic factor for age-related cataract: Identification and characterization of a novel galactokinase variant, "osaka" in asians. *Am. J. Hum. Genet.* **68**, 1036-1042.
- 25 Sangiuolo, F., Magnani, M., Stambolian, D. and Novelli, G. (2004) Biochemical characterization of two GALK1 mutations in patients with galactokinase deficiency. *Hum. Mutat.* **23**, 396.
- 26 Ai, Y., Zheng, Z., O'Brien-Jenkins, A., Bernard, D. J., Wynshaw-Boris, T., Ning, C., Reynolds, R., Segal, S., Huang, K. and Stambolian, D. (2000) A mouse model of galactose-induced cataracts. *Hum. Mol. Genet.* **9**, 1821-1827
- 27 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. Inher. Metab. Dis.* **25**, 629-634
- 28 Timson, D. J. and Reece, R. J. (2003) Functional analysis of disease-causing mutations in human galactokinase. *Eur. J. Biochem.* **270**, 1767-1774.
- 29 Timson, D. J. (2006) The structural and molecular biology of type III galactosemia. *IUBMB Life*. **58**, 83-89.
- 30 Slepak, T., Tang, M., Addo, F. and Lai, K. (2005) Intracellular galactose-1-phosphate accumulation leads to environmental stress response in yeast model. *Mol. Genet. Metab.* **86**, 360-371.

- 31 de Jongh, W. A., Bro, C., Ostergaard, S., Regenber, B., Olsson, L. and Nielsen, J. (2008) The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in *saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **101**, 317-326.
- 32 Ballard, F. J. (1975) Galactokinase from pig liver. *Methods Enzymol.* **42**, 43-47
- 33 Thoden, J. B. and Holden, H. M. (2005) The molecular architecture of human N-acetylgalactosamine kinase. *J. Biol. Chem.* **280**, 32784-32791.
- 34 Agnew, A. and Timson, D. (2010) Mechanistic studies on human N-acetylgalactosamine kinase. *J. Enzyme Inhib. Med. Chem.* **25**, 370-376.
- 35 Lai, K., Boxer, M. B. and Marabotti, A. (2014) GALK inhibitors for classic galactosemia. *Future Med. Chem.* **6**, 1003-1015.
- 36 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.
- 37 Wierenga, K. J., Lai, K., Buchwald, P. and Tang, M. (2008) High-throughput screening for human galactokinase inhibitors. *J. Biomol. Screen.* **13**, 415-423.
- 38 Tang, M., Wierenga, K., Elsas, L. J. and Lai, K. (2010) Molecular and biochemical characterization of human galactokinase and its small molecule inhibitors. *Chem. Biol. Interact.* **188**, 376-385.
- 39 Odejinmi, S., Rascon, R., Tang, M., Vankayalapati, H. and Lai, K. (2011) Structure-activity analysis and cell-based optimization of human galactokinase inhibitors. *ACS Med. Chem. Lett.* **2**, 667-672.
- 40 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.
- 41 Chiappori, F., Merelli, I., Milanese, L. and Marabotti, A. (2013) Static and dynamic interactions between GALK enzyme and known inhibitors: Guidelines to design new drugs for galactosemic patients. *Eur. J. Med. Chem.* **63**, 423-434.
- 42 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.
- 43 Thoden, J. B., Sellick, C. A., Timson, D. J., Reece, R. J. and Holden, H. M. (2005) Molecular structure of *saccharomyces cerevisiae* Gal1p, a bifunctional galactokinase and transcriptional inducer. *J. Biol. Chem.* **280**, 36905-36911.
- 44 Thoden, J. B. and Holden, H. M. (2003) Molecular structure of galactokinase. *J. Biol. Chem.* **278**, 33305-33311.
- 45 Hartley, A., Glynn, S. E., Barynin, V., Baker, P. J., Sedelnikova, S. E., Verhees, C., De Geus, D., Van Der Oost, J., Timson, D. J., Reece, R. J. and Rice, D. W. (2004) Substrate specificity and mechanism from the structure of *pyrococcus furious* galactokinase. *J. Mol. Biol.* **337**, 387-398.
- 46 Timson, D. J. and Reece, R. J. (2003) Sugar recognition by human galactokinase. *BMC Biochem.* **4**, 16.
- 47 Ballard, F. J. (1966) Kinetic studies with liver galactokinase. *Biochem. J.* **101**, 70-75

