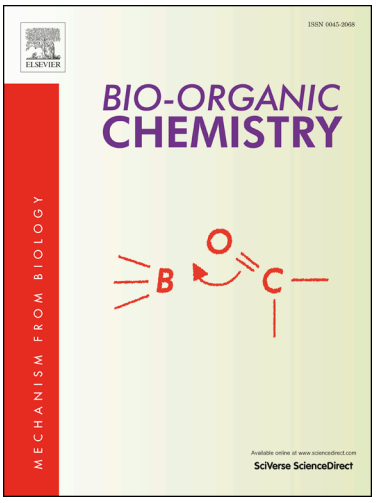
Accepted Manuscript



Modulation of the mobility of a key region in human galactokinase: Impacts on catalysis and stability

Margaret McAuley, Meilan Huang, David J. Timson

PII: S0045-2068(18)30575-3
DOI: https://doi.org/10.1016/j.bioorg.2018.09.002
Reference: YBIOO 2499

To appear in: *Bioorganic Chemistry*

Received Date: 22 June 2018
Revised Date: 29 August 2018
Accepted Date: 6 September 2018

Please cite this article as: M. McAuley, M. Huang, D.J. Timson, Modulation of the mobility of a key region in human galactokinase: Impacts on catalysis and stability, *Bioorganic Chemistry* (2018), doi: https://doi.org/10.1016/j.bioorg.2018.09.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Modulation of the mobility of a key region in human galactokinase: impacts on catalysis and stability

Margaret McAuley[1], Meilan Huang[2] and David J. Timson[1,3]*

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

[2] School of Chemistry and Chemical Engineering, Queen's University Belfast, David Keir Building, Stranmillis Road, Belfast, BT9 5AG. UK.

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

* Author to whom correspondence should be addressed.

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

Telephone +44(0)1273641623

Fax +44(0)1273642090

Email d.timson@brighton.ac.uk

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

Galactokinase catalyses the phosphorylation of α-D-galactose and some structurally related monosaccharides. The enzyme is of interest due to its potential as a biocatalyst for the production of sugar 1-phosphates and due to its involvement in the inherited metabolic disease type II galactosemia. It has been previously shown that a region (residues 231-245) in human galactokinase often has altered mobility when active site residues are varied. We hypothesised that the reverse may be true and that designing changes to this region might affect the functioning of the active site of the enzyme. Focussing on four residues (Leu-231, Gln-242, Glu-244 and Glu-245) we conducted molecular dynamics simulations to explore the effects of changing these residues to glycine or serine. In most cases the variations resulted in local changes to the 231-245 region and global changes to the root mean squared fluctuation (RMSF) of the protein. The four serine variants were expressed as recombinant proteins. All had altered steady state enzyme kinetic parameters with α-D-galactose as a substrate. However, these changes were generally less than ten-fold in magnitude. Changes were also observed with 2-deoxy-α-D-galactose, α-D-galactosamine and α-D-talose as substrates, including (in some cases) loss of detectable activity, suggesting that these variations can tune the specificity of the enzyme. This study demonstrates that activity and specificity of human galactokinase can be modulated by variations designed to affect active site flexibility. It is likely that this principle can be generalised to other enzymes.

Keywords: enzyme engineering; GHMP kinase; molecular dynamics; protein flexibility; α-D-galactose;

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Introduction

It has become well-established that protein mobility is critical for function, including catalysis [1, 2]. This mobility can be local to specific elements of the protein's structure, or more global. In particular the mobility of active site residues can be important in mediating catalysis and determining the specificity of enzymes [1, 3-5]. However, other structural elements can indirectly influence the mobility of the active site and thus need to be considered. Effects on mobility are often over-looked when protein engineering is undertaken to alter the specificity or activity of enzymes. This oversight may be a cause of the failure of some enzyme engineering projects to achieve their full potential.

We have previously shown that several regions which are not part of the active site are critical for controlling specificity and catalysis in human galactokinase (EC 2.7.1.6) [6-10]. This enzyme is a member of the GHMP kinase family and catalyses the first committed step of the Leloir pathway, the ATP-dependent phosphorylation of α-D-galactose [11-13]. The human enzyme is relatively highly selective – only α-D-galactose, α-D-galactosamine, 2-deoxy-α-D-galactose and α-D-talose are known to be substrates [8, 9, 14]. One region, located around Try-379, influences the flexibility of residues which form the active site. This appears to control specificity: variation of Tyr-379 broadens the range of substrate phosphorylated by the enzyme, albeit at a significant cost in terms of activity [9]. This region has also been targeted in bacterial galactokinases which are, in general, more promiscuous than the human enzyme [10]. Variation of the equivalent tryptophan residue in *Lactococcus lactis* and *Escherichia coli* galactokinases have resulted in considerable broadening of the substrate range combined with reasonable activity [15-20]. These broad specificity galactokinases are able to catalyse the site and stereo-specific phosphorylation of a wide range of monosaccharides, reactions which may have utility in *in vitro* glycorandomisation and the subsequent development of new drug-like molecules which incorporate sugar moieties [21].

Two other critical regions have recently been identified in human galactokinase. Residues 174–179 and 231–240 are both affected by variations to active site residues, despite these regions not being part of the active site. Molecular dynamics simulations predicted that the structure and dynamics of both of these regions are disrupted when Arg-37 or Asp-186 are varied. The propensity for α-helical structure in the 231-240 region is reduced by variation of active site residues, whereas the 174-179 region becomes more likely to form an ordered turn rather than random coil [6]. Arg-37 and Asp-186 are implicated in the catalytic mechanism of the enzyme, although their exact roles remain uncertain [6, 22-25]. Similar results were obtained when residues identified as important in the mechanism by a QM-MM study were varied (Arg-105, Glu-174 and Arg-228) [6, 26].

Galactokinase is also of interest due its role in the inherited metabolic disease type II galactosemia (OMIM #230200) [27-31]. This relatively mild genetic disease normally results in childhood cataracts which can be resolved by surgery or the exclusion of galactose (and precursors such as lactose) from the diet. It is generally considered to have minimal longer term consequences, although this assumption has been questioned recently [27, 32]. Further interest comes from the proposal to target galactokinase in the treatment of type I and type III galactosemia [12, 27]. This would, in effect, convert these two severe diseases into the much milder type II galactosemia by reducing the build-up of α-D-galactose 1-phosphate which is assumed to be toxic. Considerable progress has been made in the identification of inhibitors which are selective for human galactokinase and which show some promise in cell culture models [25, 33-36]. There has also been a recent suggestion that galactokinase inhibition may be a viable, novel chemotherapy for some forms of cancer [37].

ACCEPTED MANUSCRIPT

We reasoned that if changes in the active site affected the dynamical and structural properties of other parts of the protein, the reverse may also be true and that this might be a useful approach when attempting the engineer the specificity of this enzyme. In particular, we hypothesised that variations designed to change the flexibility of the region around residues 231-245 may affect the active site with consequences for catalysis. We noted that previous work suggested that alterations to active site residues caused this region to be less likely to adopt an α-helical conformation [6]. Therefore, we focussed on changes likely to reduce the α-helical propensity of this region. We selected four residues in this region of human GALK1 and studied the effects of glycine and serine substitutions by molecular dynamics simulations. The serine variants were also studied in *in vitro* kinetic and stability experiments in order to verify some of the *in silico* predictions. This enabled us to demonstrate the concept that catalysis can be affected by variations designed to affect the dynamics of this protein. The results broadly support our proposition and suggest an approach to enzyme engineering which might be more broadly applied.

Materials and Methods

Molecular dynamics

Molecular dynamics (MD) simulations were carried out essentially as previously described [6, 8] using Amber 10 and the Amber Parm 99 forcefield [38]. The crystal structure of human galactokinase (PDB: 1WUU; chain A) was used as the basis for all simulations [39]. Residue numbering was based on the canonical protein sequence in Uniprot (P51570). This structure was "cleaned" to remove selenomethione residues, fill gaps, convert AMP.PNP to ATP and minimise the energy as previously described [6]. The mutate feature in Biovia Discovery studio was used to generate variant structures. Partial and RESP charges for the ligands (galactose, Mg^{2+} and ATP) were determined using the Gaussian 09 package and antechamber encoded in Amber 10 respectively [26, 40, 41]. Before the simulation, the protein-ATP-galactose complex was soaked in a TIP3P water box (71.071 Å × 80.431 Å × 87.517 Å; minimum distance to the boundary of the protein, 8 Å) using tleap (as encoded in Amber 10). The system was neutralised by adding sodium ions. The system was subjected to 1250 steps of First steepest descent minimisation (1250 steps) was carried out on the system, followed by conjugate gradient minimisation (carried out with a gradient of 0.5 kcal mol^{-1} $Å^{-2}$) [42]. The system was then heated (0-300 K for 50 ps; collision frequency of 5.0 ps^{-1}) using the Langevin dynamics method [43]. The system was equilibrated using an NVT ensemble, a periodic boundary was applied and the system held at 300 K for 50 ps. Production simulation was then carried out at 300 K and 1 atm for 7 ns with a time step of 1 fs. The cut-off distance for Van der Waal's interactions (calculated using the particle mesh Ewald method [44]) was 10 Å. Hydrogen covalent bonds were constrained using SHAKE [45]. This length of simulation is consistent with recent studies on engineered proteins (for example, [46-49]).

