CMLS, Cell. Mol. Life Sci. 61 (2004) 2471–2484
1420-682X/04/202471-14
DOI 10.1007/s00018-004-4160-6

CMLS Cellular and Molecular Life Sciences

Review

Galactokinase: structure, function and role in type II galactosemia

H. M. Holden[a,*], J. B. Thoden[a], D. J. Timson[b] and R. J. Reece[c,*]

[a] Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706 (USA), Fax: +1 608 262 1319, e-mail: hazel_holden@biochem.wisc.edu
[b] School of Biology and Biochemistry, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, (United Kingdom)
[c] School of Biological Sciences, The University of Manchester, The Michael Smith Building, Oxford Road, Manchester M13 9PT, (United Kingdom), Fax: +44 161 275 5317, e-mail: richard.reece@man.ac.uk

Received 13 April 2004; accepted 7 June 2004

Abstract. The conversion of beta-D-galactose to glucose 1-phosphate is accomplished by the action of four enzymes that constitute the Leloir pathway. Galactokinase catalyzes the second step in this pathway, namely the conversion of alpha-D-galactose to galactose 1-phosphate. The enzyme has attracted significant research attention because of its important metabolic role, the fact that defects in the human enzyme can result in the diseased state referred to as galactosemia, and most recently for its utilization via 'directed evolution' to create new natural and unnatural sugar 1-phosphates. Additionally, galactokinase-like molecules have been shown to act as sensors for the intracellular concentration of galactose and, under suitable conditions, to function as transcriptional regulators. This review focuses on the recent X-ray crystallographic analyses of galactokinase and places the molecular architecture of this protein in context with the extensive biochemical data that have accumulated over the last 40 years regarding this fascinating small molecule kinase.

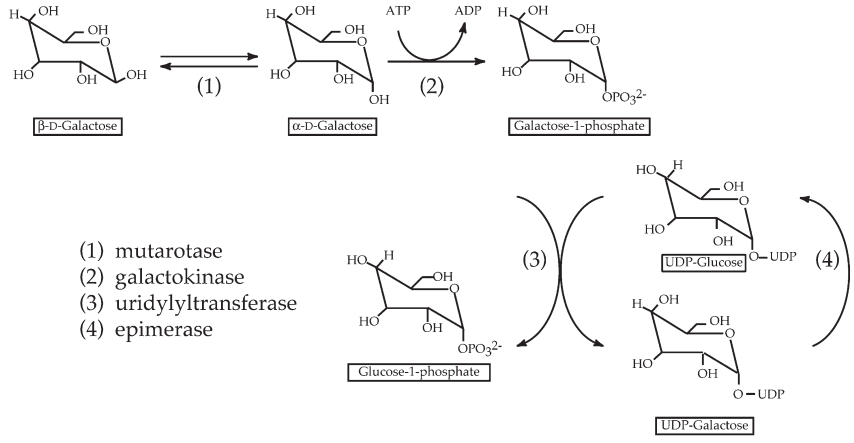
Key words. Galactokinase; GHMP superfamily; Leloir pathway; galactose regulation; galactosemia.

Introduction

In most organisms, the conversion of β-D-galactose to the more metabolically useful glucose 1-phosphate is accomplished by the action of four enzymes that constitute the Leloir pathway [1, 2]. Galactokinase, the topic of this review, catalyzes the second step in this pathway, namely the conversion of α-D-galactose to galactose 1-phosphate as indicated in scheme 1. The enzyme has attracted significant research attention for well over 45 years in part because of its important metabolic role, the fact that defects in the human enzyme can result in the diseased state referred to as galactosemia [3], and most recently for its utilization via 'directed evolution' to create new natural and unnatural sugar 1-phosphates [4]. In addition, and perhaps most surprisingly, galactokinase-like molecules can act as sensors for the intracellular concentration of galactose and, under suitable conditions, function as transcriptional regulators [5].

Early reports of galactokinase activity began to appear in the literature in the early 1960s with enzymes being isolated from sources such as *Escherichia coli* [6, 7], yeast [8] and pig liver [9]. At present, 104 different amino acid sequences for galactokinases from various organisms have been deposited in the SwissProt data bank. Additionally, the X-ray structures of the galactokinases from *Lactococcus lactis* [10] and *Pyrococcus furiosus* have been determined within the last year [11]. This review fo-

* Corresponding authors.



Scheme 1. The Leloir pathway. The enzymatic steps involved in the conversion of β-D-galactose 1-phosphate.

cuses on the recent X-ray crystallographic analyses of galactokinase and places the molecular architecture of this protein in context with the extensive biochemical data that have accumulated over the last 40 years regarding this fascinating small molecule kinase.

The molecular architecture of galactokinase

The first three-dimensional structure of galactokinase to be reported was that from *L. lactis* [10]. The bacterial enzyme is monomeric and consists of 399 amino acid residues. Its amino acid sequence is 34% identical and 47% similar to that of human galactokinase. A ribbon representation of the *L. lactis* enzyme is displayed in figure 1a. As can be seen, the molecule is elongated with overall dimensions of ~70 Å × 54 Å × 57 Å and folds into two regions referred to as the N- and C-terminal domains. These domains are formed between Thr 9 to Gly 190 and Glu 191 to Gly 390, respectively. The last amino acid residues at the C-terminus curl back onto the N-terminal domain. Five strands of mixed β-sheet and five α-helices characterize the N-terminal domain, whereas two layers of antiparallel β-sheet and six α-helices dominate the C-terminal motif. The enzyme from *L. lactis* was crystallized in the presence of inorganic phosphate and D-galactose. These ligands bind at the interface between the N- and C-terminal domains as indicated by the ball-and-stick representations in figure 1a. A close-up view of the region surrounding the α-D-galactose moiety is presented in figure 1b. Amino acid side chains responsible for anchoring the sugar ligand to the protein include Arg 36, Glu 42, Asp 45, Asp 183 and Tyr 233. Both Arg 36 and Asp 183 are strictly conserved in the amino acid sequences available in the literature thus far for galactokinases. The carboxylate side chain of Asp 183 is positioned within 3.5 Å of the C-1 hydroxyl group of galactose, whereas the guanidinium group of Arg 36 is situated between both the C-1 hydroxyl group and the inorganic phosphate.

Following the structural analysis of the *L. lactis* galactokinase, the molecular architecture of the enzyme from *P. furiosus* was reported [11]. This was a particularly important investigation in that the structure was determined in the presence of both D-galactose and MgADP. The nucleotide binds at the interface between the N- and C-terminal domains with the adenine ring located in a hydrophobic pocket formed by Phe 52, Trp 69, Ile 94, Leu 100, and Phe 110 and the 2- and 3-hydroxyls of the ribosyl group exposed to the solvent. The β-phosphate sits within hydrogen bonding distance to the backbone peptidic groups of Gly 101 and Ser 107 and the side chain hydroxyl groups of Ser 106 and Ser 107. The magnesium ion is coordinated by the side chain oxygens of Ser 107 and Glu 130 and by α- and β-phosphoryl oxygens of the nucleotide.

Not surprisingly, given the amino acid sequence similarity of 52% between the *P. furiosus* and *L. lactis* galactokinases, the polypeptide chains for these two enzymes superimpose with a root-mean-square deviation of 1.3 Å for 339 structurally equivalent α-carbons. Presented in figure 1c is a model of the *L. lactis* enzyme with bound α-D-galactose and ATP. This model was constructed on the basis of the observed binding modes for galactose and inorganic phosphate in the *L. lactis* enzyme and the binding orientation of ADP in the *P. furiosus* structure. As can be seen, Asp 183 is in the proper location to abstract the hydrogen from the C-1 hydroxyl group of galactose. Additionally, the C-1 hydroxyl group lies closely and in line with the γ-phosphate of ATP. The positively charged guanidinium group of Arg 36 might serve to both lower the pK_a of the C-1 hydroxyl group and to stabilize the penta-coordinated γ-phosphate transition state. Site-directed mutagenesis experiments to test this hypothesis are currently under way in this laboratory.

