Mechanistic Implications of *Escherichia coli* Galactokinase Structure-Based Engineering

Dirk Hoffmeister and Jon S. Thorson*[a]

Small-molecule kinases are of fundamental importance with respect to a wide range of cellular processes, and the emerging availability of small-molecule-kinase crystal structures has profoundly enhanced our mechanistic and regulatory understanding of these fascinating catalysts.[1,2] As an exemplary member of this family, the galactokinases (GalKs) have received attention for their physiological role in cellular sugar metabolism[3] and disease (galactosemia) impact within the medical sciences.[4] More recently, the enzymatic potential of GalKs has been harnessed for a biotechnological process referred to as glycorandomization. Specifically for glycorandomization, the tandem action of a promiscuous GalK and nucleotidylyltransferase is exploited to present nucleotide-sugar-donor libraries to natural-product glycosyltransferases, thereby providing a rapid chemoenzymatic means to diversify natural-product-based drug leads.[5]

Toward this goal, *E. coli* GalK directed evolution recently provided a single amino acid exchange (Y371H) far removed from the active site (~20 Å) that dramatically decreased hexopyranosyl specificity.[6] However, the hexopyranose C-4 axial requirement, that is, *galacto*-configuration, remained as a stringent substrate requirement of this Y371H mutant. Structural data from *Lactococcus lactis* GalK suggest that the strongly conserved canonical active-site amino acid residues Asp45 and Tyr233 hydrogen bond with this critical C-4 axial hydroxyl of D-galactose (**1**).[1] To further probe the role of the equivalent residues Asp37 and Tyr223 in *E. coli* GalK, and potentially decrease the reliance of substrate specificity upon this specific interaction, we report the results of independent saturation mutagenesis at both positions.

The saturation-mutagenesis strategy involved the creation of a mini gene library to introduce codons for all proteinogenic amino acids at each defined position (codon 37 or 223) of the *E. coli galK* gene. By restricting the third codon base to guanine or thymine during this process, most codons rare in *E. coli* along with the *ochre* and *opal* stop codons, were eliminated from the *galK* mini library. By using this approach, 19 substitutions for Asp37 and 19 substitutions for Tyr223 were rapidly constructed, and all 38 *E. coli* GalK variants were subsequently examined for catalytic ability.

All exchanges for Asp37 proved to be detrimental to the enzyme; this underlines a crucial role in the catalytic cycle, most likely upon substrate binding, as previously suggested.[1] Consistent with this result, a recently published study on human GalK revealed that an Asp-to-Ala mutation also inactivates the human enzyme.[7] In contrast, Tyr223 exchanges in the *E. coli* GalK resulted in enzymes with differing degrees of activity toward **1.** Specific activities found were: wild-type (wt): 23.5 U mg^{-1} enzyme, Y223F: 23.42 U mg^{-1}, Y223W: 0.87 U mg^{-1}, Y223L: 0.65 U mg^{-1}, Y223H: 0.61 U mg^{-1}, Y223R: 0.63 U mg^{-1} (Figure 1). This surprising result clearly prompted a thorough reassessment of the role of the putative hydrogen bond between the GalK Y223 hydroxyl and D-Gal C-4-OH as the bulk of the active Y223 GalK variants lack this specific hydrogen-bonding character. Thus, Y223F and Y223W were selected as representative mutants for further substrate profiling and kinetic characterization.