Simulations were monitored with perl (encoded in Amber). The ptraj command in Amber 10 was used for secondary structure analysis and RMSF calculations. and structures visualised using Biovia Discovery Studio. Trajectories were visualised using VMD software and the timeline plugin was used to visualise secondary structure [50]. These were calculated using the equilibrated part of the trajectory (typically 2-4 ns), determined by monitoring RMSD over time (Supplementary Fig. S1) [51].

Expression, purification and mutagenesis of human galactokinase

ACCEPTED MANUSCRIPT

Human galactokinase was expressed in, and purified from, *E. coli* HMS174(DE3) as previously described [6, 8, 9, 23, 52]. Briefly, cultures (1 l) of recombinant bacteria were grown, lysed by sonication and galactokinase purified from the resulting extract by cobalt affinity chromatography (His-Select, Sigma, UK). Purified protein was dialysed against 50 mM Hepes-OH, pH 7.5, 150 mM NaCl, 1 mM DTT, 10% (v/v) glycerol and stored, frozen in aliquots, at -80 °C. Mutations of the coding sequence were made using the QuikChange method of site-directed mutagenesis [53]. All coding sequences were verified by DNA sequencing (GATC Biotech, London, UK).

Galactokinase assay and kinetics

Galactokinase activity was assayed by coupling the production of ADP to the reactions catalysed by pyruvate kinase (EC 2.7.1.4) and lactate dehydrogenase (EC 1.1.1.27). This enables the reaction to be followed via the consumption of NADH and the consequent reduction in absorbance at 340 nm [6, 8, 54]. Reactions were carried out in a total volume of 160 µl in 96 well plates at 37 °C. Assay mixes contained 50 mM Hepes-OH, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.4 mM phospho*enol*pyruvate, 1 mM NADH, 10%(v/v) glycerol, pyruvate kinase (7.5 U), lactate dehydrogenase (10 U) and varying, sub-saturating amounts of ATP and galactose. Reactions were initiated by the addition of enzyme (70 nM). They were arrayed in an eight by ten matrix so that galactose concentration varied along one axis and ATP concentration along the other. This enabled the determination of eight Michaelis-Menten curves with sub-saturating galactose concentration and ten with sub-saturating ATP concentrations facilitating subsequent analysis. It is known that human galactokinase has an ordered, ternary complex mechanism [52]. Therefore, each Michaelis-Menten curve was analysed by non-linear curve fitting (as implemented in GraphPad Prism 6.0) to give apparent values for the Michaelis constant ($K_{m,app}$) and the turnover number ($k_{cat,app}$; equal to the maximum rate, $V_{max,app}$, divided by the enzyme concentration). These values were then used to determine the absolute values for the turnover number (k_{cat}) and the two Michaelis constants ($K_{m,ATP}$ and $K_{m,gal}$) according to the equation: $k_{cat,app}=k_{cat}[S]/(K_m+[S])$ where [S] is the substrate concentration [55]. The same approach was applied with monosaccharides other than galactose were used as a substrate.

Analytical methods

Protein concentrations were estimated by the method of Bradford, using BSA as a standard [56]. Protein thermal stability was assessed using differential scanning fluorimetry (DSF) as previously described [6].

Results and Discussion

Variation of residues 231-245 is predicted to affect the active site

We focussed on four residues in this region – Leu-231, Gln-242, Glu-244 and Glu-245 (Figure 1a). Leucine, glutamine and glutamate residues were selected due to their high propensity towards α-helical structure compared to serine and glycine [57]. In the crystal structure of human GALK1, the last three of these residues form part of an α-helix and the first part of a random coil structure which leads into this helix [39]. Leu-231 undergoes substantial structural and dynamic changes in response to active site residue variation, typically becoming more likely to be part of random coil structures

ACCEPTED MANUSCRIPT

[6]. Thus varying them to residues with lower α-helical propensity (i.e. glycine and serine) is likely to disrupt the helix, with potential effects on the active site and catalysis [57, 58]. Since glycine and serine also impart the greatest local flexibility on polypeptide chains, these variations are also likely to increase the mobility of the 231-245 region and may have wider implications for the overall dynamics of the protein [59]. It should be noted that Leu-231 is located in a less well-resolved region of the crystal structure. This residue (and Ser-230) is missing in some chains and was modelled is the structure used in the simulations reported here (and in our previous work) [23]. Thus the predictions concerning this residue should be viewed with greater caution than the others in this study.

Variation of Leu-231 to glycine (L231G) was predicted to result in the initiation of the α-helix at residue 233 and reduced secondary structure propensity at residues 230 and 231 (Figure 2). Analysis of the root mean squared fluctuation (RMSF) for the whole protein revealed that the region 230-245 had substantially increased fluctuation, suggesting that this variation results in substantially greater mobility for the whole helical segment (Figure 3; Supplementary Figure S3). The variation is also predicted to displace the helix comprising residues 260-265, which lies on the surface of the protein (data not shown). Interestingly, this variant is predicted to increase the number of hydrogen bonds but not result in a significant change in the number of salt bridges in the protein (Table 1). The effects of varying this residue to serine (L231S) are less dramatic, as might be expected since serine has lower flexibility than glycine [59]. Like L231G, this variant is predicted to initiate the helix at residue 233 (Figure 2; Figure 4a, left panel). It also results in displacement of the helix comprising residues 171-180 (Figure 4a, right panel). Note that this region includes Glu-174 which has previously been shown to be critical for the catalytic mechanism [6, 26]. The variant is predicted to result in a small reduction in the RMSF of the 230-245 region (Figure 3; Supplementary Figure S3). It is predicted to result in no significant change in in the overall number of salt bridges, but no change in the number of hydrogen bonds (Table 1). A new salt bridge involves a residue in the region under investigation (Glu-244 to Arg-248). This interaction is created in all the variants studied here, except those involving Glu-244. This salt bridge results in a slight "kink" in the end of the α–helix. Paradoxically, it might be expected to rigidify this region of the protein, despite the introduction of a residue (serine) which is generally associated with increased local flexibility. This new interaction may provide a partial explanation for the lack of expected effects on overall flexibility and catalysis (see below) of some of these variants.

Variation of Gln-242 to glycine (Q242G) is predicted to result in initiation of the α-helix at residue 237 (Figure 2). The change is predicted to increase the RMSF towards the C-terminus of the protein (i.e. after residue ~300; Figure 3).It is also predicted to result in no significant change in the number of salt bridges and a decrease in the number of hydrogen bonds (Table 1). Variation of this residue to serine (Q242S) also causes shortening of the α-helix by three residues (Figure 4b). It is predicted to slightly decrease the RMSF in the 230-245 region, and to cause more substantial reductions in this parameter between residues 100 and 200 (Figure 3; Supplementary Figure S3). Q242S is predicted to result in a decrease in the number of salt bridges but no change in the number of hydrogen bonds (Table 1).

When Glu-244 was varied to glycine (E244G), the α-helix is predicted to initiate at residue 237 (Figure 2). Changes also occur in the 171-180 helix, which is displaced compared to the wild-type. The surface-exposed 261-265 helix is elongated to include residue 266 and the 255-259 turn is altered in conformation and displaced. This variant has increased RMSF throughout the whole of the protein, suggesting a much more mobile and flexible structure (Figure 3). The number of salt bridges is predicted not to change, and there is no change in the number of hydrogen bonds (Table 1).

ACCEPTED MANUSCRIPT

Variation of Glu-244 to serine (E244S) resulted in disruption to the first part of the 230-245 region with lower helical propensity in residues 230-236. However, the C-terminal part of this region (which includes the varied residue) was not affected with 100% predicted α-helical conformation (Figure 2; Figure 4c, left panel). The 171-180 helix is also disrupted but there was no predicted change in the 261-266 region (Figure 4c, right panel). However, the overall RMSF of E244S is more similar to the wild-type in magnitude with some differences in the first 180 residues (Figure 3). A decrease in the number of salt bridges and hydrogen bonds is predicted for this variant.

Changing Glu-245 to glycine (E245G) also resulted in initiation of the α-helix at residue 237 (Figure 2). No changes were predicted in the 171-180 region, but similar changes to E244G were observed in the 255-266 region. Interestingly, despite these changes, the RMSF in the 231-245 region was predicted to be reduced (Figure 3; Supplementary Figure S3). The number of salt bridges was reduced (to the greatest extent of all the variants studied here) and the number of hydrogen bonds increased (Table 1). Variation of this residue to serine (E245S) resulted in similar structural changes to E244S, with the 231-240 α-helix beginning at residue 236 (Figure 4d, left panel). Unlike E245G, little change is predicted in the RMSF of the 231-245 region, but the 171-180 region is displaced including the catalytically important Glu-174 (Figure 3; Supplementary Figure S3; Figure 4d, right panel). The number of salt bridges is unchanged and the number of hydrogen bonds is predicted to be increased (Table 1).