A space-filling representation of the *L. lactis* enzyme with bound α-D-galactose and ATP is shown in figure 1d. Note that the triphosphate moiety of ATP is completely buried, whereas portions of the ribosyl group and the ade-

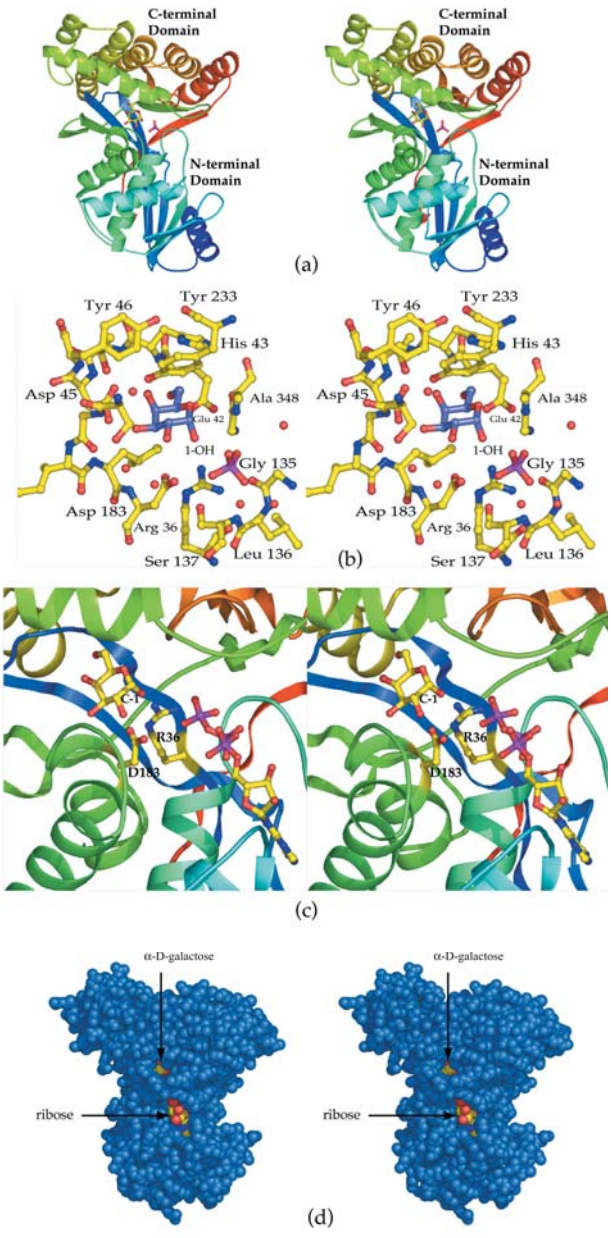


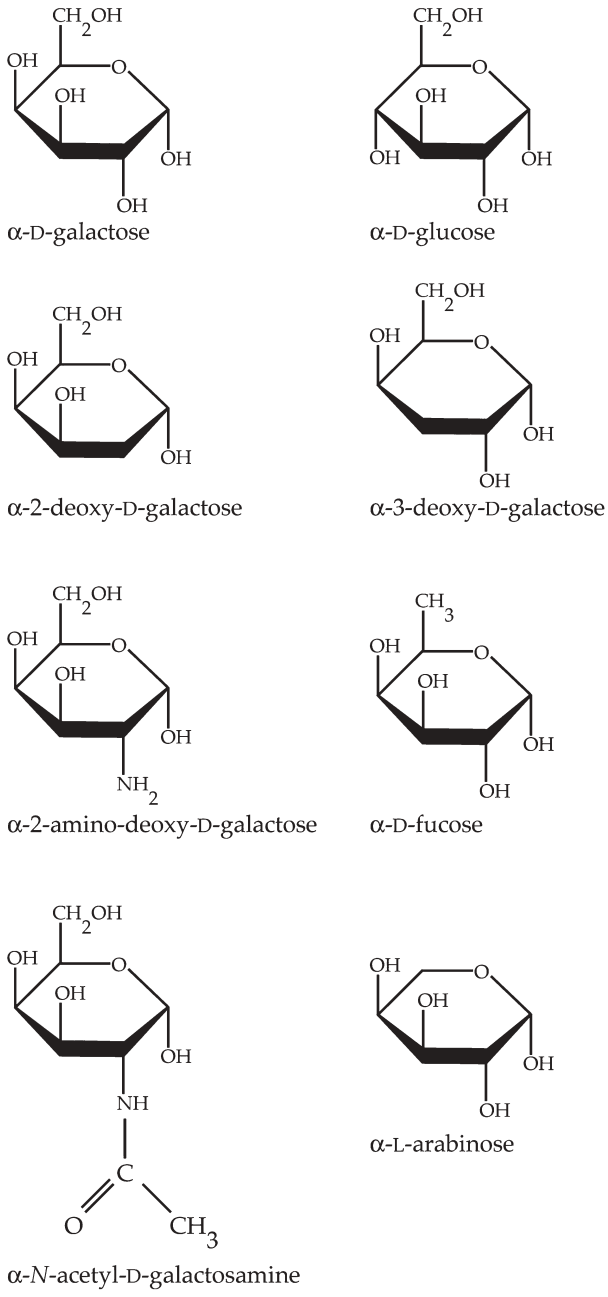
Figure 1. The structure of galactokinase from *L. lactis*. A ribbon representation of galactokinase is shown in (*a*) with the bound α-D-galactose and inorganic phosphate ligands depicted in ball-and-stick representations. A close-up view of the active site within ~3.2 Å of the sugar is presented in (*b*). The sugar moiety is highlighted in blue-gray bonds. A model of the enzyme with bound α-D-galactose and ATP is displayed in (*c*), while the corresponding space-filling representation is depicted in (*d*).

nine base are exposed to the solvent. The solvent accessible surface area for the nucleotide is ~270 Å^2 whereas that for galactose is ~100 Å^2. In the case of the *L. lactis* enzyme, the structure was solved in the presence of D-galactose and inorganic phosphate, whereas with the *P. furiosus* galactokinase, crystals were obtained in the presence of MgADP and D-galactose. If there are major conformational changes that occur upon substrate binding, at least in the case of these two bacterial galactokinases, it most likely is triggered by sugar rather than by nucleotide binding. A structural analysis of the unliganded form of the enzyme would help to clarify this issue of conformational change.

Kinetics and substrate specificity of galactokinase

One of the first galactokinases to be studied with regard to kinetics was that isolated from rat liver [12, 13]. The catalytic mechanism was shown to be sequential with ATP binding first. Following that investigation, a thorough examination of the kinetics of highly purified galactokinase from *E. coli* was reported [14]. It was shown that the bacterial enzyme displays a sequential mechanism in which both substrates add to the protein before either product is released and that the reaction mechanism is apparently random, with either ATP or galactose binding first. Contrastingly, in a subsequent study, galactokinase isolated from dormant *Vicia faba* seeds was demonstrated to follow an ordered reaction mechanism whereby galactose binds first followed by ATP [15]. In recent years, the enzyme from *Saccharomyces cerevisiae* (Gal1p) has been targeted for investigation, in part because it and the closely related protein Gal3p act as ligand sensors in regulating the transcription of genes required for galactose metabolism (reviewed in [5]). In the case of the yeast enzyme, the reaction mechanism is ordered but with ATP rather than galactose binding in the first step [16]. Similar results have been observed with the human enzyme (GALK1) [17].

With respect to sugar specificity, in an initial study on rat liver galactokinase, it was demonstrated that the enzyme phosphorylates 2-deoxy-D-galactose as well as its natural substrate [13]. A subsequent study revealed that the enzyme from *S. cerevisiae* shows a high degree of specificity for D-galactose with no ability to catalyze the phosphorylation of glucose, mannose, galactitol, arabinose, fucose, lactose or even 2-deoxy-D-galactose [18]. Interestingly, the galactokinase isolated from fenugreek seeds can use 2-deoxy-D-galactose and fucose as substrates [19], while the enzyme obtained from *Vicia faba* seeds cannot phosphorylate fucose [15]. Recent years have witnessed a renewed interest in the substrate specificities of galactokinases from various sources, due in part to their potential application for providing enzymatic routes in the preparation of unique sugar phosphates. As an example, the substrate specificity of the enzyme from *E. coli* was recently examined with 12 different sugars [20]. Of these sugars, only four proved to be substrates for the enzyme: 2-deoxy-D-galactose, 2-amino-deoxy-D-galactose, 3-deoxy-D-galactose and D-fucose, as shown in scheme 2. Clearly, the *E. coli* enzyme can accommodate substitutions at position C-2 up to a certain bulkiness and can act upon sugars



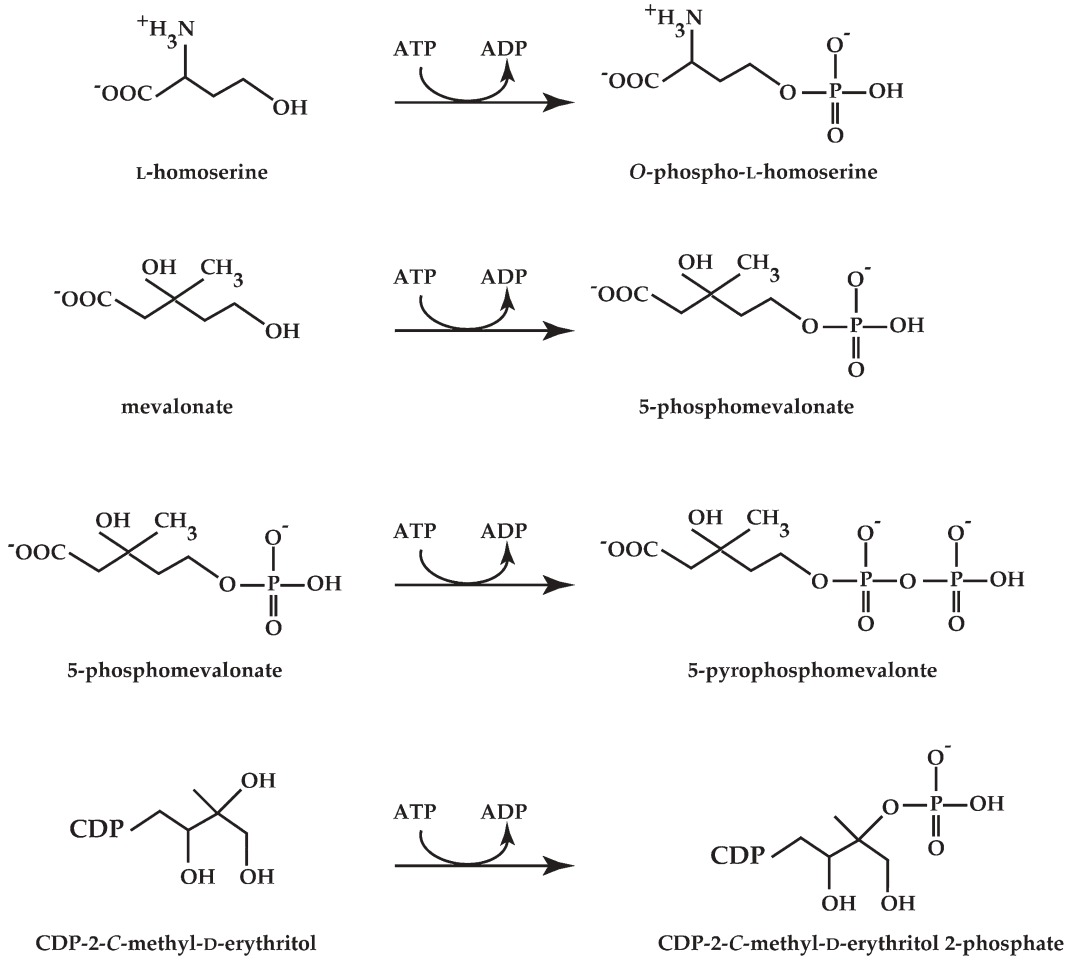
Scheme 2. The structure of galactose and other sugars that have been used to define substrate specificity amongst galactokinases.