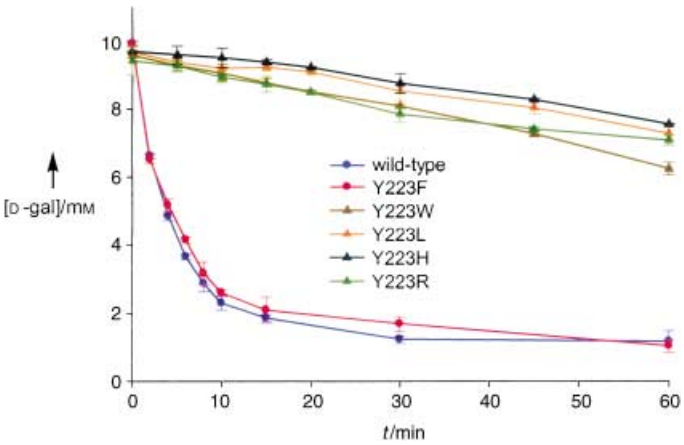


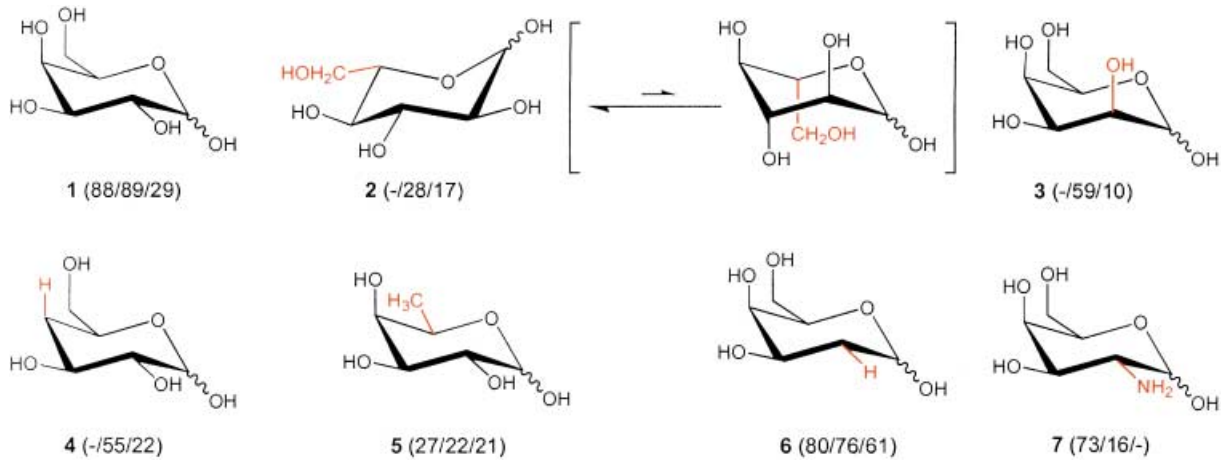
Figure 1. *Progress curves of GalK wt and mutants Y223F, Y223W, Y223L, Y223H, and Y223R for the enzymatic conversion of* ***1*** *into α-D-galactose-1-phosphate.*

Unlike wt, Y223F surprisingly converts L-glucose (**2**), D-talose (**3**), and 4-deoxy-D-galactose (**4**) into their α-1-phosphates (Scheme 1), while maintaining the basal activity toward 6-deoxy-galactose (**5**) inherent to the wt. This extended substrate spectrum holds for the Y223W mutant as well, however at a lower efficiency throughout the series. Substitutions at C-2 are tolerated to different degrees. For example, 2-deoxy-D-galactose (**6**) is an efficient substrate for wt, Y223F, and Y223W, while the catalytic turnover for D-galactosamine (**7**) dropped significantly for both Y223F and Y223W in comparison to wt.

To better understand the impact of these substitutions, the kinetics of Y223F and Y223W were compared to wt GalK (Table 1). For both ATP and **1**, the plot of V_0 against substrate concentration revealed a shallow sigmoidal relation in the wt (Figure 2). This finding contrasts earlier kinetic evaluations and may suggest either an induced fit or allosteric mechanism,[8] or possibly reflect limitations of the assay at low substrate concentrations. Interestingly, while an induced-fit mechanism has been reported for other sugar kinases, for example hexose kinase,[9] such a phenomenon has not yet been reported for *E. coli* GalK. For ATP with the wt enzyme, the calculated $K_{0.5} = 2.24 \pm 0.35$ mM, is in good accordance with previously reported values,[10] while the value for **1** was twofold higher than previous findings ($K_{0.5} = 5.2 \pm 0.21$ mM). The V_{max} value was determined as $V_{max} = 1.84 \pm 0.08$ mM min^{-1}.

[a] *Dr. D. Hoffmeister, Prof. J. S. Thorson*
Laboratory for Biosynthetic Chemistry, School of Pharmacy
University of Wisconsin–Madison
777 Highland Avenue, Madison, WI 53705 (USA)
Fax: (+1) 608-262-5345
E-mail: jsthorson@pharmacy.wisc.edu

 DOI: 10.1002/cbic.200400003



Scheme 1. *Structures of substrates accepted by Y223F or Y223W. Structural deviations from the genuine substrate* ***1*** *are indicated in red. The figures in brackets below the chemical structures indicate the percentage of sugar conversion by wt, Y223F, and Y223W, respectively, after 1 h of enzymatic reaction.*

Table 1. *Kinetic parameters for wt GalK, Y223F, and Y223W. $K_{0.5}$ values are in mM, V_{max} values in mM min^{-1}.*

	wt	Y223F	Y223W
V_{max}	1.84 (±0.08)	1.97 (±0.11)	0.096 (±0.02)
$K_{0.5\ (ATP)}$	2.24 (±0.35)	3.58 (±0.09)	1.78 (±0.16)
$K_{0.5\ (D\text{-}gal)}$	5.2 (±0.21)	4.46 (±0.28)	5.63 (±0.96)