In general, changing these residues to ones with lower α-helical propensity and increased backbone flexibility affected other parts of the protein. Interestingly, in the case of changes to Gln-243 and Glu-244, there was no change in the immediate vicinity of the residue and the α-helix in which the residues are situated was not predicted to be disrupted. Most of the variants resulted in a later start to this α-helix and an increased amount of other forms of secondary structure, including 3_{10} helices, in residues 230-236. There was no clear link between predicted changes in RMSF and the numbers of predicted salt bridges and hydrogen bonds.

Activity and stability of four variants with altered mobility in the 230-245 region

Previous work has shown that human galactokinase appears to be particularly sensitive to point mutations, with many resulting in insoluble or inactive enzyme [14, 52, 60]. Given this sensitivity, we focussed on the serine variants (L231S, Q242S, E244S and E245S) for experimental studies since these, in general, resulted in less radical changes to the overall dynamics of the protein [14, 52, 60]. Each of these four variants could be expressed in, and purified from, *E. coli* with yields in the range 1.0-1.5 mg protein per litre of culture (Supplementary Figure S2).

All four variants were active galactokinases (Table 2; Figure 5). Alteration of the dynamic behaviour of a helix adjacent to the active site would be expected to change the kinetic constants of the enzyme. Broadly this is what was observed. E244S had a significantly increased (five-fold) Michaelis constant for ATP, indicating that this variation impacts on the protein's interaction with ATP. Both E244S and E245S had significantly increased Michaelis constants for galactose indicating that variation of these two residues impacts galactose binding. The turnover number was significantly increased for E244S, E245S and Q242S. However, catalytic efficiency (as measured by the specificity constants, k_{cat}/K_m) was only significantly affected for galactose in E245S and L231S. In both cases this measure was reduced indicating that these variants are less efficient at catalysing the phosphorylation of galactose than the wild-type. Interestingly, in E244S, E245S and Q242S there is a trade-off between increased turnover number and increased Michaelis constants resulting in variant

ACCEPTED MANUSCRIPT

enzymes with either similar or slightly reduced catalytic efficiency when compared to wild-type galactokinase. A similar trade-off was seen in previous studies on consensus variants [8].

We also tested the ability of these four variants to catalyse the phosphorylation of other known substrates of wild-type galactokinase – 2-deoxy-α-D-galactose, α-D-galactosamine and α-D-talose (the C_2-epimer of galactose; Figure 1b). With 2-deoxy-D-galactose, there was no significant change in the Michaelis constants for either ATP or the monosaccharide substrate with any of the variants (Table 2; Supplementary Figure S4). For Q242S the turnover number and the catalytic efficiency with ATP were increased, whereas L231S had reduced turnover and decreased catalytic efficiency with both substrates. Both L231S and Q242S were inactive with D-galactosamine, indicating these variations have implications for which substrates can be processed by the active site. The two variants that were active with D-galactosamine were affected differently. E245S was only significantly impacted in terms of the Michaelis constant for the sugar which was more than doubled (Table 2; Supplementary Figure S5). E244S is more interesting, the Michaelis constant for ATP was significantly increased (four-fold), but k_{cat} was nine-fold higher resulting in a slight, but not significant, increase in specificity constant for ATP (three-fold) and a significant increase (almost seven-fold) in the specificity constant for the sugar. Variation of the 231-240 helix also impacts the phosphorylation of α-D-talose (Table 2; Supplementary Figure S6). L231S was inactive with this sugar indicating that variation of this residue has implications for access of substrates to the active site. Compared to the wild-type, E244S was only significantly affected in terms of catalytic efficiency of the sugar (seven-fold decrease). E245S had a significant increase in the Michaelis constant for ATP (almost four-fold), a significant increase in turnover number (over ten-fold), yet a significant reduction on catalytic efficiency for both ATP and the sugar (five-fold and seven-fold, respectively). Q242S had an increased Michaelis constant for the sugar (three-fold) and a doubled (but not significantly different) turnover number. Both of these factors contribute to give different implications for the specificity constants: the catalytic efficiency for ATP was significantly increased (three-fold), yet the catalytic efficiency for the sugar was significantly decreased (two-fold).

The variants were also tested with other monosaccharide sugars as possible substrates. However, like the wild-type, none of the variants showed any detectable activity with D-mannose, D-glucose, D-fructose, L-glucose, L-mannose or L-arabinose as substrates (0.3 µM enzyme, 80 µM ATP and 2 mM monosaccharide; data not shown). Several of these sugars have previously been shown to be substrates of different variant forms of human galactokinase [8, 9].

In addition to steady-state kinetics, we also investigated the stability of the four variants in terms of their resistance to chaotropic denaturation and to thermal denaturation. The stability of each variant towards denaturants was determined by incubation with increasing concentrations of urea (Figure 6). The wild-type proteins gave a discrete band up to 1 M Urea. Only L231S behaved in a similar manner. This is likely to be due to the localised effect of this variant on a small portion of the helix, as predicted from analysis of secondary structure. Q242S was stable up to 0.5 M Urea; this is likely to be the result of more substantial changes to the protein's structure. E244S and E245S are unstable under these conditions even in the absence of urea, indicating a large effect of these variants on the protein, consistent with the molecular dynamics analysis of these two variants. The melting temperature of each variant was determined using differential scanning flourimetry (Table 3). Broadly, results from this assay agree with those from the urea denaturation experiment. The melting temperature (T_m) of L231S was not significantly different to the wild-type protein. Both E244S and E245S have significantly reduced T_m values compared to the wild-type. However, Q242S has the lowest T_m value despite being relatively stable towards denaturation by urea.

ACCEPTED MANUSCRIPT

Conclusions

Variations to the 231-245 region of human galactokinase do have an effect on catalysis at the active site. However, they did not result in a broadening of the substrate specificity based on the range of monosaccharides tested. The effects on catalysis were variable and not clearly correlated with changes predicted in the molecular dynamics simulations. In general the effects were relatively small: almost all changes in steady-state parameters were less than ten-fold. Furthermore, the effects were not consistent between different monosaccharide substrates. Indeed some variants had no detectable activity with α-D-galactosamine and α-D-talose, despite retaining reasonable activity with α-D-galactose (Table 2). Human galactokinase is only known to show activity with α-D-galactose and monosaccharides which differ at carbon-2 of the pyranose ring. However, these variants appear to be able to tune the selectivity of the enzyme at this position enabling it to discriminate in favour of α-D-galactose. The effects of these variations on protein stability were not straightforward. Although the four serine variants all have reduced overall flexibility (as judged by RMSF plots, Figure 3; Supplementary Figure S3), three of them (Q242S, E244S and E245S) have significantly reduced thermal stability (Table 2). Two (Q242S and E245S) are also less stable to chaotropic denaturation. This suggests that reduced overall flexibility does not, necessarily, correlate with increased overall stability in human galactokinase which is consistent with studies on other enzymes [61]. Overall, these results support the hypothesis that variation of residues in this region have effects on catalysis and overall protein stability. They also suggest that using molecular dynamics to identify sites in proteins which modulate global or local mobility may be a valuable approach to identify strategies to improve activity or change substrate ranges. Such strategies might be used in conjunction with more "traditional" enzyme engineering approaches such as the rational modification of residues at the active site.

Acknowledgements

MM was in receipt of a PhD studentship funded by the Department for Employment and Learning, Northern Ireland (UK). MH acknowledges financial support from InvestNI (Northern Ireland, UK; grant number: RD0314092). We thank Queen's University Belfast for access to the High Performance Computing Cluster and Prof Aaron Maule (School of Biological Sciences, Queen's University Belfast) for access to qPCR machines used for DSF analyses.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure legends

Figure 1: (a) Wild type GALK structure (PDB: 1WUU [39]) showing residues studied in this paper. Leu-231 is shown in red, Gln-242 in orange, Glu-244 in purple and Glu-245 in blue. The substrates (ATP and galactose) are shown in the active site in stick format. (b) Structures of the monosaccharide substrates of human galactokinase.

Figure 2: Molecular dynamics predicts effects on key regions of human galactokinase. The percentage time spent in different forms of secondary structure for residues 230-244 in the wild-type and variant forms of human galactokinase is shown in histogram format. These were determined using the every tenth frame from the entire stable trajectory and the 'secstruct' command of 'ptraj' in AmberTools.

Figure 3: Root mean square fluctuation (RMSF) of wild-type and variant galactokinases calculated using the 'atomic fluct' command of 'ptraj' in AmberTools on 2 ns of stable simulation. Wild-type values are shown in dashed red lines on each graph for comparison.