lacking hydroxyl groups at positions C-3 and C-6. Strikingly, however, the enzyme cannot tolerate any modifications at C-4. Substrate specificity studies with the human enzyme have demonstrated that this galactokinase recognizes D-galactose and 2-deoxy-D-galactose but not *N*-acetyl-D-galactosamine, L-arabinose, D-fucose or D-glucose (scheme 2) [21]. Taken together, these investigations highlight the diversity of the galactokinases with respect to both their kinetics and substrate specificities and emphasize the need to consider various sources of the enzyme for future research aimed at directed evolution.

The GHMP superfamily of proteins

More than 10 years ago it was recognized, on the basis of amino acid sequence alignments, that galactokinase belongs to a larger family of proteins which includes homoserine kinase, mevalonate kinase and phosphomevalonate kinase, among others [22]. This family soon became known as the GHMP superfamily (with the abbreviation referring to the original members of the group). Proteins belonging to the family are now known to be structurally unique from the well-characterized P-loop containing kinases [23], the actin-like ATPases [24] and the ATP-grasp enzymes [25, 26]. All members of the GHMP group contain three common structural/functional motifs (I, II and III), the second of which is the most conserved with a typical sequence Pro-X-X-X-Gly-Leu-X-Ser-Ser-Ala [22]. This region is involved in nucleotide binding.

Nearly 7 years following the identification of the GHMP superfamily, the structure of one of its members, homoserine kinase from *Methanococcus jannaschii*, was solved [27, 28]. This enzyme plays a key role in bacterial threonine biosynthesis by catalyzing the phosphorylation of L-homoserine to *O*-phospho-L-homoserine, as indicated in scheme 3. A superposition of the α-carbons for homoserine kinase and galactokinase from *L. lactis* is shown in figure 2a. These enzymes superimpose with a root mean square deviation of 2.4 Å for 259 structurally equivalent α-carbon positions. In sharp contrast to that observed for the *P. furiosus* galactokinase, the bound nucleotide in homoserine kinase adopts a rare *syn* conformation [27, 28]. Interestingly, X-ray investigations of ternary complexes of homoserine kinase with homoserine and non-hydrolysable ATP analogs have revealed the lack of a catalytic base positioned near the δ-hydroxyl group of homoserine. Indeed, Asp 183 in the *L. lactis* galactokinase is structurally equivalent to Asn 141 in homoserine kinase. The apparent absence of an active site base has led to the proposal that the catalytic mechanism of homoserine kinase proceeds through stabilization of the transition state [28]. Specifically it has been postulated that the close positioning of the substrate and the γ-phosphate of the nucleotide within the active site results in direct transfer of the proton from the δ-OH group of homoserine to a γ-phosphoryl oxygen of ATP and attack of the δ-oxygen on the γ-phosphorus as indicated in scheme 4. The side chain hydroxyl group of Thr 183 has been shown to interact with the β-phosphate of the nucleotide and thus may help to stabilize the transition state. This type of direct proton transfer from substrate to nucleotide is not unprecedented and has been proposed, for example, in the hydrolysis reactions catalyzed by transducin α [29], p21ras [30] and myosin subfragment-1 [31]. The structure of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase also suggests a catalytic mecha-



Scheme 3. The reactions of GHMP kinase superfamily members.

nism involving stabilization of the transition state alone [32].

Following the successful X-ray analysis of homoserine kinase, the structures of mevalonate kinase from *M. jannaschii* and rat were reported [33, 34]. As indicated in scheme 3, mevalonate kinase catalyzes the phosphorylation of the C-5 hydroxyl oxygen of mevalonate to yield mevalonate 5-phosphate, a key intermediate in the biosynthesis of isoprenoids and sterols. The structure of the enzyme from rat was solved in the presence of ATP, and in this particular example the adenine ring of the nucleotide is in the *anti* conformation. As observed in galactokinase, there is a conserved aspartate residue (Asp 204) that may serve to abstract the proton from the C-5 hydroxyl group of mevalonate for its subsequent attack on the γ-phosphorus of ATP [35]. Additionally, there is a lysine residue, Lys 13, which is hypothesized to lower the pK_a of the substrate C-5 hydroxyl group. This residue corresponds to Arg 36 in the *L. lactis* galactokinase. The proposed catalytic mechanism for mevalonate kinase is shown in scheme 4.

As is so often the case, following the initial reports on the structures of homoserine kinase and mevalonate kinase, a flurry of papers subsequently appeared in the literature regarding the molecular architectures of other members of the GHMP superfamily. Within the last 2 years, the structures of phosphomevalonate kinase from *Streptococcus pneumoniae*, CDP-2-*C*-methyl-D-erythritol kinase from *Thermus thermophilus* and XOL-1 from *Caenorhabditis elegans* have been described [36–38]. Phosphomevalonate kinase functions in isoprenoid biosynthesis by promoting the transfer of the γ-phosphoryl group of ATP to the phosphate oxygen of 5-phosphomevalonate to yield 5-pyrophosphomevalonate (scheme 3). Likewise, CDP-2-*C*-methyl-D-erythritol kinase is involved in isoprenoid biosynthesis in several pathogenic microorganisms where it catalyzes the phosphorylation of the 2-hydroxyl group of CDP-2-*C*-methyl-D-erythritol (scheme 3). Both of these enzymes contain the conserved Lys/Arg and Asp residues as found in mevalonate kinase and galactokinase and, thus, it can be speculated that their catalytic mechanisms operate in a similar manner.

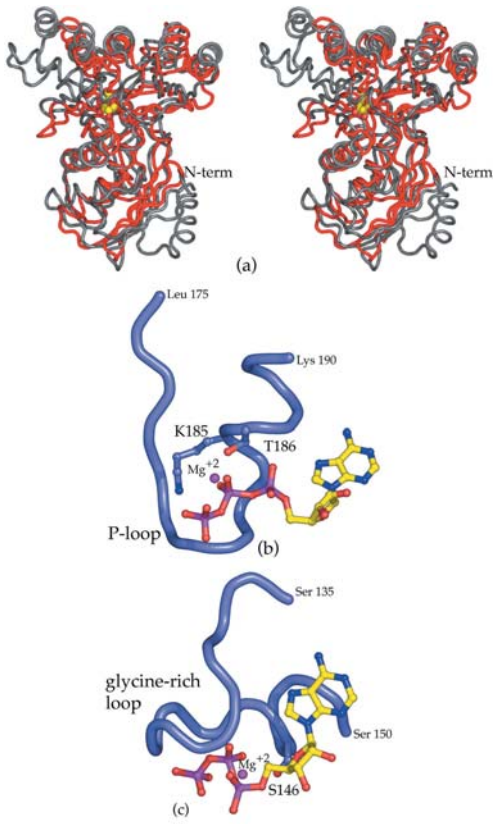
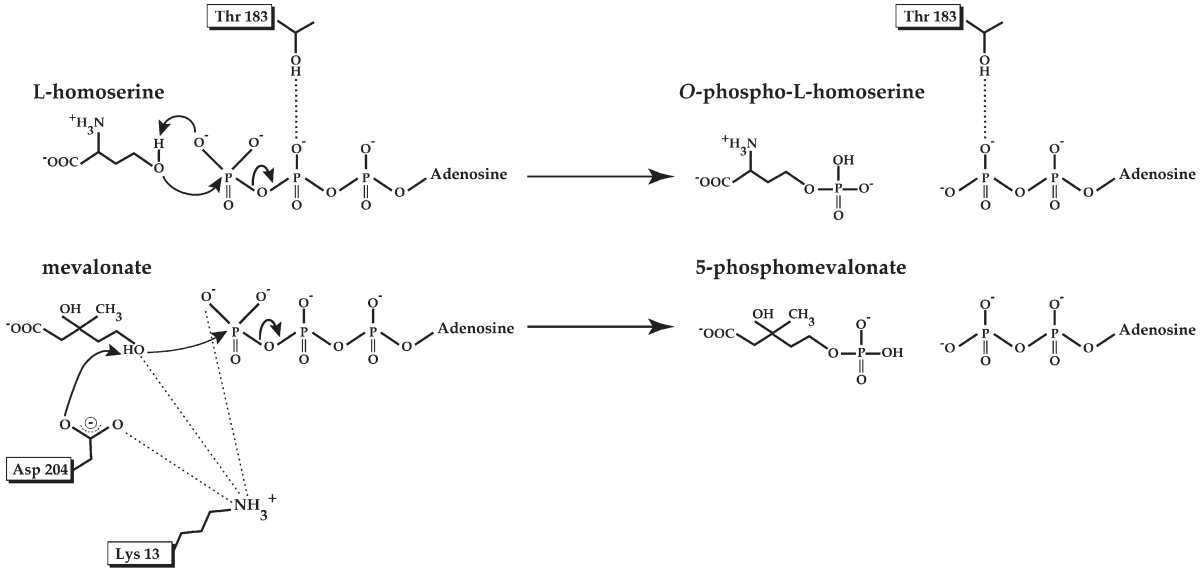