The same sigmoidal relationship for **1** was also found for Y223F and Y223W (Figure 2), with values comparable to wt (for **1** $K_{0.5} = 4.46 \pm 0.28$ mM and 5.63 ± 0.96 mM, respectively, and for ATP $K_{0.5} = 3.58 \pm 0.09$ mM and $K_{0.5} = 1.78 \pm 0.16$ mM, respectively). While the calculated V_{max} only minimally changed for Y223F ($V_{max} = 1.97 \pm 0.11$ mM min^{-1}) compared to wt, we found a drastic decrease for Y223W ($V_{max} = 0.096 \pm 0.02$ mM min^{-1}). From the above findings we conclude that the Y223 side chain hydroxyl-**1** C-4-OH hydrogen bond is not critical for catalysis. Instead, we propose the role of the Y223 to be to participate in a nonpolar stacking interaction, for example as in the *Bacillus stearothermophilus* maltogenic α-amylase,[11] to secure the pyranose ring. Such an interaction might also explain the observed activity for Y223W and, on a significantly reduced level, Y223L and Y223H. Other prominent examples of similar carbohydrate enzyme stacking interactions have been reviewed by Vyas.[12]

While the Y223 side chain hydroxyl-**1** C-4-OH hydrogen bond is clearly not critical for catalysis, our results suggest this putative hydrogen bond might still be essential in controlling substrate specificity; thereby dictating sugar entrance into the Leloir pathway. Interestingly, human and murine *N*-acetylgalactosamine kinases possess a phenylalanine on the position equivalent to 223 in *E. coli* GalK[13] and are, in fact, unspecific in that they also display GalK activity. Our work also shows that not only electrostatics but also steric considerations are critical. For example, the Y223I mutant is completely inactive while Y223L retains basal catalytic activity.

In summary, we reveal that *E. coli* GalK tolerates active-site engineering to expand its sugar-substrate range. To provide either a single variant or a set covering a maximum number of sugar substrates for biotechnological application in glycorandomization, further work on the role of other amino residues along with cloning and substrate profiling of GalKs of other species is in progress.

Experimental Section

Materials and bacterial strains: Chemicals and enzymes were purchased from Promega and Novagen (both Madison, WI), Sigma (St. Louis, MO), Fisher/Acros Organics (Hanover Park, IL), or Fluka (Milwaukee, WI). *E. coli* strains XL1-blue for cloning and BL21-Gold (DE3) for protein over-expression were from Stratagene (La Jolla, CA). The complete series of hexopyranoses (both D- and L-configured) for substrate profiling was of commercial origin, except for D- and L-idose.[6] Modified sugars tested were D-galacturonic acid, D-galactosamine, 2-deoxy-, 4-deoxy-, and 6-deoxy-D-galactose (commercial suppliers as above).

Saturation mutagenesis: Targeted codon replacements within the *galK* gene were accomplished by using the QuikChange II mutagenesis kit (Stratagene). Template DNA was the plasmid pGalK, which is the expression vector pET15b (Novagen) that harbors *E. coli galK* cloned into the *Nde*I and *Bam*HI sites, the primer pair 5′-GATTGGTGAACACACCNNKTACAACGACGGTTTCGTTCTG-3′/5′-CAGAACGAAACCGTCGTTGTAMNNGGTGTGTTCACCAATC-3′ to saturate codon 37, and primer pair 5′-CCCTGGTTGGCAGCGAANNKAACACCCGTCGTGAACAG-3′/5′-CTGTTCACGACGGGTGTTMNNTTCGCTGCCAACCAGGG-3′ for codon 223 of the wt reading frame. Screening for the desired mutants was done by DNA-sequencing on an ABI 310 automatic DNA sequencer (Perkin–Elmer, Foster City, CA).