Figure 4: Key predicted, structural changes in variants of human galactokinase. (a) Predicted changes in the L231S variant. Left, the 230-240 helix begins at residue 231, not 230. Right, the 171-180 helix adjacent to the active site is disrupted. The catalytically important residue Glu-174 is shown in stick form. (b) Predicted changes in the Q242S variant. The 230-240 helix begins at residue 233, not 230. (c) Predicted changes in the E244S variant. Left, the 230-240 helix begins at residue 237, not 230. Right, the 171-180 helix adjacent to the active site is disrupted. The catalytically important residue Glu-174 is shown in stick form. (d) Predicted changes in the E245S variant. Left, the 230-240 helix begins at residue 236, not 230 and the helix is displaced compared to the wild-type protein. Right, the 171-180 helix adjacent to the active site is disrupted. The catalytically important residue Glu-174 is shown in stick form. These figures show the mean structure obtained using representative testcluster analysis of 2 equilibrated ns of simulation. Five average clusters were produced from this timeframe and the cluster which represented the most frames chosen for comparison and display.

Figure 5: Steady-state kinetics analysis of wild-type and variant galactokinase with D-galactose as a substrate. Apparent V_{max} values are plotted against substrate concentration for each of the variants. The lines represent the non-linear curve fit calculated using GraphPad Prism 6.0. Values shown are means ± standard error. For all the variants 0.07 µM enzyme was used.

Figure 6: Native PAGE showing the results of exposure of wild-type and variant galactokinases to increasing concentrations of urea. Enzymes (4.5 µM) were incubated at 37 °C for 30 min with increasing concentrations of Urea (shown above the gels), before loading onto a 15% native gel (pH 8.8) and electrophoresed at 20 mA for 3 h. A discrete band indicates a largely folded, non-denatured protein. The disappearance of this band, or the blurring of the band is associated with chaotropic denaturation of the protein.

ACCEPTED MANUSCRIPT

Tables

Table 1: The number of predicted hydrogen bonds and salt bridges in wild-type and variant galactokinases.

Variant	Number of salt bridges	Number of hydrogen bonds
Wild-type	34 ± 4.7^{ns}	87.3 ± 7.4
L231G	33.8 ± 2.4^{ns}	$89.4 \pm 7.6^{***}$
L231S	31.6 ± 1.8^{ns}	87.9 ± 7.7^{ns}
Q242G	32.2 ± 2.4^{ns}	$82.8 \pm 7.4^{***}$
Q242S	$30.4 \pm 3.6^{*}$	87.6 ± 7.3^{ns}
E244G	34.8 ± 3.6^{ns}	88.0 ± 7.7^{ns}
E244S	$30.2 \pm 2.8^{*}$	$84.5 \pm 7.4^{***}$
E245G	$28.4 \pm 1.1^{***}$	$89.9 \pm 7.9^{***}$
E245S	32.2 ± 2.3^{ns}	$90.2 \pm 7.7^{***}$

The number of salt bridges was calculated using five frames from 1 ns of equilibrated trajectory, shown here are means and errors estimated as standard deviations of these means. Mean numbers of hydrogen bonds present in each variant were calculated using 1 ns of stable trajectory and the 'hydrogen bonds' plugin in VMD. Error values shown are standard deviation. Significance of both data sets was determined using a one-way ANOVA and compared to WT using Dunnett's post-test analysis in GraphPad Prism. ns, not significant; *** $p \leq 0.001$ (n=1000); * $p = \leq 0.05$.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2: Steady state kinetic data for variant galactokinases with various substrates.

Monosaccharide substrate	Variant	$K_{m,ATP}$ (μM)	$K_{m,sugar}$ (μM)	k_{cat} (s^{-1})	$k_{cat}/K_{m,ATP}$ ($mM^{-1}s^{-1}$)	$k_{cat}/K_{m,sugar}$ ($mM^{-1}s^{-1}$)
D-Galactose	Wild-type	3.9 ± 1.1	170 ± 30	1.7 ± 0.1	430 ± 110	10.1 ± 1.2
	L231S	2.4 ± 0.5	260 ± 80	1.0 ± 0.1	360 ± 70	3.9 ± 1.1**
	Q242S	9.4 ± 3.4	440 ± 70	5.9 ± 0.3***	610 ± 200	13.5 ± 1.5
	E244S	19.8 ± 7.1*	1000 ± 200**	6.1 ± 0.4***	310 ± 90	6.2 ± 0.9
	E245S	13.2 ± 2.6	1100 ± 120**	4.0 ± 0.2***	300 ± 50	3.9 ± 0.3**
2-Deoxy-D-galactose	Wild-type	43.2 ± 5.6	970 ± 130	6.8 ± 0.4	190 ± 10	7.1 ± 0.6
	L231S	14.8 ± 10.3	1100 ± 280	0.4 ± 0.1***	30 ± 18**	0.5 ± 0.1***
	Q242S	44.6 ± 13.4	2700 ± 990	15.0 ± 4.0*	420 ± 110*	5.6 ± 0.7
	E244S	36.7 ± 4.4	1800 ± 280	8.0 ± 1.0	210 ± 20	4.5 ± 0.5*
	E245S	39.3 ± 12.2	1900 ± 540	11.0 ± 2.0	320 ± 60	5.8 ± 0.7
D-Galactosamine	Wild-type	16.2 ± 5.3	710 ± 180	2.1 ± 0.2	110 ± 250	2.9 ± 0.5
	E244S	59.4 ± 25.2*	920 ± 150	18.1 ± 1.3***	360 ± 80	19.7 ± 1.9***
	E245S	7.7 ± 4.0	2900 ± 910*	3.4 ± 0.5	440 ± 180	1.2 ± 0.1
D-Talose	Wild-type	16.2 ± 5.3	710 ± 180	1.7 ± 0.2	110 ± 25	2.9 ± 0.5

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Q242S	12.8 ± 2.7	2300 ± 470*	3.4 ± 0.2	380 ± 70**	1.5 ±0.3*
E244S	6.6 ± 4.1	830 ± 380	0.5 ± 0.0	73 ± 42	0.4 ± 0.1***
E245S	59.4 ± 2.5***	920 ± 150	18.1 ± 1.3***	20 ± 2*	0.4 ± 0.1***

All measurements taken using 0.07 µM of enzyme. Errors shown are standard errors, parameters were all estimated by fitting graphs of $V_{max, app}$ against substrate concentration as described in Materials and Methods using GraphPad Prism version 6.0. Significance calculated via one-way ANOVA and Dunnett post-test analysis to compare with wild-type parameters. * p≤0.05; ** p≤0.01; *** p≤0.001; all other values not significantly different to wild-type. Results with the wild-type enzyme and these substrates have been reported previously [6, 8].

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3: Stability of variant galactokinases, as measured by differential scanning fluorimetry

Variant	T_m (°C)
Wild-type	56.1 ± 0.3
L231S	55.8 ± 0.1^{ns}
Q242S	52.7 ± 0.1***
E244S	55.1 ± 0.7**
E245S	54.0 ± 0.4***

Variant enzyme (4.5 μM) was mixed with 5× sypro orange (manufacturer's concentration definition). Readings were taken using a Rotor Gene Q qPCR machine from 25 °C to 95 °C in increments of 1 °C and steps of 5 s. Melt analysis was used to determine T_m values. Errors shown are standard deviation and significance was determined using a one-way ANOVA in GraphPad Prism (n=9). ns, not significant; ** p≤0.01; *** p≤0.001. The result with the wild-type enzyme has been reported previously [6].