Figure 2. The GHMP superfamily. Galactokinase belongs to a family of proteins referred to as the GHMP superfamily. Shown in (*a*) is a superposition of the α-carbon traces for galactokinase from *L. lactis* (in black) and homoserine kinase from *M. jannaschii* (in red). Members of this superfamily accommodate ATP in a completely different manner from that observed for the P-loop enzymes. The P-loop found in the motor domain of *D. discoideum* myosin II is presented in (*b*), while the glycine-rich loop found in rat mevalonate kinase is displayed in (*c*). X-ray coordinates for homoserine kinase, myosin II and mevalonate kinase were obtained from the Protein Data Bank (1FWK, 1KVK and 1FMW, respectively).

Especially intriguing is the report on the X-ray crystallographic analysis of XOL-1 from *C. elegans* [38]. While the biological role of this protein is still largely uncharacterized, it is believed to function as a developmental regulator in the determination of sexual fate in the organism. Despite remarkably low amino acid sequence identities of between 9 and 10%, the fold of XOL-1 is similar to members of the GHMP superfamily. In XOL-1, however, residues corresponding to Arg 36 and Asp 183 in the *L. lactis* galactokinase are replaced with Thr 53 and Val 178, respectively. XOL-1 apparently does not bind ATP, and it has been suggested that the molecular scaffold of the GHMP superfamily has evolved to execute a completely new function [38].

A similar story has been noted within the short chain dehydrogenase/reductase (SDR) superfamily. Members belonging to this family typically catalyze oxidation-reduction reactions using $NAD(P)^+$ as the cofactor, employ a tyrosine residue as the catalytic base and function as dehydrogenases, dehydratases, isomerases or epimerases [39, 40]. One of the first members of the SDR superfamily to be extensively studied by high-resolution X-ray crystallographic analyses was UDP-galactose 4-epimerase from *E. coli*, which catalyzes the final step in the Leloir pathway (scheme 1) [41]. Each subunit of the dimeric enzyme folds into two domains: an N-terminal motif that is involved in dinucleotide binding and a C-terminal region that is responsible for substrate positioning. While it was assumed that the molecular scaffold employed by epimerase and other members of the SDR superfamily was ideally suited for catalysis only, subsequent amino acid sequence alignments suggested that this may not be strictly true [42]. The high-resolution structural investigation of NmrA, a negative transcriptional regulator in *Aspergillus nidulans* with no known catalytic



Scheme 4. Proposed catalytic mechanisms for homoserine kinase (upper) and mevalonate kinase (lower).

activity, confirmed that the molecular architecture employed by members of the SDR superfamily is also suited for regulatory functions [43]. While the catalytic tyrosine is not conserved in NmrA, nevertheless the protein binds NAD^+, at least in the crystalline state. Another protein of interest is human TIP30/CC3, a putative metastasis suppressor that promotes apoptosis and inhibits angiogenesis [44–48]. TIP30 shows significant amino acid sequence similarity to *E. coli* UDP-galactose 4-epimerase, and modeling studies have suggested that the unique 20 amino acids at the N-terminus of the protein may be important for its interaction with cellular targets [49]. It is interesting that the scaffolds for two of the enzymes in the Leloir pathway, galactokinase and UDP-galactose 4-epimerase, provide three-dimensional platforms that are suitable for regulatory functions as well. Furthermore, galactose mutarotase, the first enzyme of the Leloir pathway (scheme 1), has a very similar overall molecular architecture to domain 5 of β-galactosidase [50, 51]. However, the role of domain 5 in β-galactosidase function is unknown at the present time.

The manner in which members of the GHMP superfamily accommodate ATP is distinctively different from that observed in the well-characterized P-loop proteins [23]. The P-loop enzymes contain a characteristic signature sequence of Gly-X-X-X-X-Gly-Lys-Thr/Ser. Shown in figure 2b is the P-loop from the motor domain of *Dictyostelium discoideum* myosin II with bound MgATP [52]. For this motor domain, the P-loop begins at Gly 179 and has an amino acid sequence of Gly-Glu-Ser-Gly-Ala-Gly-Lys-Thr. Backbone peptidic groups provide most of the electrostatic interactions between the P-loop and the β- and γ-phosphoryl groups of ATP. The side chain functional groups of Lys 185 and Thr 186 provide additional electrostatic or hydrogen bonding interactions. Note that the magnesium ion is coordinated by the β- and γ-phosphoryl groups of the nucleotide. Members of the GHMP superfamily contain a glycine-rich motif rather than the P-loop. In rat mevalonate kinase, shown in figure 2c, this region begins at Gly 140 with the amino acid sequence Gly-Ala-Gly-Leu-Gly-Ser-Ser-Ala-Ala. Again, most of the interactions between the β- and γ-phosphates of the nucleotide and the protein are provided by backbone peptidic groups in the glycine-rich loop. The magnesium ion, like that observed for the myosin motor domain, is coordinated by the β- and γ-phosphoryl oxygens. As can be seen by comparing figures 2b and 2c, the conformations of the nucleotide phosphate backbones are quite different. Note that in both the P-loop kinases and the GHMP enzymes, the ribosyl hydroxyl groups are solvent exposed. This is in sharp contrast to that observed for the ATP-grasp proteins where the 2- and 3-hydroxyl groups are typically bridged by the carboxylate of a glutamate residue [26].

The fold of the GHMP superfamily provides an ideal milieu for the phosphorylation of small molecules. Within the superfamily, however, significant differences are now being seen. In the case of homoserine kinase, the reaction mechanism appears to proceed without the employment of a catalytic base, while in enzymes such as mevalonate kinase and phosphomevalonate kinase, a conserved base has been postulated to function in catalysis. Whether base catalysis is operative in galactokinase is still unclear, however. While Asp 183 is located in an appropriate place to function as a base, detailed kinetics studies on the yeast enzyme suggest that the reaction mechanism proceeds through transition state stabilization [16].

The regulation of galactose metabolism in yeast

The expression of the *GAL* genes of the baker's yeast *S. cerevisiae* has, for many years, been studied as a paradigm for eukaryotic transcriptional regulation. The *GAL* genetic switch provides the transcriptional control of genes whose products are required for the metabolism of galactose (including the enzymes of the Leloir pathway) [5]. For yeast cells grown without galactose, the *GAL* genes are essentially inert. They are not expressed because a transcriptional activator (Gal4p) upon which they depend is rendered inactive through its interaction with a transcriptional repressor (Gal80p) [53]. The complex of the activator and repressor remain bound to DNA upstream of the *GAL* genes, but the activator cannot recruit RNA polymerase II to initiate transcription. This means that the yeast cells will metabolize other sources of carbon, e.g., glucose, in preference to galactose even if a mixture of glucose and galactose is available to the cell [54]. If, however, the only source of carbon available to the cell is galactose, then the *GAL* genes will become transcriptionally active, both rapidly and to a high level, and galactose is converted to glucose 1-phosphate (see scheme 1) [55]. The presence of galactose within the cell is not detected by either Gal4p or Gal80p, but by another protein. This transcriptional inducer, Gal3p, interacts with the transcriptional repressor, Gal80p, in a galactose- and ATP-dependent manner [56]. The net result of this interaction is that Gal4p becomes active and transcription of the *GAL* genes proceeds. Gal3p requires both galactose and ATP to interact with Gal80p [57], but unlike a galactokinase, which also uses galactose and ATP as substrates, it does not possess an enzymatic activity [58]. Rather, the protein appears to require galactose and ATP so that it can adopt a conformation to allow it to interact with Gal80p [56]. The genome of *S. cerevisiae* contains two galactokinase-like genes. The first (encoding Gal1p) is a well-characterized galactokinase enzyme [16, 59], while the other gene encodes Gal3p. Gal1p and Gal3p are highly similar, with the encoded proteins being ~70% identical and ~90% similar. Other galactokinases, e.g., those from *E. coli* [60], human [D. J. Timson and R. J.