Protein over-expression and kinase assay/kinetics: For initial prescreening of the minilibrary comprising all variants of *galK* codons 37 and 223, the appropriate *E. coli* transformants were first grown as 1 mL miniature cultures to an OD_{600} ~0.7 in LB-medium selecting with ampicillin (100 μg mL^{-1} final). Protein over-expression was induced by adding IPTG (1 mM final) over 2 h. Cells were harvested by centrifugation. To prescreen for kinase activity supernatants

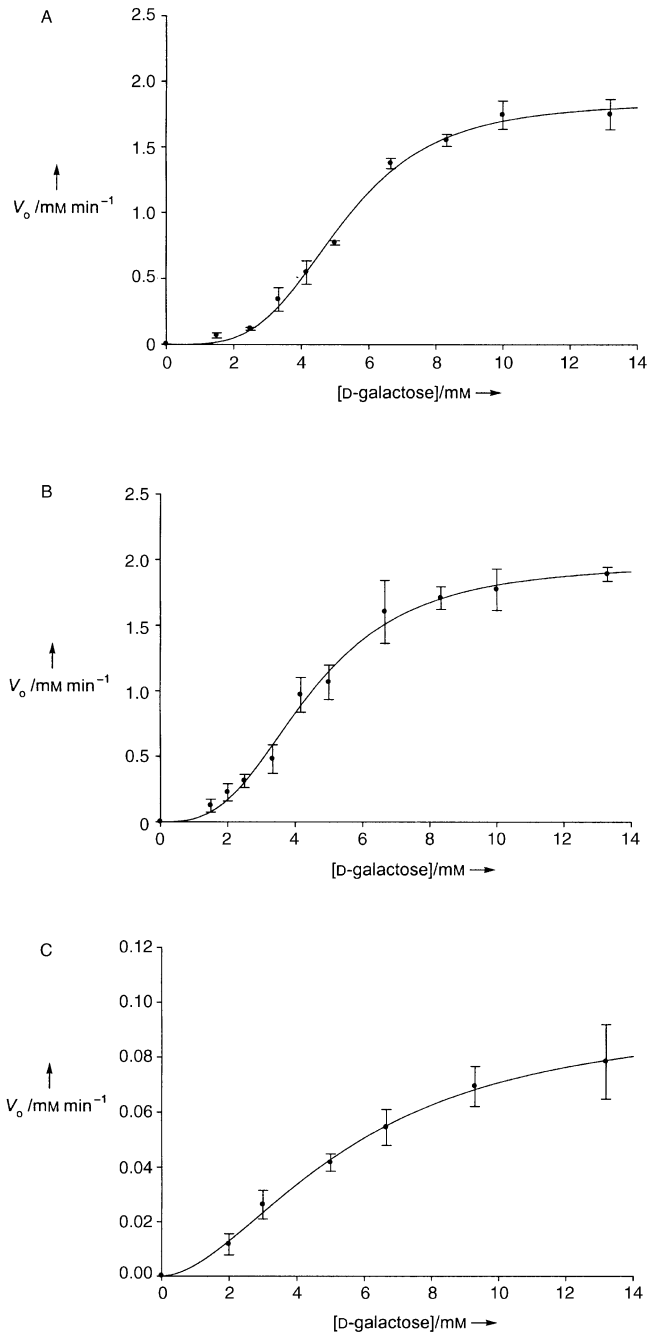


Figure 2. *Plots of initial velocities (V_0) of the kinase reaction for different D-galactose concentrations. A) wild-type, B) Y223F, C) Y223W.*

from lysonase-treated *E. coli* cultures were used. To record kinetic data or time courses of the GalK variants Y223F, Y223W, Y223L, Y223H, and Y223R, 70 mL cultures were used, and the over-expressed protein purified to apparent homogeneity by metal affinity chromatography on nickel-nitrilotriacetic acid (Ni-NTA, QIAGEN, Valencia, CA). The purification was monitored on SDS-polyacrylamide gels. The imidazole was subsequently removed by repeated diafiltration by using Centricon YM-10 (10 kDa exclusion limit) filtration devices (Millipore, Bedford, MA) and sodium phosphate buffer (50 mM, pH 7.5).

Enzymatic and prescreening reactions were performed in multiwell plates on a Biomek FX automated liquid-handling workstation (Beckman Coulter, Fullerton, CA) equipped with a Fluostar Optima plate reader (BMG, Durham, NC) following a previously described protocol[10] but including the following modifications: for specific activities, [ATP] = [D-galactose] = 10 mM. Kinetic data were obtained by determining the slope of the linear phase of the progress curve over 2 min in 30 s intervals and then by calculating the kinetic parameters as described.[10] Enzyme concentrations were determined by the colorimetric assay described by Bradford, with BSA as protein standard.[14]

Note added in proof: While this manuscript was under review, David W. Rice and co-workers reported the crystal structure of the *Pyrococcus furiosus* galactokinase[15] and found Tyr200 (equivalent to Tyr223 in *E. coli* GalK) to hydrogen bond with the C-4-OH of D-galactose.