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

[1] N.M. Goodey, S.J. Benkovic, Allosteric regulation and catalysis emerge via a common route, Nature chemical biology 4(8) (2008) 474-482.
[2] M. McAuley, D.J. Timson, Modulating Mobility: a Paradigm for Protein Engineering?, Appl Biochem Biotechnol 181(1) (2016) 83-90.
[3] P.K. Agarwal, Role of protein dynamics in reaction rate enhancement by enzymes, J Am Chem Soc 127(43) (2005) 15248-56.
[4] E.Z. Eisenmesser, O. Millet, W. Labeikovsky, D.M. Korzhnev, M. Wolf-Watz, D.A. Bosco, J.J. Skalicky, L.E. Kay, D. Kern, Intrinsic dynamics of an enzyme underlies catalysis, Nature 438(7064) (2005) 117-21.
[5] A. Stank, D.B. Kokh, J.C. Fuller, R.C. Wade, Protein Binding Pocket Dynamics, Acc Chem Res 49(5) (2016) 809-15.
[6] M. McAuley, M. Huang, D.J. Timson, Insight into the mechanism of galactokinase: Role of a critical glutamate residue and helix/coil transitions, Biochim Biophys Acta 1865(3) (2017) 321-328.
[7] M. McAuley, H. Kristiansson, M. Huang, A.L. Pey, D.J. Timson, Galactokinase promiscuity: a question of flexibility?, Biochem.Soc.Trans. 44(1) (2015) 116-122.
[8] M. McAuley, N. Mesa-Torres, A. McFall, S. Morris, M. Huang, A.L. Pey, D.J. Timson, Improving the Activity and Stability of Human Galactokinase for Therapeutic and Biotechnological Applications, Chembiochem 19(10) (2018) 1088-1095.
[9] H. Kristiansson, D.J. Timson, Increased promiscuity of human galactokinase following alteration of a single amino acid residue distant from the active site, Chembiochem : a European journal of chemical biology 12 (2011) 2081-2087.
[10] H. Kristiansson, D.J. Timson, Galactokinases: potential biotechnological applications as biocatalysts, Current Biotechnology 1 (2012) 148-154.
[11] H.M. Holden, J.B. Thoden, D.J. Timson, R.J. Reece, Galactokinase: structure, function and role in type II galactosemia, Cellular and molecular life sciences : CMLS 61(19-20) (2004) 2471-2484.
[12] D.J. Timson, GHMP kinases - structures, mechanisms and potential for therpeutically relevant inhibition, Current Enzyme Inhibition 3 (2007) 77-94.
[13] P. Bork, C. Sander, A. Valencia, Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases, Protein science : a publication of the Protein Society 2(1) (1993) 31-40.
[14] D.J. Timson, R.J. Reece, Sugar recognition by human galactokinase, BMC biochemistry [computer file] 4(1) (2003) 16.
[15] D. Hoffmeister, J.S. Thorson, Mechanistic implications of *Escherichia coli* galactokinase structure-based engineering, Chembiochem 5(7) (2004) 989-992.
[16] D. Hoffmeister, J. Yang, L. Liu, J.S. Thorson, Creation of the first anomeric D/L-sugar kinase by means of directed evolution, Proceedings of the National Academy of Sciences of the United States of America 100(23) (2003) 13184-13189.
[17] J.S. Thorson, W.A. Barton, D. Hoffmeister, C. Albermann, D.B. Nikolov, Structure-based enzyme engineering and its impact on in vitro glycorandomization, Chembiochem : a European journal of chemical biology 5(1) (2004) 16-25.
[18] J. Yang, X. Fu, Q. Jia, J. Shen, J.B. Biggins, J. Jiang, J. Zhao, J.J. Schmidt, P.G. Wang, J.S. Thorson, Studies on the substrate specificity of *Escherichia coli* galactokinase, Organic letters 5(13) (2003) 2223-2226.
[19] J. Yang, X. Fu, J. Liao, L. Liu, J.S. Thorson, Structure-Based Engineering of *E. coli* Galactokinase as a First Step toward In Vivo Glycorandomization, Chemistry & biology 12(6) (2005) 657-664.
[20] J. Yang, L. Liu, J.S. Thorson, Structure-based enhancement of the first anomeric glucokinase, Chembiochem : a European journal of chemical biology 5(7) (2004) 992-996.
[21] J. Yang, D. Hoffmeister, L. Liu, X. Fu, J.S. Thorson, Natural product glycorandomization, Bioorganic & medicinal chemistry 12(7) (2004) 1577-1584.

ACCEPTED MANUSCRIPT

[22] X. Chu, N. Li, X. Liu, D. Li, Functional studies of rat galactokinase, Journal of Biotechnology 141(3-4) (2009) 142-146.
[23] C.F. Megarity, M. Huang, C. Warnock, D.J. Timson, The role of the active site residues in human galactokinase: Implications for the mechanisms of GHMP kinases, Bioorganic Chemistry 39(3) (2011) 120-126.
[24] L.A. Reinhardt, J.B. Thoden, G.S. Peters, H.M. Holden, W.W. Cleland, pH-rate profiles support a general base mechanism for galactokinase (*Lactococcus lactis*), FEBS letters 587(17) (2013) 2876-2881.
[25] M. Tang, K. Wierenga, L.J. Elsas, K. Lai, Molecular and biochemical characterization of human galactokinase and its small molecule inhibitors, Chemico-biological interactions 188(3) (2010) 376-385.
[26] M. Huang, X. Li, J.W. Zou, D.J. Timson, Role of Arg228 in the Phosphorylation of Galactokinase: The Mechanism of GHMP Kinases by Quantum Mechanics/Molecular Mechanics Studies, Biochemistry 52(28) (2013) 4858-4868.
[27] A.M. Bosch, H.D. Bakker, A.H. van Gennip, J.V. van Kempen, R.J. Wanders, F.A. Wijburg, Clinical features of galactokinase deficiency: a review of the literature, Journal of inherited metabolic disease 25(8) (2002) 629-634.
[28] D.J. Timson, The molecular basis of galactosemia - Past, present and future, Gene 589 (2016) 133-141.
[29] R. Gitzelmann, Hereditary Galactokinase Deficiency, a Newly Recognized Cause of Juvenile Cataracts, Pediatric research 1(1) (1967) 14-23.
[30] W.R. Pickering, R.R. Howell, Galactokinase deficiency: clinical and biochemical findings in a new kindred, The Journal of pediatrics 81(1) (1972) 50-55.
[31] O. Thalhammer, R. Gitzelmann, M. Pantlitschko, Hypergalactosemia and galactosuria due to galactokinase deficiency in a newborn, Pediatrics 42(3) (1968) 441-445.
[32] J.B. Hennermann, P. Schadewaldt, B. Vetter, Y.S. Shin, E. Monch, J. Klein, Features and outcome of galactokinase deficiency in children diagnosed by newborn screening, Journal of inherited metabolic disease 34(2) (2011) 399-407.
[33] L. Liu, M. Tang, M.J. Walsh, K.R. Brimacombe, R. Pragani, C. Tanega, J.M. Rohde, H.L. Baker, E. Fernandez, B. Blackman, J.M. Bougie, W.H. Leister, D.S. Auld, M. Shen, K. Lai, M.B. Boxer, Structure activity relationships of human galactokinase inhibitors, Bioorganic & medicinal chemistry letters 25(3) (2015) 721-727.
[34] S. Odejinmi, R. Rascon, M. Tang, H. Vankayalapati, K. Lai, Structure-activity analysis and cell-based optimization of human galactokinase inhibitors, ACS medicinal chemistry letters 2(9) (2011) 667-672.
[35] K.J. Wierenga, K. Lai, P. Buchwald, M. Tang, High-throughput screening for human galactokinase inhibitors, Journal of Biomolecular Screening : the official journal of the Society for Biomolecular Screening 13(5) (2008) 415-423.
[36] K. Lai, M.B. Boxer, A. Marabotti, GALK inhibitors for classic galactosemia, Future medicinal chemistry 6(9) (2014) 1003-1015.
[37] M. Tang, E. Etokidem, K. Lai, The Leloir Pathway of Galactose Metabolism - A Novel Therapeutic Target for Hepatocellular Carcinoma, Anticancer Res 36(12) (2016) 6265-6271.
[38] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, Journal of computational chemistry 25(9) (2004) 1157-1174.
[39] J.B. Thoden, D.J. Timson, R.J. Reece, H.M. Holden, Molecular structure of human galactokinase: implications for type II galactosemia, Journal of Biological Chemistry 280(10) (2005) 9662-9670.
[40] M.J. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, Gaussian 09, revision A. 1, Gaussian Inc., Wallingford, CT, 2009.
[41] D. Case, T. Darden, T. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, M. Crowley, R. Walker, W. Zhang, AMBER 10 University of California, San Francisco, 2008.