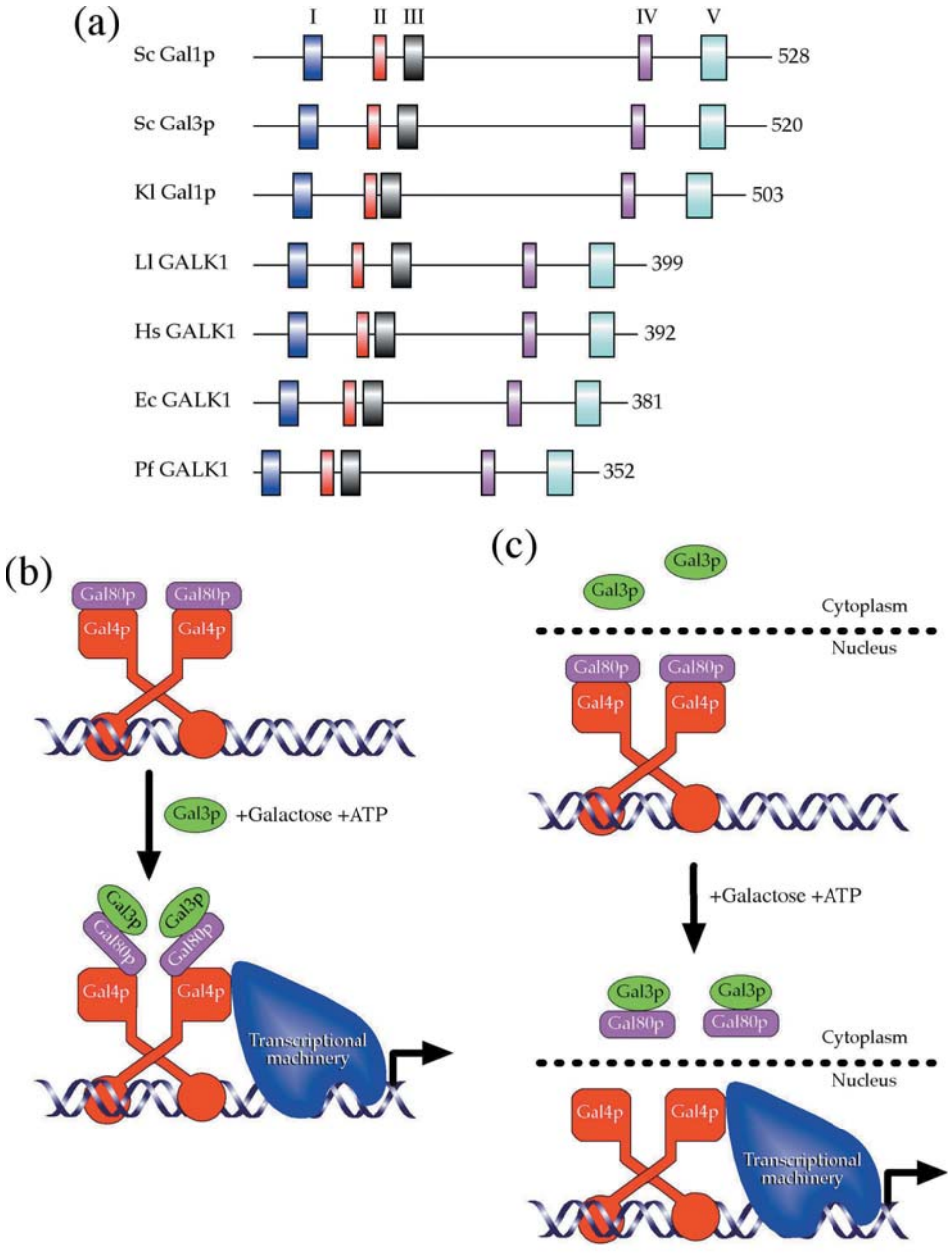


Figure 3. Galatokinase enzymes and transcriptional regulation. The locations of the amino acid sequence elements common to galactokinase-like molecules are shown in (*a*). The positions of each of five amino acid sequence motifs (I–V) originally identified by [58] are indicated. Motifs I and V (which correspond to the structural motifs I and III in the GHMP superfamily) form part of the galactose binding site, while motif III (defined as motif II in the GHMP superfamily) forms part of the ATP binding site. Sc, *S. cerevisiae*; Kl, *K. lactis*, Ll; *L. lactis*; Hs, *Homo sapiens*; Ec, *E. coli*; Pf, *P. furiosus*. Two potential mechanisms for the activation of the *GAL* genetic switch are shown in (*b*) and (*c*). In (*b*) the association of Gal3p with Gal80p results in an alteration of the interaction between Gal80p and Gal4p. Gal4p is then able to recruit the transcriptional machinery (by a direct contact with Tra1p [68]) such that gene expression occurs. In (*c*), Gal80p shuttles between the nucleus, where it can interact with DNA-bound Gal4p, and the cytoplasm. In the presence of galactose and ATP, Gal80p is retained in the cytoplasm through its association with Gal3p, thereby allowing Gal4p to activate transcription.

Reece, unpublished data], and pea [E. Taylor and R. J. Reece, unpublished data] are able to substitute for the loss of Gal1p in yeast cells, but are unable to overcome the transcriptional defect in the *GAL* genes resulting from the loss of Gal3p. In the absence of Gal3p, the induction of the *GAL* genes occurs slowly, with full induction only occurring 3–4 days after the addition of galactose to the cultures rather than within 30 min for wild-type strains [60]. The *GAL3* defect can be overcome if Gal1p is produced at higher than wild-type levels. Yeast cells bearing a *GAL3* mutation and a *GAL1* mutation are unable to induce the *GAL* genes at all. Combined, these data suggest

that while Gal1p is primarily a galactokinase enzyme, it can also function as a weak transcriptional inducer. This conclusion has also been supported through a number of in vitro experiments [53, 58]. Other yeasts, e. g., the milk yeast *Kluyveromyces lactis*, contain a single galactokinase-like molecule that is both an efficient transcriptional inducer (Gal3p-like function) and a galactokinase enzyme (Gal1p-like function) [61].

Gal3p shares all the hallmarks of a galactokinase enzyme, and the protein contains each of the five amino acid sequence homology motifs that are common to galactokinases (fig. 3 a). Note that motifs I, III and V in figure 3 a correspond to the structural motifs I, II and III identified in the GHMP superfamily. Despite the fact that Gal3p contains all of these sequence motifs, it is unable to turnover galactose. Strong evidence does, however, indicate that Gal1p and Gal3p have evolved from each other. Gal3p appears to be a defective galactokinase because it contains a two amino acid deletion in homology region III (fig. 3 a). The reintroduction of these residues (a serine and an alanine starting at position 164) imparts the resulting protein with, albeit weak, galactokinase activity [58]. The equivalent residue to the serine in the *P. furiosus* galactokinase structure interacts with the magnesium ion and the β-phosphate of the nucleotide [11].

An especially intriguing question concerns the manner in which a galactokinase evolved to function as a transcriptional regulator. An inspection of the yeast galactokinase-like proteins that can act as transcriptional inducers (Gal1p and Gal3p from *S. cerevisiae* and Gal1p from *K. lactis*) show that they have a relatively large, and relatively well conserved, region between sequence homology motifs III and IV which is not found in bacterial or mammalian galactokinases (fig. 3 a). Although it is tempting to speculate that this region of the protein is responsible for the interaction with Gal80p, and hence transcriptional induction, mutations in *GAL3* that give rise to a constitutive phenotype (that is, activation of transcription in the absence of galactose) are spread throughout its length [62]. It is, however, clear that a monomeric form of Gal3p interacts with Gal80p only in the presence of galactose and ATP [63], but the precise nature of this interaction remains obscure. Some reports have suggested that induction occurs through the association of a tripartite complex formed between Gal4p, Gal80p and Gal3p [53]. Others have proposed that Gal80p dissociates from Gal4p and interacts with Gal3p in the cytoplasm of yeast cells [64], thereby freeing Gal4p from the inhibitory effects of Gal80p and allowing transcriptional activation to occur [65]. In either model (figs 3 b, c), the ability of Gal3p to interact with Gal80p is essential for the transcriptional induction of the *GAL* genes. Biochemical and genetic evidence have been used to support both potential models for *GAL* gene activation. In favor of the nuclear association of Gal4p, Gal3p and Gal80p are the observations that Gal4p purified from yeast grown in the presence and absence of galactose is associated with Gal80p [66], artificially constructed Gal80p molecules that contain an activation domain are able to regulate transcription in the presence and absence of galactose [67], and the recent observation using Fluorescence Resonance Energy Transfer (FRET) analysis that Gal4p and Gal80p do not dissociate from each other [68]. The dissociation model (fig. 3 c) is supported by data indicating that Gal3p is predominately cytoplasmic [64] and that the expression of a myristoylated version of the protein (which will be targeted to the cell plasma membrane) does not unduly impair the induction of the *GAL* genes [65]. In addition, chromatin immunoprecipitation experiments [65] and pull-down assays [69] suggest that the Gal4p-Gal80p complex is weakened when cells are grown in the presence of galactose.

As discussed here, the mechanism of action of Gal3p in its role as a galactose sensor and transcriptional regulator is not well understood at the molecular level. The changes to the protein that are presumed to occur upon interaction with its ligands (galactose and ATP) are not defined, and the mechanism by which it controls the function of Gal80p is supported by apparently contradictory data. Clearly further experimentation is required. However, the similarity between Gal3p and galactokinases should provide one means by which further analysis can be achieved.