Acknowledgements

This work was supported in part by the National Institutes of Health grants GM58196, CA84374, and AI52218. We thank the Deutsche Forschungsgemeinschaft (DFG) for postdoctoral fellowship support (D.H.). Lesley Liu (School of Pharmacy, University of Wisconsin–Madison) is gratefully acknowledged for the chemical synthesis of L-idose and D-idose.

Keywords: carbohydrates · enzymes · glycorandomization · kinases · phosphosugars

[1] J. B. Thoden, H. M. Holden, *J. Biol. Chem.* **2003**, *278*, 33305–33311.
[2] a) A. M. Mulichak, J. E. Wilson, K. Padmanabhan, R. M. Garavito, *Nat. Struct. Biol.* **1998**, *5*, 555–560; b) A. E. Aleshin, C. Zeng, G. P. Bourenkov, H. D. Bartunik, H. J. Fromm, R. B. Honzatko, *Structure* **1998**, *6*, 39–50; c) S. Ito, S. Fushinobu, J. J. Jeong, I. Yoshioka, S. Koga, H. Shoun, T. Wakagi, *J. Mol. Biol.* **2003**, *331*, 871–883.
[3] a) J. R. Sherman, J. Adler, *J. Biol. Chem.* **1963**, *238*, 873–878; b) J. S. Gulbinsky, W. W. Cleland, *Biochemistry* **1968**, *7*, 566–575; c) D. B. Wilson, D. S. Hogness, *J. Biol. Chem.* **1969**, *244*, 2137–2142.
[4] a) O. Thalhammer, R. Gitzelmann, M. Pantlitschko, *Pediatrics* **1968**, *42*, 441–445; b) E. Beutler, F. Matsumoto, W. Kuhl, A. Krill, N. Levy, R. Sparkes, M. Degnan, *N. Engl. J. Med.* **1973**, *288*, 1203–1206; c) J. T. Prchal, M. E. Conrad, H. W. Skalka, *Lancet* **1978**, *8054*, 12–13.
[5] a) X. Fu, C. Albermann, J. Jiang, J. Liao, C. Zhang, J. S. Thorson, *Nat. Biotechnol.* **2003**, *21*, 1467–1469; b) J. S. Thorson, W. A. Barton, D. Hoffmeister, C. Albermann, D. B. Nikolov, *ChemBioChem* **2004**, *5*, 16–25; c) J. Yang, D. Hoffmeister, L. Liu, X. Fu, J. S. Thorson, *Bioorg. Med. Chem.* **2004**, *12*, 1577-1584.
[6] D. Hoffmeister, J. Yang, L. Liu, J. S. Thorson, *Proc. Natl. Acad. Sci USA* **2003**, *100*, 13184–13189.
[7] D. J. Timson, R. J. Reece, *BMC Biochem.* **2003**, *4*, 16.
[8] I. H. Segel, *Enzyme Kinetics*, Wiley, New York, **1993**, pp. 346–464.
[9] a) G. De la Fuente, R. Lagunas, A. Sols, *Eur. J. Biochem.* **1970**, *16*, 226–233; b) W. S. Bennett, Jr., T. A. Steitz, *Proc. Natl. Acad. Sci USA* **1978**, *75*, 4848–4852.
[10] J. Yang, X. Fu, Q. Jia, J. Shen, J. B. Biggins, J. Jiang, J. Zhao, J. J. Schmidt, P. G. Wang, J. S. Thorson, *Org. Lett.* **2003**, *5*, 2223–2226.
[11] Z. Dauter, M. Dauter, A. M. Brzozowski, S. Christensen, T. V. Borchert, L. Beier, K. S. Wilson, G. J. Davies, *Biochemistry* **1999**, *38*, 8385–8392.
[12] N. K. Vyas, *Curr. Opin. Struct. Biol.* **1991**, *1*, 732–740.
[13] a) R. T. Lee, C. L. Peterson, A. F. Calman, I. Herskowitz, J. J. O'Donnell, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10887–10891; b) The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team, *Nature* **2002**, *420*, 563–575; c) I. Pastuszak, J. O'Donnell, A. D. Elbein, *J. Biol. Chem.* **1996**, *271*, 23653–23656.

[14] M. Bradford, *Anal. Biochem.* **1976**, *72*, 248 – 254.
[15] A. Hartley, S. E. Glynn, V. Barynin, P. J. Baker, S. E. Sedelnikova, C. Verhees, D. de Geus, J. van der Oost, D. J. Timson, R. J. Reece, D. W. Rice, *J. Mol. Biol.* **2004**, *337*, 387 – 398.

Received: January 7, 2004