- 48 Chu, X., Li, N., Liu, X. and Li, D. (2009) Functional studies of rat galactokinase. *J. Biotechnol.* **141**, 142-146.
- 49 Timson, D. J. and Reece, R. J. (2002) Kinetic analysis of yeast galactokinase: Implications for transcriptional activation of the *GAL* genes. *Biochimie.* **84**, 265-272.
- 50 Lee, L., Kinoshita, S., Kumagai, H. and Tochikura, T. (1980) Galactokinase of *bifidobacterium bifidum*. *Agric. Biol. Chem.* **44**, 2961-2966
- 51 Foglietti, M. J. and Percheron, F. (1976) [Purification and mechanism of action of a plant galactokinase]. *Biochimie.* **58**, 499-504
- 52 Dey, P. M. (1983) Galactokinase of *vicia faba* seeds. *Eur. J. Biochem.* **136**, 155-159
- 53 Gulbinsky, J. S. and Cleland, W. W. (1968) Kinetic studies of *escherichia coli* galactokinase. *Biochemistry.* **7**, 566-575
- 54 Fu, Z., Wang, M., Potter, D., Miziorko, H. M. and Kim, J. J. (2002) The structure of a binary complex between a mammalian mevalonate kinase and ATP: Insights into the reaction mechanism and human inherited disease. *J. Biol. Chem.* **277**, 18134-18142.
- 55 Krishna, S. S., Zhou, T., Daugherty, M., Osterman, A. and Zhang, H. (2001) Structural basis for the catalysis and substrate specificity of homoserine kinase. *Biochemistry.* **40**, 10810-10818.
- 56 Megarity, C. F., Huang, M., Warnock, C. and Timson, D. J. (2011) The role of the active site residues in human galactokinase: Implications for the mechanisms of GHMP kinases. *Bioorg. Chem.* **39**, 120-126.
- 57 Reinhardt, L. A., Thoden, J. B., Peters, G. S., Holden, H. M. and Cleland, W. W. (2013) pH-rate profiles support a general base mechanism for galactokinase (*lactococcus lactis*). *FEBS Lett.* **587**, 2876-2881.
- 58 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.
- 59 Kristiansson, H. and Timson, D. J. (2012) Galactokinases: Potential biotechnological applications as biocatalysts. *Current Biotechnology.* **1**, 148-154.
- 60 Yang, J., Hoffmeister, D., Liu, L., Fu, X. and Thorson, J. S. (2004) Natural product glycorandomization. *Bioorg. Med. Chem.* **12**, 1577-1584
- 61 Langenhan, J. M., Peters, N. R., Guzei, I. A., Hoffmann, F. M. and Thorson, J. S. (2005) Enhancing the anticancer properties of cardiac glycosides by neoglycorandomization. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12305-12310.
- 62 Thorson, J. S., Barton, W. A., Hoffmeister, D., Albermann, C. and Nikolov, D. B. (2004) Structure-based enzyme engineering and its impact on *in vitro* glycorandomization. *Chembiochem.* **5**, 16-25.
- 63 Yang, J., Fu, X., Jia, Q., Shen, J., Biggins, J. B., Jiang, J., Zhao, J., Schmidt, J. J., Wang, P. G. and Thorson, J. S. (2003) Studies on the substrate specificity of *escherichia coli* galactokinase. *Org. Lett.* **5**, 2223-2226.
- 64 Hoffmeister, D., Yang, J., Liu, L. and Thorson, J. S. (2003) Creation of the first anomeric D/L-sugar kinase by means of directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13184-13189.