Directed evolution of galactokinases

Many of the clinically important medicines on today's market, such as erythromycin, vancomycin, novobiocin and digitoxin, contain various D- or L-sugar substituents. Without these appended sugar(s), the pharmacological properties of such compounds are often either completely lost or markedly decreased [70]. It has been estimated that approximately half of the drug leads around the world are derived directly from natural products, many of which are glycosylated secondary metabolites [71]. In that sugars often play major roles in the biological effectiveness of these compounds, there is an increasing effort directed towards altering glycosylation patterns to produce new and potentially important therapeutics. One approach that has been championed within the last few years is the so-called in vitro glycorandomization method whereby chemically synthesized sugars are activated and subsequently attached to natural products using promiscuous nucleotidyltransferases and glycosyltransferases [71]. The starting materials for in vitro glycorandomization are sugar phosphates. While there are synthetic routes available for the preparation of such compounds, they are often tedious and time consuming. The renewed interest in galactokinase as a biological catalyst has, in-

deed, been sparked by the need to prepare modified sugar phosphates in an expedient manner [72]. To this end, a recent directed-evolution approach using error-prone polymerase chain reation (PCR) with the *E. coli* galactokinase gene has produced a site-directed mutant protein, Y371H, which displays new kinase activity against D-galacturonic acid, D-talose, L-altrose and L-glucose. On the basis of the X-ray model for the *L. lactis* enzyme, it is believed that Tyr 371 in the *E. coli* galactokinase is situated ~20 Å from the active site. Most likely, this residue would have never been targeted for investigation simply by a 'rational' redesign of the active site region. Undoubtedly, new sugar kinases of alternate substrate specificity will be produced in the near future as such direct-evolution approaches are utilized towards the development of the next generation of drugs.

Diseases arising from galactokinase deficiencies

Approximately 20 mutations (including base substitutions, base deletions and larger deletions) have been described in human galactokinase which result in the disease type II galactosemia (MIM#230200). This disease, the main symptom of which is early onset cataracts, is less severe than either type I or type III galactosemia, which are caused by defects in the galactose-1-phosphate uridyltransferase or UDP-galactose 4-epimerase genes, respectively [3, 73–75]. Deficiency in these genes can result in kidney, liver and brain damage in addition to cataracts. Provided galactokinase deficiency is detected early in life, the condition can be managed by restricting patients to a galactose (and lactose) free diet [3].

The absence of functional galactokinase results in the buildup of unmetabolized galactose in cells and in blood plasma. This is a particular problem in the lens cells of the eye where (in humans, but not in mice [76]) the enzyme aldose reductase converts galactose to its corresponding sugar alcohol galactitol (dulcitol). Although galactose can be transported across the cell membrane, galactitol cannot, and consequently it accumulates in the lens cells. This leads to the osmotic uptake of water, swelling of the cells, and eventual apoptosis and lysis [77–79]. The process is believed to be similar to that seen in some diabetic patients where glucose accumulates in the lens and is reduced to its corresponding sugar alcohol, sorbitol, resulting in similar osmotic and lytic effects [80, 81].

The first disease-causing mutations in human galactokinase to be described were a point mutation in structural motif I of the GHMP superfamily, V32M, and a nonsense mutation at codon 80. In both cases, the patients were homozygous for the mutations, and their blood galactokinase activities were essentially zero [82]. Although the V32M mutant protein could be expressed in *E. coli*, it was not present in the soluble fraction following disruption of the bacterial cells. The most likely cause of this was a failure of the mutant protein to fold correctly [17, 83]. A study of galactokinase-deficient individuals from several Romani families in Bulgaria identified a second point mutation, P28T [84]. The mutation in this population has been dated to a single founder mutation ~750 years ago [85]. It also occurs in Bosnian populations [86] and has been detected in at least one Turkish patient [87]. Like V32M, the mutant protein could be expressed in *E. coli* but the protein could not be purified [17]. The positions of these mutations, mapped onto the *L. lactis* enzyme model, are indicated in figure 4.

Two additional point mutations have been described in or near motif I of the GHMP superfamily, G36R and H44Y [87]. The positions of these mutations are indicated in figure 4. The patient with the G36R mutation was heterozygous for this and a frameshift mutation. Blood galactokinase activity was zero, and expression of the mutant protein in *E. coli* resulted in insoluble protein [17]. H44Y was found in a patient with zero blood galactokinase activity who was heterozygous for this and another point mutation, G349S [87]. The G349S mutation (fig. 4) was also detected in a Japanese patient who was homozygous for the mutation and had a low (but non-zero) blood galactokinase activity [88]. Both these mutant proteins could be expressed in and purified from *E. coli*. Kinetic analysis of the mutant proteins showed that H44Y was defective in the turnover number (k_{cat}), the K_m for α-D-galactose, and the specificity constant (k_{cat}/K_m) for both galactose and MgATP [17]. This highly conserved residue forms part of the galactose binding site, and the main chain carbonyl group is within hydrogen bonding distance of O-6 of the sugar [10, 11]. Replacement of this residue with the bulkier tyrosine is likely to disrupt the local structure and may well result in steric hindrance in the active site, consequences of which would be entirely consistent with the kinetic changes observed

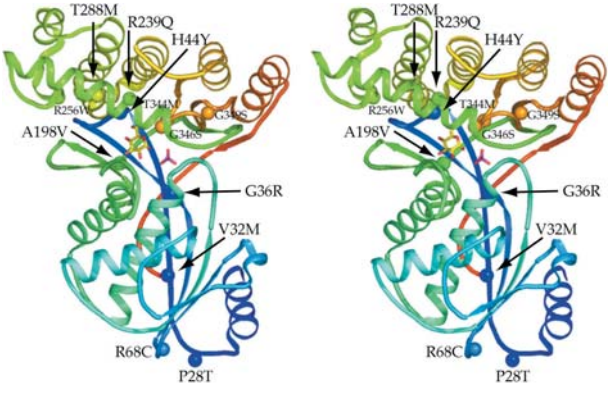


Figure 4. Mutations giving rise to type II galactosemia. The positions of some of the known disease-causing mutations in human galactokinase are mapped onto the *L. lactis* enzyme model on the basis of amino acid sequence alignments.

in this mutant protein. The importance of this histidine residue in maintaining the structure of the active site has been underlined by studies in which it was substituted with an alanine or isoleucine; both substitutions resulted in insoluble protein during expression in *E. coli* [21]. Regarding Gly 349, the kinetic consequences of this mutation are a reduction in the turnover number and in the specificity constants for both substrates [17]. It is another highly conserved residue which lies close to the active site. The ϕ,ψ angles adopted by glycine in this position are incompatible with other amino acids, and hence mutation to a serine most likely causes alterations in the backbone conformation, which affects the nearby active site [10, 11]. Another point mutation, G346S, was detected in a patient who also had a 7-bp deletion in the gene [87]. The mutant enzyme shows defects in turnover number and in the specificity constant for galactose [17]. This residue is well conserved and is usually either a glycine or an alanine. In both the *L. lactis* and *P. furiosus* structures, the equivalent residue is an alanine (Ala 348 or Ala 305, respectively). Although this residue does not directly contact the sugar, its main chain carbonyl lies within hydrogen bonding distance to the conserved histidine (His 44 in the human enzyme), which itself is close to the sugar [11].

Several point mutations outside the conserved GHMP superfamily motifs have been detected as well (fig. 4). One individual was heterozygous for two such mutations – R68C and A394P [89]. Enzymological analysis of the R68C mutant showed that it was modestly impaired in all the kinetic parameters apart from the specificity constant for α-D-galactose. The equivalent residues in the two known structures (Tyr 42 in *P. furiosus* and Arg 67 in *L. lactis*) are located ~30 Å from the active site [10, 11]. This distance suggests that any effects on the kinetic parameters are due to structural changes which are propagated through the molecule. A384P was insoluble when expressed in *E. coli* [17], and there is no clearly homologous residue in either known structure. Another mutation, T288M, which was detected in a heterozygous patient who had a frameshift in the other allele [89], gives rise to insoluble protein during expression in *E. coli* [17]. Although there are equivalent residues in the *L. lactis* and *P. furiosus* structures (Thr 290 and Val 248, respectively), the lack of conservation of residue type among other galactokinases makes it difficult to predict the structural consequences of this mutation. R256W and T344M were detected in a Japanese patient heterozygous for these mutations, and a T344M homozygote was also found [88]. Recently, an R239Q mutation was detected in a patient heterozygous for this and the known disease-causing mutation V32M [83]. Biochemical studies on the *E. coli* expressed R239Q mutant enzyme showed it to have reduced activity and thermal stability compared to wild type [83]. The mutation M1I (found in a patient homozygous for this mutation [87] and also in a heterozygote [90]) abolishes the start codon of the gene. Assuming that, in this case, translation begins from the next available methionine residue, this would mean a protein beginning at Met 55 which would completely lack structural motif I of the GHMP superfamily.