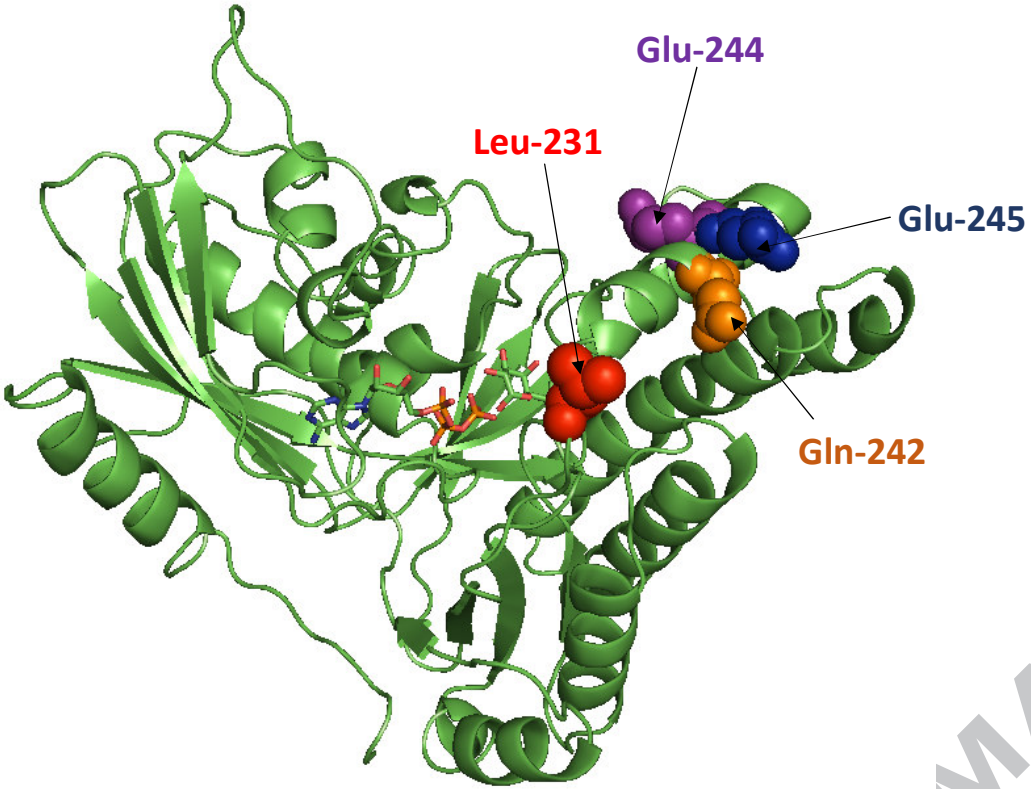
ACCEPTED MANUSCRIPT

[42] P.K. Weiner, P.A. Kollman, AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions, Journal of Computational Chemistry 2(3) (1981) 287-303.
[43] R.W. Pastor, B.R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms, Molecular Physics 65(6) (1988) 1409-1419.
[44] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, The Journal of chemical physics 98(12) (1993) 10089-10092.
[45] J. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, Journal of Computational Physics 23(3) (1977) 327-341.
[46] F. Chaudhry, A.Q. Ather, M.J. Akhtar, A. Shaukat, M. Ashraf, M. Al-Rashida, M.A. Munawar, M.A. Khan, Green synthesis, inhibition studies of yeast α-glucosidase and molecular docking of pyrazolylpyridazine amines, Bioorg Chem 71 (2017) 170-180.
[47] D. Chand, P. Panigrahi, N. Varshney, S. Ramasamy, C.G. Suresh, Structure and function of a highly active Bile Salt Hydrolase (BSH) from *Enterococcus faecalis* and post-translational processing of BSH enzymes, Biochim Biophys Acta 1866(4) (2018) 507-518.
[48] D. Lou, Y. Wang, J. Tan, L. Zhu, S. Ji, B. Wang, Functional contribution of coenzyme specificity-determining sites of 7α-hydroxysteroid dehydrogenase from *Clostridium absonum*, Comput Biol Chem 70 (2017) 89-95.
[49] A. Adamu, R.A. Wahab, M.S. Shamsir, F. Aliyu, F. Huyop, Deciphering the catalytic amino acid residues of l-2-haloacid dehalogenase (DehL) from *Rhizobium* sp. RC1: An *in silico* analysis, Comput Biol Chem 70 (2017) 125-132.
[50] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, Journal of Molecular Graphics 14(1) (1996) 33-8, 27-8.
[51] A. Bakan, L.M. Meireles, I. Bahar, ProDy: protein dynamics inferred from theory and experiments, Bioinformatics 27(11) (2011) 1575-7.
[52] D.J. Timson, R.J. Reece, Functional analysis of disease-causing mutations in human galactokinase, European journal of biochemistry / FEBS 270(8) (2003) 1767-1774.
[53] W. Wang, B.A. Malcolm, Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis, BioTechniques 26(4) (1999) 680-682.
[54] A. Platt, H.C. Ross, S. Hankin, R.J. Reece, The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase, Proceedings of the National Academy of Sciences of the United States of America 97(7) (2000) 3154-3159.
[55] D.J. Timson, Quantitative enzymology, Current Enzyme Inhibition 11(1) (2015) 12-31.
[56] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Analytical Biochemistry 72 (1976) 248-254.
[57] C.N. Pace, J.M. Scholtz, A helix propensity scale based on experimental studies of peptides and proteins, Biophys J 75(1) (1998) 422-7.
[58] P.Y. Chou, G.D. Fasman, Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins, Biochemistry 13(2) (1974) 211-22.
[59] F. Huang, W.M. Nau, A conformational flexibility scale for amino acids in peptides, Angewandte Chemie (International ed.in English) 42(20) (2003) 2269-2272.
[60] F. Sangiuolo, M. Magnani, D. Stambolian, G. Novelli, Biochemical characterization of two GALK1 mutations in patients with galactokinase deficiency, Human mutation 23(4) (2004) 396.
[61] A. Karshikoff, L. Nilsson, R. Ladenstein, Rigidity versus flexibility: the dilemma of understanding protein thermal stability, Febs j 282(20) (2015) 3899-917.

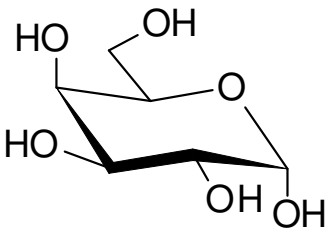
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

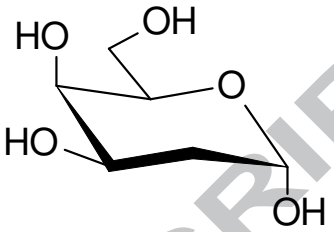
(a)



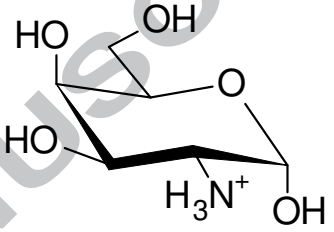
(b)