- 65 Hoffmeister, D. and Thorson, J. S. (2004) Mechanistic implications of *escherichia coli* galactokinase structure-based engineering. *Chembiochem.* **5**, 989-992. doi:10.1002/cbic.200400003 [doi]
- 66 Yang, J., Fu, X., Liao, J., Liu, L. and Thorson, J. S. (2005) Structure-based engineering of *E. coli* galactokinase as a first step toward in vivo glycorandomization. *Chem. Biol.* **12**, 657-664.
- 67 Kristiansson, H. and Timson, D. J. (2011) Increased promiscuity of human galactokinase following alteration of a single amino acid residue distant from the active site. *ChemBioChem.* **12**, 2081-2087.
- 68 Kristiansson, H. and Timson, D. J. (2012) N-acetylgalactosamine kinase: A naturally promiscuous small molecule kinase. *Appl. Biochem. Biotechnol.* **166**, 57-63.
- 69 Stockbridge, R. B. and Wolfenden, R. (2009) The intrinsic reactivity of ATP and the catalytic proficiencies of kinases acting on glucose, N-acetylgalactosamine, and homoserine: A thermodynamic analysis. *J. Biol. Chem.* **284**, 22747-22757.
- 70 Goodey, N. M. and Benkovic, S. J. (2008) Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **4**, 474-482.
- 71 Lehmann, M., Pasamontes, L., Lassen, S. F. and Wyss, M. (2000) The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta.* **1543**, 408-415.
- 72 Risso, V. A., Manssour-Triedo, F., Delgado-Delgado, A., Arco, R., Barroso-delJesus, A., Ingles-Prieto, A., Godoy-Ruiz, R., Gavira, J. A., Gaucher, E. A., Ibarra-Molero, B. and Sanchez-Ruiz, J. M. (2015) Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol. Biol. Evol.* **32**, 440-455.
- 73 Mesa-Torres, N., Yunta, C., Fabelo-Rosa, I., Gonzalez-Rubio, J. M., Sanchez-Ruiz, J. M., Salido, E., Albert, A. and Pey, A. L. (2014) The consensus-based approach for gene/enzyme replacement therapies and crystallization strategies: The case of human alanine-glyoxylate aminotransferase. *Biochem. J.* **462**, 453-463.
- 74 Thoden, J. B., Timson, D. J., Reece, R. J. and Holden, H. M. (2004) Molecular structure of human galactose mutarotase. *J. Biol. Chem.* **279**, 23431-23437.
- 75 Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L. and Holden, H. M. (2000) Crystallographic evidence for tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. *Biochemistry.* **39**, 5691-5701.
- 76 Wedekind, J. E., Frey, P. A. and Rayment, I. (1996) The structure of nucleotidylated histidine-166 of galactose-1-phosphate uridylyltransferase provides insight into phosphoryl group transfer. *Biochemistry.* **35**, 11560-11569.

Figure legends

Figure 1: The Leloir pathway. Representative structures of the enzymes are shown next to the corresponding reactions. The human structures of galactose mutarotase (GALM, 1SO0 [74]), galactokinase (GALK1, 1WUU [42]) and UDP-galactose 4'-epimerase (GALE, 1EK6 [75]) are shown. Since there is currently no structure of the human galactose 1-phosphate uridylyltransferase (GALT) known, the *E. coli* structure (1HXQ [76]) is shown for this enzyme.

Figure 2: Dysfunction of each of the enzymes of the Leloir pathway can result in different types of galactosemia. Of these, types I and III are more severe than type II. Inhibition of GALK1 would reduce the build-up of galactose 1-phosphate in galactosemia types I and III. Thus it may be possible to pharmacologically convert these severe diseases into something resembling a much milder one [27].

Figure 3: Active site residues in human and bacterial galactokinases. Residues in close proximity to the substrates in (a) *L. lactis* (PDB: 1PIE; [44]) and (b) human galactokinase (PDB: 1WUU; [42]) are shown. Since no crystal structure with ATP bound is available for the *L. lactis* enzyme, this is shown with inorganic phosphate (P_i).