The mutation A198V (the so-called Osaka variant) was first detected in three Japanese infants with mild galactokinase deficiency [90]. Although all three patients were heterozygotes with other mutations, none developed cataracts in childhood. Population studies showed a prevalence of 4.1% in the Japanese population and lower, but detectable incidences in Korean, Taiwanese and Chinese populations. No incidence was detected in black or white populations from the United States [90]. Individuals who have this mutation have a statistically significantly higher chance of developing cataracts later in life. Comparison of the enzymes isolated from erythrocytes of individuals who were homozygous for either the A198V mutation or for the wild-type gene showed unchanged K_m values for both α-D-galactose and MgATP. However the V_{max} was reduced about fivefold [90]. Given that V_{max} is the product of the enzyme concentration and the turnover number, this reduction could be caused by an effect on either (or both) of these parameters. Immunoblot analysis showed roughly 20% of the amount of galactokinase in erythrocytes from the A198V homozygotes [90]. Furthermore, this mutant protein could be expressed in and purified from *E. coli* in yields comparable to the wild type, and its kinetic parameters were not substantially different from the wild type either [17]. The equivalent residue in the *L. lactis* enzyme, Ala 195, is ~17 Å from the active site [10]. Interestingly, the equivalent residue in *P. furiosus* is a valine [11, 91]. The biological consequences of the A198V mutation in human galactokinase are not clear, but these data suggest that this mutation does not greatly affect either its solubility or enzymology. Obviously there are subtle effects that cannot be accounted for on the basis of our current knowledge regarding galactokinase structure and function. Additionally, although some residues (Val 32, Gly 36, His 44, Gly 346 and Gly 349) are well conserved between human galactokinase and the two bacterial enzymes for which structures are known, there are many others (Pro 28, Arg 68, Ala 198, Thr 288 and Ala 384) which are less well conserved. Thus, in order to understand the manner in which these disease-causing mutations affect the structure of galactokinase, it will be necessary to determine the three-dimensional structure of the human enzyme. Such an analysis is presently under way by this group.

Acknowledgement. This research was supported in part by grants from the NIH (DK47814 to H. M. H.) and from the Wellcome Trust and the BBSRC (to R. J. R).

1 Caputto R., Leloir L. F., Trucco R. E., Cardini C. E. and Paladini A. C. (1949) Enzymatic transformations of galactose into glucose derivatives. J. Biol. Chem. **179:** 497–498
2 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
3 Segal S. and Berry G. T. (1995) Disorders of galactose metabolism. In: The Metabolic and Molecular Basis of Inherited Disease, pp. 967–1000, Scriver C. R., Beaudet A. L., Sly W. S. and Valle D. (eds), McGraw-Hill, New York
4 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. USA **100:** 13184–13189
5 Reece R. J. (2000) Molecular basis of nutrient-controlled gene expression in *Saccharomyces cerevisiae*. Cell. Mol. Life Sci. **57:** 1161–1171
6 Sherman J. R. and Adler J. (1963) Galactokinse from *Escherichia coli*. J. Biol. Chem. **238:** 873–878
7 Wilson D. B. and Hogness D. S. (1969) The enzymes of the galactose operon in *Escherichia coli*. 3. The size and composition of galactokinase. J. Biol. Chem. **244:** 2137–2142
8 Heinrich M. R. (1964) The purification and properties of yeast galactokinase. J. Biol. Chem. **239:** 50–53
9 Ballard F. J. (1966) Purification and properties of galactokinase from pig liver. Biochem. J. **98:** 347–352
10 Thoden J. B. and Holden H. M. (2003) Molecular structure of galactokinase. J. Biol. Chem. **278:** 33305–33311
11 Hartley A., Glynn S. E., Barynin V., Baker P. J., Sedelnikova S. E., Verhees C. et al. (2004) Substrate specificity and mechanism from the structure of *Pyrococcus furiosus* galactokinase. J. Mol. Biol. **337:** 387–398
12 Ballard F. J. (1966) Kinetic studies with liver galactokinase. Biochem. J. **101:** 70–75
13 Walker D. G. and Khan H. H. (1968) Some properties of galactokinase in developing rat liver. Biochem. J. **108:** 169–175
14 Gulbinsky J. S. and Cleland W. W. (1968) Kinetic studies of *Escherichia coli* galactokinase. Biochemistry **7:** 566–575
15 Dey P. M. (1983) Galactokinase of *Vicia faba* seeds. Eur. J. Biochem. **136:** 155–159
16 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
17 Timson D. J. and Reece R. J. (2003) Functional analysis of disease-causing mutations in human galactokinase. Eur. J. Biochem. **270:** 1767–1774
18 Schell M. A. and Wilson D. B. (1979) Purification of galactokinase mRNA from *Saccharomyces cerevisiae* by indirect immunoprecipitation. J. Biol. Chem. **254:** 3531–3536
19 Foglietti M. J. and Percheron F. (1976) Purification and mechanism of action of a plant galactokinase. Biochimie **58:** 499–504
20 Yang J., Fu X., Jia Q., Shen J., Biggins J. B., Jiang J. et al. (2003) Studies on the substrate specificity of *Escherichia coli* galactokinase. Org. Lett. **5:** 2223–2226
21 Timson D. J. and Reece R. J. (2003) Sugar recognition by human galactokinase. BMC Biochem. **4:** 16
22 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
23 Smith C. A. and Rayment I. (1996) Active site comparisons highlight structural similarities between myosin and other P-loop proteins. Biophys. J. **70:** 1590–1602
24 Bork P., Sander C. and Valencia A. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. Proc. Natl. Acad. Sci. USA **89:** 7290–7294
25 Galperin M. Y. and Koonin E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci. **6:** 2639–2643
26 Thoden J. B., Firestine S. M., Benkovic S. J. and Holden H. M. (2002) PurT-encoded glycinamide ribonucleotide transformylase. Accommodation of adenosine nucleotide analogs within the active site. J. Biol. Chem. **277:** 23898–23908
27 Zhou T., Daugherty M., Grishin N. V., Osterman A. L. and Zhang H. (2000) Structure and mechanism of homoserine kinase: prototype for the GHMP kinase superfamily. Structure Fold Des. **8:** 1247–1257
28 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
29 Sondek J., Lambright D. G., Noel J. P., Hamm H. E. and Sigler P. B. (1994) GTPase mechanism of G proteins from the 1.7-A crystal structure of transducin alpha-GDP-AlF_4. Nature **372:** 276–279
30 Schweins T., Geyer M., Scheffzek K., Warshel A., Kalbitzer H. R. and Wittinghofer A. (1995) Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins. Nat. Struct. Biol. **2:** 36–44
31 Fisher A. J., Smith C. A., Thoden J. B., Smith R., Sutoh K., Holden H. M. et al. (1995) X-ray structures of the myosin motor domain of *Dictyostelium discoideum* complexed with MgADP.BeFx and MgADP.AlF_4. Biochemistry **34:** 8960–8972
32 Hasemann C. A., Istvan E. S., Uyeda K. and Deisenhofer J. (1996) The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. Structure **4:** 1017–1029
33 Yang D., Shipman L. W., Roessner C. A., Scott A. I. and Sacchettini J. C. (2002) Structure of the *Methanococcus jannaschii* mevalonate kinase, a member of the GHMP kinase superfamily. J Biol. Chem. **277:** 9462–9467
34 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
35 Potter D. and Miziorko H. M. (1997) Identification of catalytic residues in human mevalonate kinase. J. Biol. Chem. **272:** 25449–25454
36 Romanowski M. J., Bonanno J. B. and Burley S. K. (2002) Crystal structure of the *Streptococcus pneumoniae* phosphomevalonate kinase, a member of the GHMP kinase superfamily. Proteins **47:** 568–571
37 Wada T., Kuzuyama T., Satoh S., Kuramitsu S., Yokoyama S., Unzai S. et al. (2003) Crystal structure of 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase, an enzyme in the non-mevalonate pathway of isoprenoid synthesis. J. Biol.Chem. **278:** 30022–30027
38 Luz J. G., Hassig C. A., Pickle C., GodzikA., 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
39 Duax W. L., Ghosh D. and Pletnev V. (2000) Steroid dehydrogenase structures, mechanism of action and disease. Vitam. Horm. **58:** 121–148
40 Oppermann U., Filling C., Hult M., Shafqat,N., Wu X., Lindh M. et al. (2003) Short-chain dehydrogenases/reductases (SDR): the 2002 update. Chem. Biol. Interact. **143–144:** 247–253
41 Thoden J. B. and Holden H. M. (1998) Dramatic differences in the binding of UDP-galactose and UDP-glucose to UDP-galactose 4-epimerase from *Escherichia coli*. Biochemistry **37:** 11469–11477
42 Baker M. E. (1999) TIP30, a cofactor for HIV-1 Tat-activated transcription, is homologous to short-chain dehydrogenases/reductases. Curr. Biol. **9:** R471
43 Stammers D. K., Ren J., Leslie K., Nichols C. E., Lamb H. K., Cocklin S. et al. (2001) The structure of the negative transcrip-