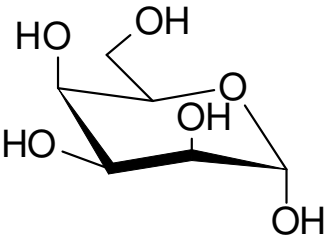
α-D-galactose



2-dexoy-α-D-galactose



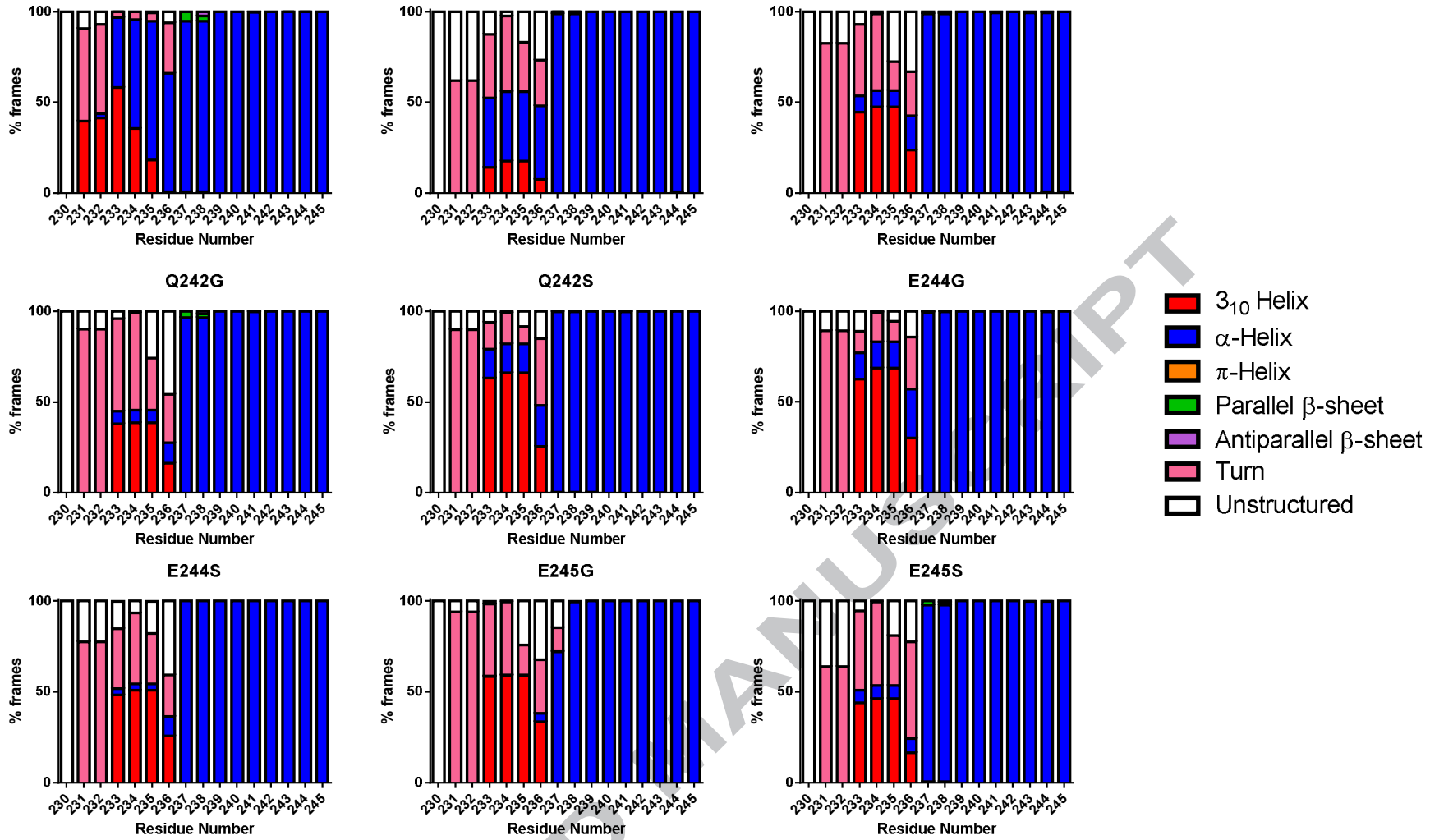
α-D-galactosamine



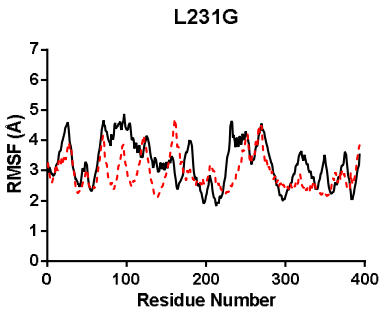
α-D-talose

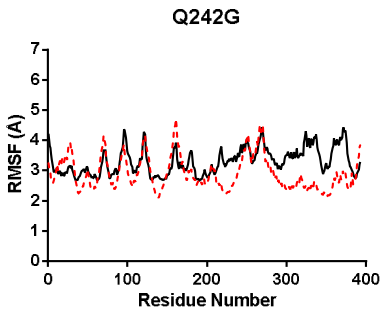
ACCEPTED MANUSCRIPT

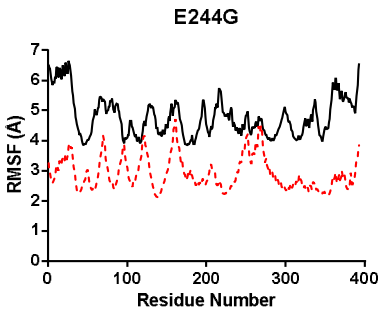
ACCEPTED MANUSCRIPT

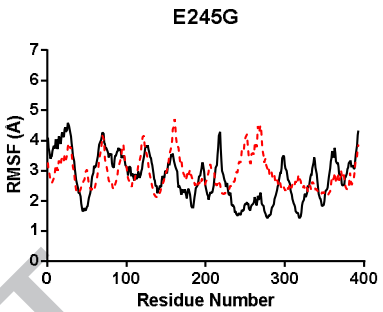


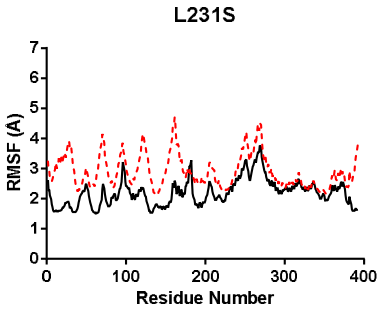
ACCEPTED MANUSCRIPT

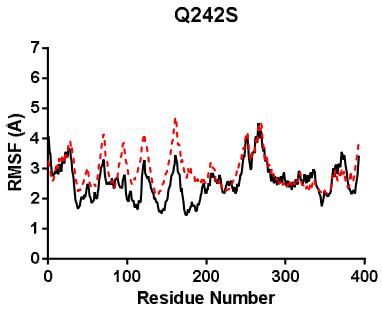


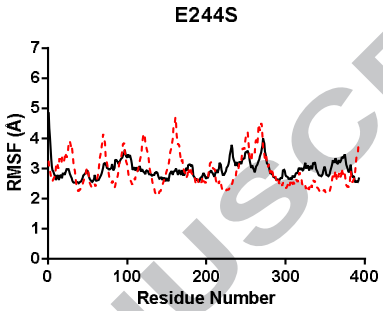


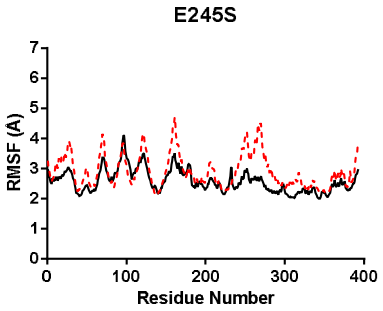








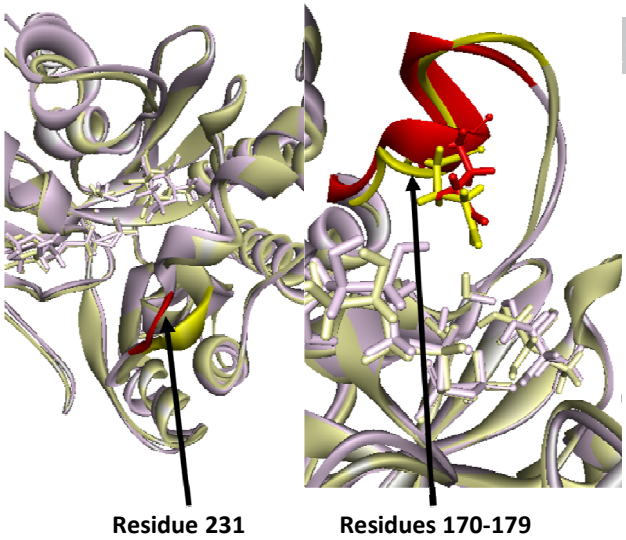




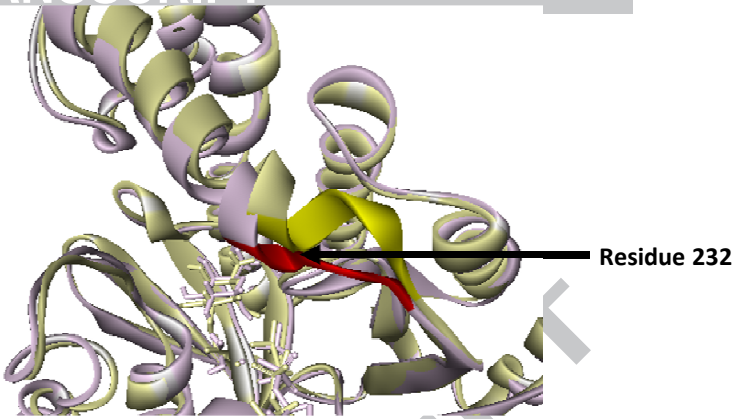
ACCEPTED MANUSCRIPT

CCEPTED MANUSCRIPT

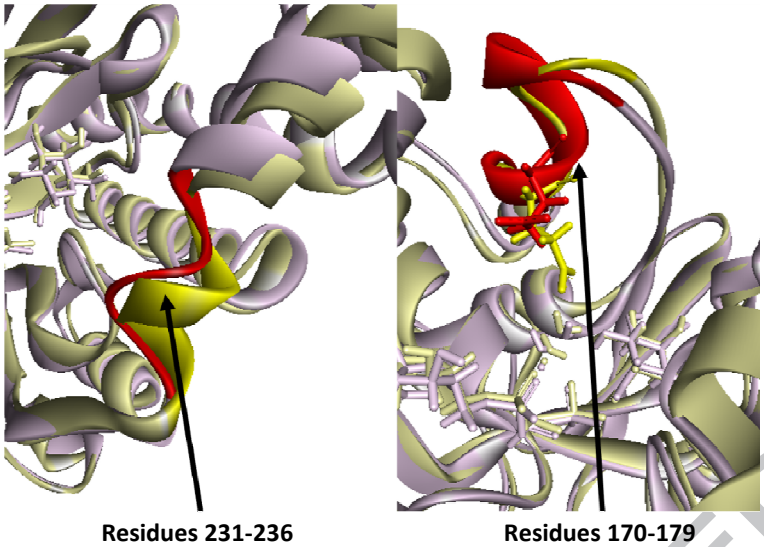
(a)



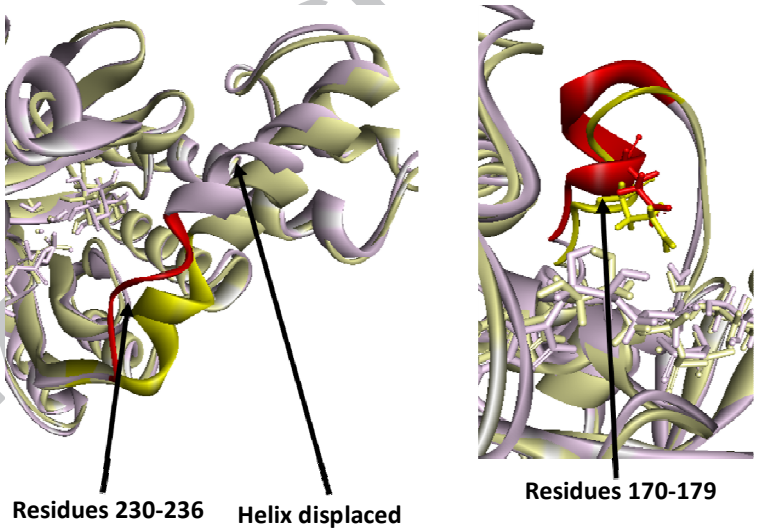
(b)