tional regulator NmrA reveals a structural superfamily which includes the short-chain dehydrogenase/reductases. EMBO J. **20:** 6619–6626
44 Xiao H., Tao Y., Greenblatt J. and Roeder R. G. (1998) A cofactor, TIP30, specifically enhances HIV-1 Tat-activated transcription. Proc. Natl. Acad. Sci. USA **95:** 2146–2151
45 Shtivelman E. (1997) A link between metastasis and resistance to apoptosis of variant small cell lung carcinoma. Oncogene **14:** 2167–2173
46 Xiao H., Palhan V., Yang Y. and Roeder R. G. (2000) TIP30 has an intrinsic kinase activity required for up-regulation of a subset of apoptotic genes. EMBO J. **19:** 956–963
47 NicAmhlaoibh R. and Shtivelman E. (2001) Metastasis suppressor CC3 inhibits angiogenic properties of tumor cells in vitro. Oncogene **20:** 270–275
48 Ito M., Jiang C., Krumm K., Zhang X., Pecha J., Zhao J. et al. (2003) TIP30 deficiency increases susceptibility to tumorigenesis. Cancer Res. **63:** 8763– 8767
49 Baker M. E., Yan L. and Pear M. R. (2000) Three-dimensional model of human TIP30, a coactivator for HIV-1 Tat-activated transcription, and CC3, a protein associated with metastasis suppression. Cell. Mol. Life Sci. **57:** 851–858
50 Thoden J. B. and Holden H. M. (2002) High resolution X-ray structure of galactose mutarotase from *Lactococcus lactis*. J. Biol. Chem. **277:** 20854–20861
51 Jacobson R. H., Zhang X. J., DuBose R. F. and Matthews B. W. (1994) Three-dimensional structure of beta-galactosidase from *E. coli*. Nature **369:** 761–766
52 Bauer C. B., Holden H. M., Thoden J. B., Smith R. and Rayment I. (2000) X-ray structures of the apo and MgATP-bound states of *Dictyostelium discoideum* myosin motor domain. J. Biol. Chem. **275:** 38494–38499
53 Platt A. and Reece R. J. (1998) The yeast galactose genetic switch is mediated by the formation of a Gal4p-Gal80p-Gal3p complex. EMBO J. **17:** 4086–4091
54 Broach J. R. (1979) Galactose regulation in *Saccharomyces cerevisiae*. The enzymes encoded by the *GAL*7, 10, 1 cluster are co-ordinately controlled and separately translated. J. Mol. Biol. **131:** 41–53
55 Yarger J. G., Halvorson H. O. and Hopper J. E. (1984) Regulation of galactokinase (*GAL*1) enzyme accumulation in *Saccharomyces cerevisiae*. Mol. Cell. Biochem. **61:** 173–182
56 Zenke F. T., Engles R., Vollenbroich V., Meyer J., Hollenberg C. P. and Breunig K. D. (1996) Activation of Gal4p by galactose-dependent interaction of galactokinase and Gal80p. Science **272:** 1662–1665
57 Yano K. and Fukasawa T. (1997) Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **94:** 1721–1726
58 Platt A., Ross H. C., Hankin S. and Reece R. J. (2000) The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase. Proc. Natl. Acad. Sci. USA **97:** 3154–3159
59 Schell M. A. and Wilson D. B. (1977) Purification and properties of galactokinase from *Saccharomyces cerevisiae*. J. Biol. Chem. **252:** 1162–1166
60 Bhat P. J., Oh D. and Hopper J. E. (1990) Analysis of the GAL3 signal transduction pathway activating *GAL*4 protein-dependent transcription in *Saccharomyces cerevisiae*. Genetics **125:** 281–291
61 Meyer J., Walker-Jonah A. and Hollenberg C. P. (1991) Galactokinase encoded by *GAL*1 is a bifunctional protein required for induction of the *GAL* genes in *Kluyveromyces lactis* and is able to suppress the gal3 phenotype in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **11:** 5454–5461
62 Blank T. E., Woods M. P., Lebo C. M., Xin P. and Hopper J. E. (1997) Novel Gal3 proteins showing altered Gal80p binding cause constitutive transcription of Gal4p-activated genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **17:** 2566–2575
63 Timson D. J., Ross H. C. and Reece R. J. (2002) Gal3p and Gal1p interact with the transcriptional repressor Gal80p to form a complex of 1:1 stoichiometry. Biochem. J. **363:** 515–520
64 Peng G. and Hopper J. E. (2000) Evidence for Gal3p's cytoplasmic location and Gal80p's dual cytoplasmic-nuclear location implicates new mechanisms for controlling Gal4p activity in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **20:** 5140–5148
65 Peng G. and Hopper J. E. (2002) Gene activation by interaction of an inhibitor with a cytoplasmic signaling protein. Proc. Natl. Acad.Sci. USA **99:** 8548–8553
66 Parthun M. R. and Jaehning J. A. (1992) A transcriptionally active form of *GAL*4 is phosphorylated and associated with *GAL*80. Mol. Cell. Biol. **12:** 4981–4987
67 Leuther K. K. and Johnston S. A. (1992) Nondissociation of *GAL*4 and *GAL*80 in vivo after galactose induction. Science **256:** 1333–1335
68 Bhaumik S. R., Raha T., Aiello D. P. and Green M. R. (2004) In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. Genes Dev. **18:** 333–343
69 Sil A. K., Alam S., Xin P., Ma L., Morgan M., Lebo C. M. et al. (1999) The Gal3p-Gal80p-Gal4p transcription switch of yeast: Gal3p destabilizes the Gal80p-Gal4p complex in response to galactose and ATP. Mol. Cell. Biol. **19:** 7828–7840
70 Thorson J. S. and Vogt T. (2002) Glycosylated natural products. In: Carbohydrate-Based Drug Discovery, Wong C.-H., (ed), Wiley-VCH, Weinheim
71 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
72 Yang J., Hoffmeister D., Liu L., Fu X. and Thorson J. S. (2004) Natural product glyorandomization. Bioorg. Med. Chem. **12:** 1577–1584
73 Petry K. G. and Reichardt J. K. (1998) The fundamental importance of human galactose metabolism: lessons from genetics and biochemistry. Trends Genet. **14:** 98–102
74 Novelli G. and Reichardt J. K. (2000) Molecular basis of disorders of human galactose metabolism: past, present, and future. Mol. Genet. Metab. **71:** 62–65
75 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
76 Ai Y., Zheng Z., O'Brien-Jenkins A., Bernard D. J., Wynshaw-Boris T., Ning C. et al. (2000) A mouse model of galactose-induced cataracts. Hum. Mol. Genet. **9:** 1821–1827
77 Kinoshita J. H., Merola L. O., Satoh K. and Dikmak E. (1962) Osmotic changes caused by the accumulation of dulcitol in the lenses of rats fed with galactose. Nature **194:** 1085–1087
78 Dvornik E., Simard-Duquesne N., Krami M., Sestanj K., Gabbay K. H., Kinoshita J. H. et al. (1973) Polyol accumulation in galactosemic and diabetic rats: control by an aldose reductase inhibitor. Science **182:** 1146–1148
79 Lin L. R., Reddy V. N., Giblin F. J., Kador P. F. and Kinoshita J. H. (1991) Polyol accumulation in cultured human lens epithelial cells. Exp. Eye Res. **52:** 93–100
80 Lee A. Y., Chung S. K. and Chung S. S. (1995) Demonstration that polyol accumulation is responsible for diabetic cataract by the use of transgenic mice expressing the aldose reductase gene in the lens. Proc. Natl.Acad. Sci. USA **92:** 2780–2784
81 Yabe-Nishimura C. (1998) Aldose reductase in glucose toxicity: a potential target for the prevention of diabetic complications. Pharmacol. Rev. **50:** 21–33
82 Stambolian D., Ai Y., Sidjanin D., Nesburn K., Sathe G., Rosenberg M. et al. (1995) Cloning of the galactokinase cDNA and identification of mutations in two families with cataracts. Nat. Genet. **10:** 307–312
83 Sangiuolo F., Magnani M., Stambolian D. and Novelli G. (2004) Biochemical characterization of two GALK1 muta-

tions in patients with galactokinase deficiency. Hum. Mutat. **23:** 396
84 Kalaydjieva L., Perez-Lezaun A., Angelicheva D., Onengut S., Dye D., Bosshard N. U. et al. (1999) A founder mutation in the GK1 gene is responsible for galactokinase deficiency in Roma (Gypsies). Am. J. Hum. Genet. **65:** 1299–1307
85 Hunter M., Heyer E., Austerlitz F., Angelicheva D., Nedkova V., Briones P. et al. (2002) The P28T mutation in the GALK1 gene accounts for galactokinase deficiency in Roma (Gypsy) patients across Europe. Pediatr. Res. **51:** 602–606
86 Reich S., Hennermann J., Vetter B., Neumann L. M., Shin Y. S., Soling A. et al. (2002) An unexpectedly high frequency of hypergalactosemia in an immigrant Bosnian population revealed by newborn screening. Pediatr. Res. **51:** 598–601
87 Kolosha V., Anoia E., de Cespedes C., Gitzelmann R., Shih L., Casco T. et al. (2000) Novel mutations in 13 probands with galactokinase deficiency. Hum. Mutat. **15:** 447–453
88 Asada M., Okano Y., Imamura T., Suyama I., Hase Y. and Isshiki G. (1999) Molecular characterization of galactokinase deficiency in Japanese patients. J. Hum. Genet. **44:** 377–382
89 Hunter M., Angelicheva D., Levy H. L., Pueschel S. M. and Kalaydjieva L. (2001) Novel mutations in the GALK1 gene in patients with galactokinase deficiency. Hum. Mutat. **17:** 77–78
90 Okano Y., Asada M., Fujimoto A., Ohtake A., Murayama K., Hsiao K. J. et al. (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
91 Verhees C. H., Koot D. G., Ettema T. J., Dijkema C., de Vos W. M. and van der Oost J. (2002) Biochemical adaptations of two sugar kinases from the hyperthermophilic archaeon *Pyrococcus furiosus*. Biochem. J. **366:** 121–127