(c)

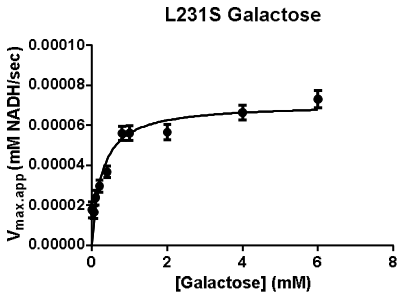


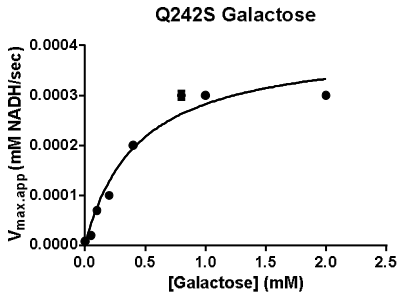
(d)

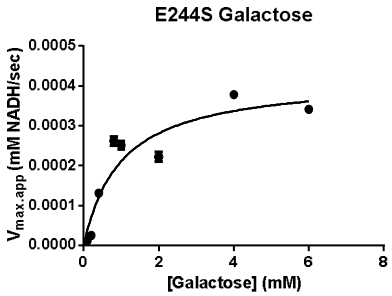


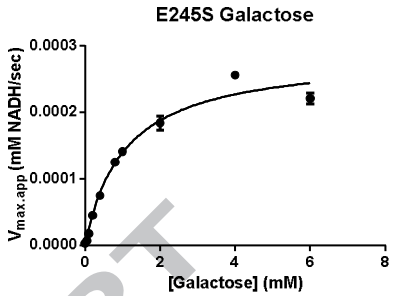
ACCEPTED MANUSCRIPT

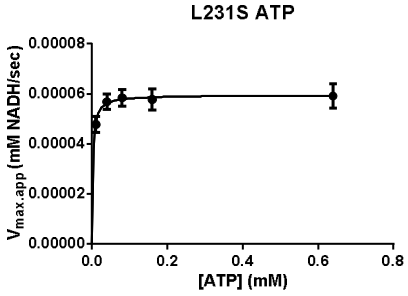
ACCEPTED MANUSCRIPT

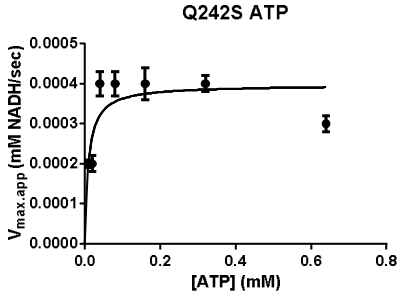


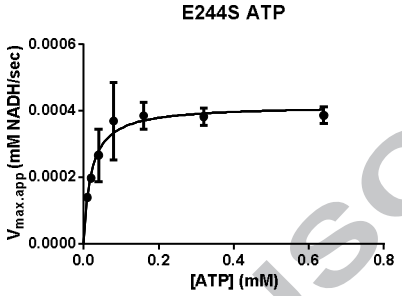


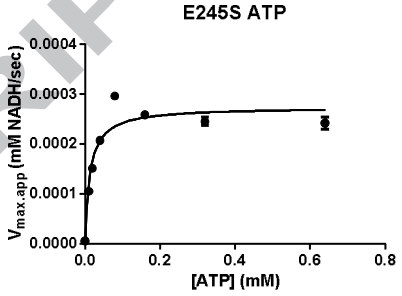






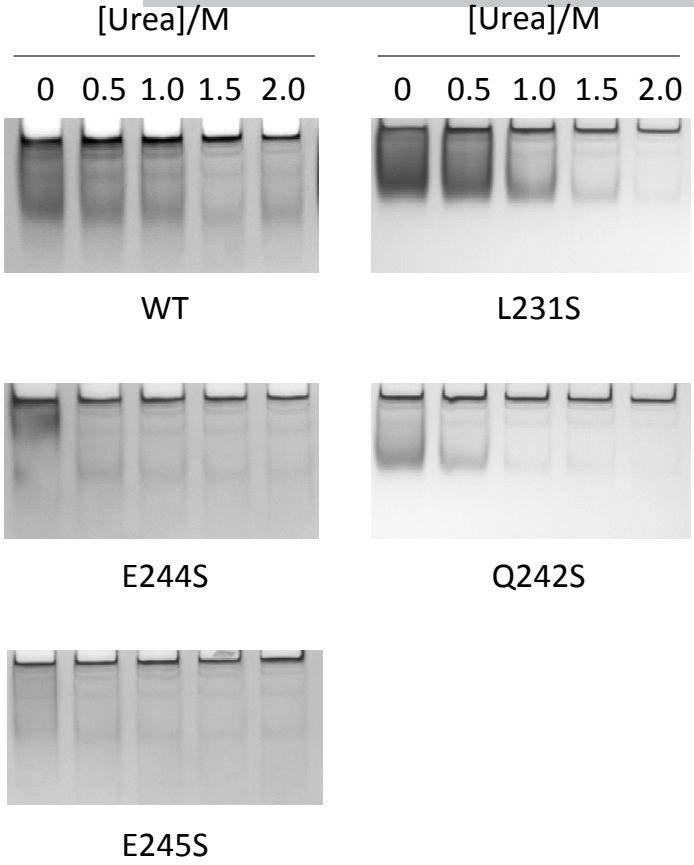






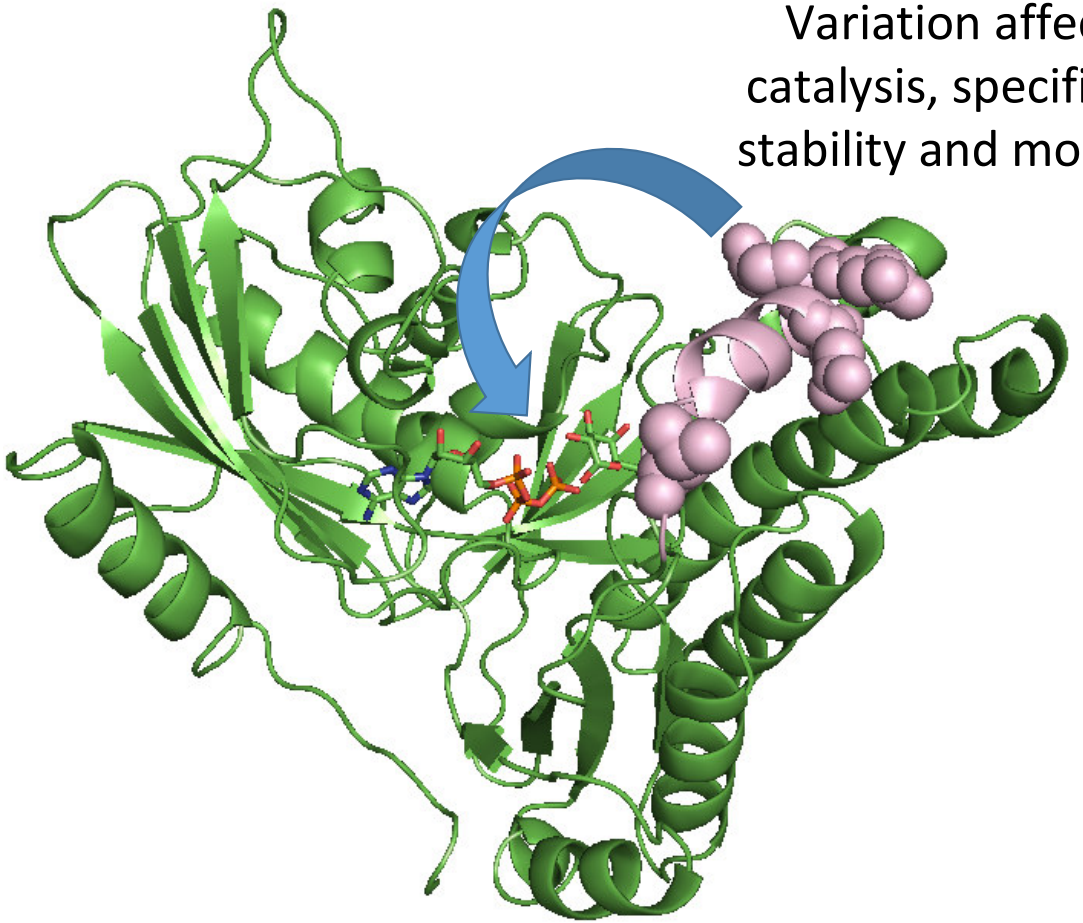
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



Galactose + ATP → Galactose 1-phosphate + ADP

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

- The specificity of galactokinase can be changed by altering active site mobility
- Residues distant from the active site were studied: Leu-231, Gln-242, Glu-244, Glu-245
- Variation of these residues is predicted to alter protein mobility
- These variations also affect specificity, stability and catalysis
- This principle could be extended to other enzymes of biotechnological interest

ACCEPTED MANUSCRIPT